Switch Dell N1524

Com alta taxa de transferência e capacidade para lidar com cargas de trabalho inesperadas, os switches da Série N1500 são soluções ideais para redes GbE de rápido crescimento que requerem agregação de alta densidade com redundância e disponibilidade contínuas.

Dell Networking N-Series N1500, N2000, N3000, and N4000 Switches

User's Configuration Guide

Version 6.3.0.0 and Later

Regulatory Models: N1524, N1524P, N1548, N1548P, N2024, N2024P, N2038, N2048P, N3024, N3024F, N3024P, N3048, N3048P, N4032, N4032F, N4064, N4064F

Notes and Cautions

NOTE: A NOTE indicates important information that helps you make better use of your computer.

CAUTION: A CAUTION indicates potential damage to hardware or loss of data if instructions are not followed.

____

Regulatory Models: N1524, N1524P, N1548, N1548P, N2024, N2024P, N2038, N2048P, N3024, N3024F, N3024P, N3048, N3048P, N4032, N4032F, N4064, N4064F

January 2016 Rev. A02

Contents

1

Introduction

The switches in the Dell Networking N1500, N2000, N3000, and N4000 Series switches Series are stackable layer-2 and layer-3 switches. These switches include the following features:

- 1U form factor, rack-mountable chassis design.
- Support for all data-communication requirements for a multi-layer switch, including layer-2 switching, IPv4 routing, IPv6 routing, IP multicast, quality of service, security, and system management features.
- High availability with hot swappable stack members.

The Dell Networking N-Series includes 17 switch models: N1524, N1524P, N1548, N1548P, N2024, N2024P, N2038, N2048P, N3024, N3024F, N3024P, N3048, N3048P, N4032, N4032F, N4064, N4064F.

NOTE: Switch administrators are strongly advised to maintain Dell Networking N-Series switches on the latest version of the Dell Networking Operating System. Dell Networking continually improves the features and functions based on feedback from you, the customer. For critical infrastructure, prestaging of the new release into a non-critical portion of the network is recommended to verify network configuration and operation with any new version of Dell Networking N-Series switch firmware.

About This Document

This guide describes how to configure, monitor, and maintain Dell Networking N1500, N2000, N3000, and N4000 Series switches Series switches by using web-based Dell OpenManage Switch Administrator utility or the command-line interface (CLI).

Audience

This guide is for network administrators in charge of managing one or more Dell Networking N-Series switches. To obtain the greatest benefit from this guide, you should have a basic understanding of Ethernet networks and local area network (LAN) concepts.

Document Conventions

Table 1-1 describes the typographical conventions this document uses.

Table 1-1. Document Conventions

Convention	Description
Bold	Page names, field names, menu options, button names, and CLI commands and keywords.
`courier font`	Command-line text (CLI output) and file names
[ ]	In a command line, square brackets indicate an optional entry.
{ }	In a command line, inclusive brackets indicate a selection of compulsory parameters separated by the | character. One option must be selected. For example: **spanning-tree mode** {**stp**|**rstp**|**mstp**} means that for the **spanning-tree mode** command, **stp**, **rstp**, or **mstp** must be entered.
Italic	In a command line, indicates a variable.
<Enter>	Any individual key on the keyboard.
CTRL + Z	A keyboard combination that involves pressing the Z key while holding the CTRL key.

Additional Documentation

The following documents for the Dell Networking N-Series switches are available at www.dell.com/support:

- *Getting Started Guide*—provides information about the switch models in the series, including front and back panel features. It also describes the installation and initial configuration procedures.
- *CLI Reference Guide*—provides information about the command-line interface (CLI) commands used to configure and manage the switch. The document provides in-depth CLI descriptions, syntax, default values, and usage guidelines.

2

Switch Feature Overview

This section describes the switch user-configurable software features.

NOTE: Before proceeding, read the release notes for this product. The release notes are part of the firmware download.

The topics covered in this section include:

- System Management Features
- Stacking Features
- Security Features
- Green Technology Features
- Power over Ethernet (PoE) Plus Features
- Switching Features
- Virtual Local Area Network Supported Features
- Spanning Tree Protocol Features
- Link Aggregation Features
- Routing Features
- IPv6 Routing Features
- Quality of Service (QoS) Features
- Layer-2 Multicast Features
- Layer-3 Multicast Features

System Management Features

Multiple Management Options

Any of the following methods can be used to manage the switch:

- Use a web browser to access the Dell OpenManage Switch Administrator interface. The switch contains an embedded Web server that serves HTML pages. Dell Networking N-Series switches support HTTP and HTTPS over IPv4 or IPv6.
- Use a Telnet client, SSH client, or a direct console connection to access the CLI. The CLI syntax and semantics conform as much as possible to common industry practice. Dell Networking N-Series switches support Telnet and SSH access over IPv4 or IPv6.
- Use a network management system (NMS), like the Dell OpenManage Network Manager, to manage and monitor the system through SNMP. The switch supports SNMP v1/v2c/v3 over the UDP/IP transport protocol.

Nearly all switch features support a preconfiguration capability, even when the feature is not enabled or the required hardware is not present. pre-configured capabilities become active only when enabled (typically via an admin mode control) or when the required hardware is present (or both). For example, a port can be pre-configured with both trunk and access mode information. The trunk mode information is applied only when the port is placed into trunk mode and the access mode information is only applied when the port is placed into access mode. Likewise, OSPF routing can be configured in the switch without being enabled on any port. This capability is present in all of the management options.

System Time Management

The switch can be configured to obtain the system time and date through a remote Simple Network Time Protocol (SNTP) server, or the time and date can be set locally on the switch. The time zone and information about time shifts that might occur during summer months can also be configured. When SNTP is used to obtain the time, communications between the switch and the SNTP server can be encrypted.

The Dell Networking SNTP client supports connection to SNTP servers over IPv4 or IPv6.

For information about configuring system time settings, see "Managing General System Settings " on page 389.

Log Messages

The switch maintains in-memory log messages as well as persistent logs. Remote logging can be configured so that the switch sends log messages to a remote syslog server. The switch can also be configured to email log messages to a configured SMTP server. This allows the administrator to receive the log message in a specified e-mail account. Switch auditing messages, CLI command logging, Web logging, and SNMP logging can be enabled or disabled.

Dell Networking N-Series switches support logging to syslog servers over IPv4 or IPv6.

For information about configuring system logging, see "Monitoring and Logging System Information " on page 353.

Integrated DHCP Server

NOTE: This feature is not supported on the Dell Networking N1500 Series switches.

Dell Networking N-Series switches include an integrated DHCP server that can deliver host-specific configuration information to hosts on the network. The switch DHCP server allows the configuration of IPv4 address pools (scopes), and when a host’s DHCP client requests an address, the switch DHCP server automatically assigns the host an address from the pool.

For information about configuring the DHCP server settings, see "DHCP Server and Relay Settings " on page 1089.

Management of Basic Network Information

The DHCP client on the switch allows the switch to acquire information such as the IPv4 or IPv6 address and default gateway from a network DHCP server. The DHCP client can also be disabled and static network information can be configured instead. Other configurable network information includes a Domain Name Server (DNS), hostname to IP address mapping, and a default domain name.

If the switch detects an IP address conflict on the management interface, it generates a trap and sends a log message.

For information about configuring basic network information, see "Setting the IP Address and Other Basic Network Information " on page 165.

IPv6 Management Features

Dell Networking N-Series switches provide IPv6 support for many standard management features including HTTP, HTTPS/SSL, Telnet, SSH, syslog, SNTP, TFTP, and traceroute on both the in-band and out-of-band management ports.

Dual Software Images

Dell Networking N-Series switches can store up to two software images. The dual image feature enables upgrading the switch without deleting the older software image. One image is designated as the active image and the other image as the backup image.

For information about managing the switch image, see "Images and File Management " on page 469.

File Management

Files, such as configuration files and system images, can be uploaded and downloaded using HTTP (web only), TFTP, Secure FTP (SFTP), or Secure Copy (SCP). Configuration file uploads from the switch to a server are a good way to back up the switch configuration. A configuration file can also be downloaded from a server to the switch to restore the switch to the configuration in the downloaded file.

Files can be copied to and from a USB Flash drive that is plugged into the USB port on the front panel of the switch. Or, the switch can be automatically upgraded by booting it with a newer firmware image on a USB drive plugged in to the switch. Dell Networking N-Series switches support file copy protocols to both IPv4 and IPv6 servers.

For information about uploading, downloading, and copying files, see "Images and File Management " on page 469.

Switch Database Management Templates

Switch Database Management (SDM) templates enable reallocating system resources to support a different mix of features based on network requirements. Dell Networking N-Series switches support the following three templates:

- Dual IPv4 and IPv6 (default)
- IPv4 Routing
- IPv4 Data Center

For information about setting the SDM template, see "Managing General System Settings " on page 389.

Automatic Installation of Firmware and Configuration

The Auto Install feature allows the switch to upgrade or downgrade to a newer software image and update the configuration file automatically during device initialization with limited administrative configuration on the device. If a USB device is connected to the switch and contains a firmware image and/or configuration file, the Auto Install feature installs the image or configuration file from USB device. Otherwise, the switch can obtain the necessary information from a DHCP server on the network.

NOTE: Automatic migration of the startup configuration to the next version of firmware from the current and previous versions of firmware is supported; the syntax is automatically updated when it is read into the running-config. Check the release notes to determine if any parts of the configuration cannot be migrated. Save the running-config to maintain the updated syntax. Migration of configuration is not assured on a firmware downgrade. When upgrading or downgrading firmware, check the configuration to ensure that it implements the desired configuration. Meta-configuration data (stack-port and slot configuration) is always reset to the defaults on a downgrade on each stack unit. As an example, Ethernet ports configured as stacking ports default back to Ethernet mode on a downgrade.

Migration of configuration information is never assured when errors are shown while the system is booting. Although the errored lines are displayed, commands that enter a sub-configuration mode followed by an exit command cause the CLI to exit Global Configuration mode, and subsequent configuration commands are ignored. Always hand-edit the startup-config if errors are shown on the screen during bootup.

For information about Auto Install, see "DHCP and USB Auto-Configuration " on page 499.

sFlow

sFlow is the standard for monitoring high-speed switched and routed networks. sFlow technology is built into network equipment and gives complete visibility into network activity, enabling effective management and control of network resources. The Dell Networking N-Series switches support sFlow version 5.

For information about configuring managing sFlow settings, see "Monitoring Switch Traffic " on page 519.

SNMP Alarms and Trap Logs

The system logs events with severity codes and timestamps. The events are sent as SNMP traps to a trap recipient list.

For information about configuring SNMP traps and alarms, see "SNMP " on page 433.

CDP Interoperability Through ISDP

Industry Standard Discovery Protocol (ISDP) allows the Dell Networking N-Series switch to interoperate with Cisco devices running the Cisco Discovery Protocol (CDP). ISDP is a proprietary layer-2 network protocol which inter-operates with Cisco network equipment and is used to share information between neighboring devices (routers, bridges, access servers, and switches).

For information about configuring ISDP settings, see "Discovering Network Devices " on page 825.

Remote Monitoring (RMON)

RMON is a standard Management Information Base (MIB) that defines current and historical MAC-layer statistics and control objects, allowing real-time information to be captured across the entire network.

For information about configuring managing RMON settings, see "Monitoring Switch Traffic " on page 519.

N3000 Series Access and Aggregation Firmware Images

There are two N3000 switch firmware images available. The Access Router image runs on both N3000 and N2000 switches. The image file is named N3000_N2000v*A.B.C.D*.stk, where *A*, *B*, *C*, and *D* are the version numbers. This image supports all switch capabilities except for BGP. The *D* (build number) in the image file name is enumerated starting with 1. N3000 series switches are loaded with the Access Router image in the factory.

The Aggregation Router image runs on N3000 model switches and is named N3000_BGPv*A.B.C.D*.stk. The *D* (build number) in the name is enumerated starting with 51. The Aggregation Router image supports MP-BGP and all other switch features except for MLAG, MVR, Auto-VoIP, GARP, GVRP, GMRP, iSCSI, and DVLAN (QinQ).

Which image type is installed can be determined by examining the first few lines of the running-config. The following example shows an Aggregation Router (BGP-enabled) firmware.

```
console#show running-config
!Current Configuration:
!System Description "Dell Networking N3048, 6.3.0.51, Linux 3.6.5-
ccbd1338"
```

```
!System Software Version 6.3.0.51
!Image File Name N3000_BGPv6.3.0.51.stk
!Software Capability AGGREGATION ROUTER
```

When migrating between the two types of images, certain commands in the startup-config may fail to execute because the relevant feature is not available. The switch firmware will identify any failed commands. It is necessary to edit the startup-config if errors are displayed and remove any failed commands. Do not simply save the running-config when commands in the startup-config fail, as the startup-config may contain modal commands that enter into a sub-mode not supported by the firmware. The exit command to exit the sub-mode may, in fact, exit Global Configuration mode, causing all subsequent commands to fail, even though those commands may be valid.

Stacking Features

For information about creating and maintaining a stack of switches, see "Stacking " on page 193.

High Stack Count

The Dell Networking N2000, N3000, and N4000 Series switches include a stacking feature that allows up to 12 switches to operate as a single unit. The Dell Networking N1500 switches allows stack configuration up to 4 units. The Dell Networking N2000 and N3000 Series switches have two fixed mini-SAS stacking connectors at the rear. Dell Networking N2000 Series switches will stack with other Dell Networking N2000 Series switches and Dell Networking N3000 Series switches stack with other Dell Networking N3000 Series switches.

Dell Networking N1500 Series switches stack with other Dell Networking N1500 Series switches over front-panel ports configured for stacking. Dell Networking N4000 Series switches stack with other Dell Networking N4000 Series switches over front-panel ports configured for stacking.

Single IP Management

When multiple switches are connected together through the stack ports, they operate as a single unit with a larger port count. The stack operates and is managed as a single entity. One switch acts as the master, and the entire stack is managed through the management interface (Web, CLI, or SNMP) of the stack master.

Master Failover with Transparent Transition

The stacking feature supports a *standby* or backup unit that assumes the stack master role if the stack master fails. As soon as a stack master failure is detected, the standby unit initializes the control plane and enables all other stack units with the current configuration. The standby unit maintains a synchronized copy of the running configuration for the stack.

Nonstop Forwarding on the Stack

The Nonstop Forwarding (NSF) feature allows the forwarding plane of stack units to continue to forward packets while the control and management planes restart as a result of a power failure, hardware failure, or software fault on the stack master and allows the standby switch to quickly takeover as the master.

Hot Add/Delete and Firmware Synchronization

Units can be added to and deleted from the stack without cycling the power on the stack. Units to be added to the stack must be powered off prior to cabling into the stack to avoid election of a new master unit and a possible downgrade of the stack. When the newly added unit is powered on, the Stack Firmware Synchronization feature, if enabled, automatically synchronizes the firmware version with the version running on the stack master. The synchronization operation may result in either an upgrade or a downgrade of firmware on the mismatched stack member. Once the firmware is synchronized on a member unit, the running-config on the member is updated to match the master switch. The startup-config on the standby and member switches is not updated to match the master switch due to configuration changes on the master switch. Saving the startup config on the master switch also saves it to the startup config on all the other stack members. The hardware configuration of every switch is updated to match the master switch (unit number, slot configuration, stack member number, etc.).

NOTE: ALWAYS POWER OFF a unit to be added to a stack prior to cabling it into the stack. Newly added units must be powered on one-at-a-time beginning with the unit directly connected to an already powered on stack member.

Security Features

Configurable Access and Authentication Profiles

Rules can be configured to limit access to the switch management interface based on criteria such as access type and source IP address of the management host. The user can also be required to be authenticated locally or by an external server, such as a RADIUS server.

For information about configuring access and authentication profiles, see "Authentication, Authorization, and Accounting " on page 229.

Password-Protected Management Access

Access to the Web, CLI, and SNMP management interfaces is password protected, and there are no default users on the system.

For information about configuring local user accounts, see "Authentication, Authorization, and Accounting " on page 229.

Strong Password Enforcement

The Strong Password feature enforces a baseline password strength for all locally administered users. Password strength is a measure of the effectiveness of a password in resisting guessing and brute-force attacks. The strength of a password is a function of length, complexity and randomness. Using strong passwords lowers overall risk of a security breach.

For information about configuring password settings, see "Authentication, Authorization, and Accounting " on page 229.

TACACS+ Client

The switch has a TACACS+ client. TACACS+ provides centralized security for validation of users accessing the switch. TACACS+ provides a centralized user management system while still retaining consistency with RADIUS and other authentication processes.

For information about configuring TACACS+ client settings, see "Authentication, Authorization, and Accounting " on page 229.

RADIUS Support

The switch has a Remote Authentication Dial In User Service (RADIUS) client and can support up to 32 named authentication and accounting RADIUS servers. The switch also supports RADIUS Attribute 4, which is the configuration of a NAS-IP address. The switch can also be configured to accept RADIUS-assigned VLANs.

For information about configuring RADIUS client settings, see "Authentication, Authorization, and Accounting " on page 229.

SSH/SSL

The switch supports Secure Shell (SSH) for secure, remote connections to the CLI and Secure Sockets Layer (SSL) to increase security when accessing the web-based management interface. The SSH server can be enabled or disabled using the **ip ssh** command.

For information about configuring SSH and SSL settings, see "Authentication, Authorization, and Accounting " on page 229.

Inbound Telnet Control

By default, the switch allows access over Telnet. The administrator can enable or disable the Telnet server using the **ip telnet** command. Additionally, the Telnet port number is configurable using the same command.

For information about configuring inbound Telnet settings, see "Authentication, Authorization, and Accounting " on page 229.

Denial of Service

The switch supports configurable Denial of Service (DoS) attack protection for eight different types of attacks.

For information about configuring DoS settings, see "Port and System Security " on page 623.

Port Protection

A port may be put into the error-disabled state for any of the following reasons:

- BPDU Storm: By default, if Spanning Tree Protocol (STP) bridge protocol data units (BPDUs) are received at a rate of 15pps or greater for three consecutive seconds on a port, the port will be error-disabled. The threshold is not configurable.
- Broadcast, Multicast, Unicast Storm: If broadcast, unknown multicast, or unknown unicast packets are received at a rate greater than the configured limit and the configured action is to disable the port, the port will be error-disabled. Storm control is not enabled by default. See the **storm-control** commands for further information. A trap is issued for ports disabled by Storm Control.
- DHCP Rate Limit: If DHCP packets are received on a port at a rate that exceeds 15 pps, the port will be error-disabled. The threshold is configurable up to 300 pps for up to 15s long using the **ip dhcp snooping limit** command. DHCP snooping is disabled by default. The default protection limit is 15 pps. A trap is issued for interfaces disabled by DHCP Snooping.
- DoS: Interfaces on which a denial of service attack is detected are error-disabled. Refer to the **dos-control** command for configuration options.
- ARP Inspection: By default, if Dynamic ARP Inspection packets are received on a port at a rate that exceeds 15 pps for 1 second, the port will be error-disabled. The threshold is configurable up to 300 pps and the burst is configurable up to 15s long using the **ip arp inspection limit** command. A trap is issued for interfaces disabled by Dynamic ARP Inspection.
- SFP Mismatch: Insertion of an unsupported SFP transceiver will error-disable the interface. This behavior can be suppressed using the **service unsupported-transceiver** command.
- SFP+ transceivers: SFP+ transceivers are not compatible with SFP slots (N3024F front-panel ports). To avoid damage to SFP+ transceivers mistakenly inserted into SFP ports, the SFP port is error-disabled when an SFP+ transceiver is detected.
- UDLD: Interfaces on which unidirectional packet flow is detected are error-disabled.

- ICMP storms: Ports on which ICMP storms are detected are error-disabled. The rate limit and burst sizes are configurable separately for IPv4 and IPv6.
- PML: Interfaces on which the port security violation is configured to shut down the interface are error-disabled when a violation occurs.
- Loop Protect: Loop protection diagnostically disables ports on which a loop is detected. A log message may be issued when a port is disabled by Loop Protection.
- BPDU Guard: An interface that receives a BPDU with BPDU guard enabled is error-disabled. Use the **spanning-tree bpdu-protection** command to enable BPDU guard.

A port that is error-disabled may be returned to service using the **no shutdown** command. Alternatively, the operator may configure the auto recovery service to return the error disabled ports to service after a configurable period of time. Refer to the **errdisable recovery** command for more information.

Captive Portal

The Captive Portal feature blocks clients from accessing the network until user verification has been established. When a user attempts to connect to the network through the switch, the user is presented with a customized Web page that might contain username and password fields or the acceptable use policy. Users can be required to be authenticated by a local or remote RADIUS database before access is granted.

For information about configuring the Captive Portal features, see "Captive Portal " on page 313.

802.1X Authentication (IEEE 802.1X)

802.1X authentication enables the authentication of system users through a local internal server or an external server. Only authenticated and approved system users can transmit and receive frames over the port. Supplicants are authenticated using the Extensible Authentication Protocol (EAP). EAP-MD5 authentication with no privacy protocol is supported for switch-initiated (server-side) authentication to remote authentication servers. Local (IAS) authentication and MAB authentication support EAP-MD5 only. Encrypted communication with authentication servers is not supported;

however, the switch will transport encrypted packets, such as PEAP or EAP-TLS packets, between the supplicant and authentication server in support of mutual authentication and privacy.

For information about configuring IEEE 802.1X settings, see "Port and System Security " on page 623.

MAC-Based 802.1X Authentication

MAC-based authentication allows multiple supplicants connected to the same port to each authenticate individually. For example, a system attached to the port might be required to authenticate in order to gain access to the network, while a VoIP phone might not need to authenticate in order to send voice traffic through the port.

For information about configuring MAC-based 802.1X authentication, see "Port and System Security " on page 623.

802.1X Monitor Mode

Monitor mode can be enabled in conjunction with 802.1X authentication to allow network access even when the user fails to authenticate. The switch logs the results of the authentication process for diagnostic purposes. The main purpose of this mode is to help troubleshoot the configuration of a 802.1X authentication on the switch without affecting the network access to the users of the switch.

For information about enabling the 802.1X Monitor mode, see "Port and System Security " on page 623.

MAC-Based Port Security

The port security feature limits access on a port to users with specific MAC addresses. These addresses are manually defined or learned on that port. When a frame is seen on a locked port, and the frame source MAC address is not tied to that port, the protection mechanism is invoked.

For information about configuring MAC-based port security, see "Port and System Security " on page 623.

Access Control Lists (ACLs)

Access Control Lists (ACLs) can help to ensure network availability for legitimate users while blocking attempts to access the network by unauthorized users or to restrict legitimate users from accessing the network. ACLs may be used to provide traffic flow control, restrict contents of routing updates, decide which types of traffic are forwarded or blocked, and above all, provide some level of security for the network. The switch supports the following ACL types:

- IPv4 ACLs
- IPv6 ACLs
- MAC ACLs

For all ACL types, the ACL rule can be configured to filter traffic when a packet enters or exits the Ethernet port, LAG, or VLAN interface. ACLs work only on switched ports. They do not operate on the out-of-band port.

ACLs can be used to implement policy-based routing (PBR) to implement packet routing according to specific organizational policies.

For information about configuring ACLs and PBR, see "Access Control Lists " on page 629.

Time-Based ACLs

With the Time-based ACL feature, the administrator can define when an ACL is in effect and the amount of time it is in effect.

For information about configuring time-based ACLs, see "Access Control Lists " on page 629.

IP Source Guard (IPSG)

IP source guard (IPSG) is a security feature that filters IP packets based on the source ID. The source ID may either be source IP address or a source IP address source MAC address pair as found in the local DHCP snooping database. IPSG depends on DHCP Snooping to associate IP address with MAC addresses.

For information about configuring IPSG, see "Snooping and Inspecting Traffic " on page 943.

DHCP Snooping

DHCP Snooping is a security feature that monitors DHCP messages between a DHCP client and DHCP server. It filters harmful DHCP messages and builds a bindings database of (MAC address, IP address, VLAN ID, port) tuples that are specified as authorized. DHCP snooping can be enabled globally and on specific VLANs. Ports within the VLAN can be configured to be trusted or untrusted. DHCP servers must be reached through trusted ports.

For information about configuring DHCP Snooping, see "Snooping and Inspecting Traffic " on page 943.

Dynamic ARP Inspection

Dynamic ARP Inspection (DAI) is a security feature that rejects invalid and malicious ARP packets. The feature prevents a class of man-in-the-middle attacks, where an unfriendly station intercepts traffic for other stations by poisoning the ARP caches of its unsuspecting neighbors. The malicious station sends ARP requests or responses mapping another station's IP address to its own MAC address.

Dynamic ARP Inspection relies on DHCP Snooping.

For information about configuring DAI, see "Snooping and Inspecting Traffic " on page 943.

Protected Ports (Private VLAN Edge)

Private VLAN Edge (PVE) ports are a layer-2 security feature that provides port-based security between ports that are members of the same VLAN. It is an extension of the common VLAN. Traffic from protected ports is sent only to the uplink ports and cannot be sent to other ports within the VLAN.

For information about configuring IPSG, see "Port-Based Traffic Control " on page 849.

Green Technology Features

For information about configuring Green Technology features, see "Port Characteristics " on page 591.

Energy Detect Mode

When the Energy Detect mode is enabled and the port link is down, the PHY automatically goes down for short period of time and then wakes up periodically to check link pulses. This mode reduces power consumption on the port when no link partner is present. Energy Detect is proprietary and operates independently from EEE.

Energy Efficient Ethernet

The switch supports the IEEE 802.3az Energy Efficient Ethernet (EEE) Lower Power Idle Mode, which enables both the send and receive sides of the link to disable some functionality for power savings when the link is lightly loaded. EEE is standardized and operates independently of Energy Detect.

EEE and Energy Detect are supported on the Dell Networking N1500, N2000, and N3000 Series 1G copper ports. EEE and energy detect are supported on the Dell Networking N4000 Series 10G copper ports.

EEE and Energy Detect are disabled by default on the Dell Networking N2000 and N3000 Series copper ports. Energy Detect is enabled by default on the Dell Networking N4000 Series switches and cannot be disabled. EEE is disabled by default on the Dell Networking N4000 Series 10G copper ports.

Power Utilization Reporting

The switch displays the current power consumption of the power supply (or power supplies). This information is available from the management interface.

Power over Ethernet (PoE) Plus Features

NOTE: The Dell Networking N1524P/N1548P, N2024P/N2048P, and N3024P/N3048P switches support PoE Plus and UPOE on selected ports. The PoE Plus and UPOE features do not apply to the other models in the Dell Networking N2000/N3000/N4000 Series.

For information about configuring PoE Plus features, see "Managing General System Settings " on page 389.

Power Over Ethernet (PoE) Plus Configuration

The Dell Networking N1524P/N1548P, N2024P/N2048P and N3024P/N3048P switches support PoE Plus configuration for power threshold, power priority, SNMP traps, and PoE legacy device support. PoE can be administratively enabled or disabled on a per-port basis. Power can also be limited on a per-port basis.

PoE Plus Support

The Dell Networking N1524P/N1548P, N2024P/N2048P and N3024P/N3048P switches implement the PoE Plus specification (IEEE 802.1at), in addition to the IEEE 802.3AF specification). This allows power to be supplied to Class 4 PD devices that require power greater than 15.4 Watts. Each port is capable of delivering up to 34.2W of power. Real-time power supply status is also available on the switch as part of the PoE Plus implementation.

UPOE Support

The Dell Networking N3024P/N3048P switches implement 4-pair Universal Power over Ethernet (UPOE) on the first 12 ports. This allows power to be supplied to Class 5 powered devices that require power up to 60 watts. UPOE power must be configured manually. The N1424P/N1548P/N2024P/N2048P switches do not support UPOE. High-power mode must be enabled in addition to four-par forced mode for the switch to deliver power on all four pairs. Class D or better cabling is required for feeds in excess of 34.2 watts. Normally, CAT 5E cabling does meet this requirement.

Switching Features

Flow Control Support (IEEE 802.3x)

Flow control enables lower speed switches to communicate with higher speed switches by requesting that the higher speed switch refrain from sending packets for a limited period of time. Transmissions are temporarily halted to prevent buffer overflows.

For information about configuring flow control, see "Port-Based Traffic Control " on page 849.

Head of Line Blocking Prevention

Head of Line (HOL) blocking prevention prevents traffic delays and frame loss caused by traffic competing for the same egress port resources. HOL blocking queues packets, and the packets at the head of the queue are forwarded before packets at the end of the queue.

Alternate Store and Forward (ASF)

NOTE: This feature is available on the Dell Networking N4000 Series switches only.

The Alternate Store and Forward (ASF) feature reduces latency for large packets. When ASF is enabled, the memory management unit (MMU) can forward a packet to the egress port before it has been entirely received on the Cell Buffer Pool (CBP) memory.

AFS, which is also known as cut-through mode, is configurable through the command-line interface. For information about how to configure the AFS feature, see the *CLI Reference Guide* available at www.dell.com/support.

Jumbo Frames Support

Jumbo frames enable transporting data in fewer frames to ensure less overhead, lower processing time, and fewer interrupts.

For information about configuring the switch MTU, see "Port Characteristics " on page 591.

Auto-MDI/MDIX Support

The switch supports auto-detection between crossed and straight-through cables. Media-Dependent Interface (MDI) is the standard wiring for end stations, and the standard wiring for hubs and switches is known as Media-Dependent Interface with Crossover (MDIX). Auto-negotiation must be enabled for MDIX to detect the wiring configuration.

VLAN-Aware MAC-based Switching

Packets arriving from an unknown source address are sent to the CPU and added to the Hardware Table. Future packets addressed to or from this address are more efficiently forwarded.

Back Pressure Support

On half-duplex links, a receiver may prevent buffer overflows by jamming the link so that it is unavailable for additional traffic. On full-duplex links, a receiver may send a PAUSE frame indicating that the transmitter should cease transmission of frames for a specified period.

NOTE: Dell Networking N2000/N3000/N4000 Series switches do not support half-duplex operation.

When flow control is enabled, the Dell Networking N-Series switches will observe received PAUSE frames or jamming signals, but will not issue them when congested.

Auto-negotiation

Auto-negotiation allows the switch to advertise modes of operation. The auto-negotiation function provides the means to exchange information between two switches that share a point-to-point link segment and to automatically configure both switches to take maximum advantage of their transmission capabilities.

Dell Networking N-Series switches enhance auto-negotiation by providing configuration of port advertisement. Port advertisement allows the system administrator to configure the port speeds that are advertised.

For information about configuring auto-negotiation, see "Port Characteristics " on page 591.

Storm Control

When layer-2 frames are processed, broadcast, unknown unicast, and multicast frames are flooded to all ports on the relevant virtual local area network (VLAN). The flooding occupies bandwidth and loads all nodes connected on all ports. Storm control limits the amount of broadcast, unknown unicast, and multicast frames accepted and forwarded by the switch.

For information about configuring Broadcast Storm Control settings, see "Port-Based Traffic Control " on page 849.

Port Mirroring

Port mirroring mirrors network traffic by forwarding copies of incoming and outgoing packets from multiple source ports to a monitoring port. Source ports may be VLANs, Ethernet interfaces, port-channels, or the CPU port. The switch also supports flow-based mirroring, which allows copying certain types of traffic to a single destination port. This provides flexibility—instead of mirroring all ingress or egress traffic on a port the switch can mirror a subset of that traffic. The switch can be configured to mirror flows based on certain kinds of layer-2, layer-3, and layer-4 information.

Dell Networking N-Series switches support RSPAN destinations where traffic can be tunneled across the operational network. RSPAN does not support configuration of the CPU port as a source.

For information about configuring port mirroring, see "Monitoring Switch Traffic " on page 519.

Static and Dynamic MAC Address Tables

Static entries can be added to the switch's MAC address table and the aging time can be configured for entries in the dynamic MAC address table. Entries can also be searched in the dynamic table based on several different criteria.

For information about viewing and managing the MAC address table, see "MAC Addressing and Forwarding " on page 1083.

Link Layer Discovery Protocol (LLDP)

The IEEE 802.1AB defined standard, Link Layer Discovery Protocol (LLDP), allows the switch to advertise major capabilities and physical descriptions. This information can be used to help identify system topology and detect bad configurations on the LAN.

For information about configuring LLDP, settings see "Discovering Network Devices " on page 825.

Link Layer Discovery Protocol (LLDP) for Media Endpoint Devices

The Link Layer Discovery Protocol for Media Endpoint Devices (LLDP-MED) provides an extension to the LLDP standard for network configuration and policy, device location, and Power over Ethernet.

For information about configuring LLDP-MED, settings see "Discovering Network Devices " on page 825.

Connectivity Fault Management (IEEE 802.1ag)

NOTE: This feature is available on the Dell Networking N4000 Series switches only.

The Connectivity Fault Management (CFM) feature, also known as Dot1ag, supports Service Level Operations, Administration, and Management (OAM). CFM is the OAM Protocol provision for end-to-end service layer instance in carrier networks. The CFM feature provides mechanisms to help perform connectivity checks, fault detection, fault verification and isolation, and fault notification per service in a network domain.

For information about configuring IEEE 802.1ag settings, see "Connectivity Fault Management " on page 923.

Priority-based Flow Control (PFC)

NOTE: This feature is available on the Dell Networking N4000 Series switches only.

The Priority-based Flow Control feature allows the switch to pause or inhibit transmission of individual priorities within a single Ethernet link. By configuring PFC to pause a congested priority (priorities) independently, protocols that are highly loss sensitive can share the same link with traffic that has different loss tolerances. Priorities are differentiated by the priority field of the 802.1Q VLAN header. The Dell Networking N4000 Series switches support lossless transport of frames on up to two priority classes.

NOTE: An interface that is configured for PFC is automatically disabled for 802.3x flow control.

For information about configuring the PFC feature, see "Data Center Bridging Features " on page 1051.

Data Center Bridging Exchange (DBCx) Protocol

NOTE: This feature is available on the Dell Networking N4000 Series switches only.

The Data Center Bridging Exchange Protocol (DCBx) is used by DCB devices to exchange configuration information with directly connected peers. The protocol is also used to detect misconfiguration of the peer DCB devices and, optionally, for configuration of peer DCB devices. For information about configuring DCBx settings, see "Data Center Bridging Features " on page 1051. DCBx is a link-local protocol and operates only on individual links.

Enhanced Transmission Selection

NOTE: This feature is available on the Dell Networking N4000 Series switches only.

Enhanced Transmission Selection (ETS) allows the switch to allocate bandwidth to traffic classes and share unused bandwidth with lower-priority traffic classes while coexisting with strict-priority traffic classes. ETS is supported on the Dell Networking N4000 Series switches and can be configured manually or automatically using the auto configuration feature. For more information about ETS, see "Enhanced Transmission Selection " on page 1067.

Cisco Protocol Filtering

The Cisco Protocol Filtering feature (also known as Link Local Protocol Filtering) filters Cisco protocols that should not normally be relayed by a bridge. The group addresses of these Cisco protocols do not fall within the IEEE defined range of the 802.1D MAC Bridge Filtered MAC Group Addresses (01-80-C2-00-00-00 to 01-80-C2-00-00-0F).

For information about configuring LLPF, settings see "Port-Based Traffic Control " on page 849.

DHCP Layer-2 Relay

This feature permits layer-3 relay agent functionality in layer-2 switched networks. The switch supports layer-2 DHCP relay configuration on individual ports, link aggregation groups (LAGs) and VLANs.

For information about configuring layer-2 DHCP relay settings see "Layer-2 and Layer-3 Relay Features " on page 1157.

Virtual Local Area Network Supported Features

For information about configuring VLAN features see "VLANs " on page 701.

VLAN Support

VLANs are collections of switching ports that comprise a single broadcast domain. Packets are classified as belonging to a VLAN based on either the VLAN tag or a combination of the ingress port and packet contents. Packets sharing common attributes can be groups in the same VLAN. The Dell Networking N-Series switches are in full compliance with IEEE 802.1Q VLAN tagging.

Port-Based VLANs

Port-based VLANs classify incoming packets to VLANs based on their ingress port. When a port uses 802.1X port authentication, packets can be assigned to a VLAN based on the result of the 802.1X authentication a client uses when it accesses the switch. This feature is useful for assigning traffic to Guest VLANs or Voice VLANs.

IP Subnet-based VLAN

This feature allows incoming untagged packets to be assigned to a VLAN and traffic class based on the source IP address of the packet.

MAC-based VLAN

This feature allows incoming untagged packets to be assigned to a VLAN and traffic class based on the source MAC address of the packet.

IEEE 802.1v Protocol-Based VLANs

VLAN classification rules are defined on data-link layer (layer-2) protocol identification. Protocol-based VLANs are used for isolating layer-2 traffic.

GARP and GVRP Support

NOTE: GARP, GVRP, and GMRP are not available when running the AGGREGATION ROUTER image.

The switch supports the Generic Attribute Registration Protocol (GARP). GARP VLAN Registration Protocol (GVRP) relies on the services provided by GARP to provide IEEE 802.1Q-compliant VLAN pruning and dynamic VLAN creation on 802.1Q trunk ports. When GVRP is enabled, the switch registers and propagates VLAN membership on all ports that are part of the active spanning tree protocol topology.

For information about configuring GARP timers see "Layer-2 Multicast Features " on page 867.

Voice VLAN

The Voice VLAN feature enables switch ports to carry voice traffic with a configured priority. The priority level enables the separation of voice and data traffic transiting the switch. Voice VLAN is the preferred solution for enterprises wishing to deploy VoIP services in their network.

Guest VLAN

The Guest VLAN feature allows the administrator to provide service to unauthenticated users, i.e., users that are unable to support 802.1X authentication.

For information about configuring the Guest VLAN see "Port and System Security " on page 623.

Unauthorized VLAN

The Unauthorized VLAN feature allows the administrator to configure a VLAN for 802.1X-aware hosts that attempt authentication and fail.

Double VLANs

NOTE: DVLAN is not available on the N3000 running the AGREGATION ROUTER image.

The Double VLAN feature (IEEE 802.1QinQ) allows the use of a second tag on network traffic. The additional tag helps differentiate between customers in the Metropolitan Area Networks (MAN) while preserving individual customer's VLAN identification when they enter their own 802.1Q domain.

Spanning Tree Protocol Features

For information about configuring Spanning Tree Protocol features, see "Spanning Tree Protocol " on page 779.

Spanning Tree Protocol (STP)

Spanning Tree Protocol (IEEE 802.1D) is a standard requirement of layer-2 switches that allows bridges to automatically prevent and resolve layer-2 forwarding loops.

Spanning Tree Port Settings

The STP feature supports a variety of per-port settings including path cost, priority settings, Port Fast mode, STP Root Guard, Loop Guard, TCN Guard, and Auto Edge. These settings are also configurable per-LAG.

Rapid Spanning Tree

Rapid Spanning Tree Protocol (RSTP) detects and uses network topologies to enable faster spanning tree convergence after a topology change, without creating forwarding loops. The port settings supported by STP are also supported by RSTP.

Multiple Spanning Tree

Multiple Spanning Tree (MSTP) operation maps VLANs to spanning tree instances. Packets assigned to various VLANs are transmitted along different paths within MSTP Regions (MST Regions). Regions are one or more interconnected MSTP bridges with identical MSTP settings. The MSTP standard lets administrators assign VLAN traffic to unique paths.

The switch supports IEEE 802.1Q-2005, which corrects problems associated with the previous version, provides for faster transition-to-forwarding, and incorporates new features for a port (restricted role and restricted TCN).

Bridge Protocol Data Unit (BPDU) Guard

Spanning Tree BPDU Guard is used to disable the port in case a new device tries to enter the already existing topology of STP. Thus devices, which were originally not a part of STP, are not allowed to influence the STP topology.

BPDU Filtering

When spanning tree is disabled on a port, the BPDU Filtering feature allows BPDU packets received on that port to be dropped. Additionally, the BPDU Filtering feature prevents a port in Port Fast mode from sending and receiving BPDUs. A port in Port Fast mode is automatically placed in the forwarding state when the link is up to increase convergence time.

RSTP-PV and STP-PV

Dell Networking N-Series switches support both Rapid Spanning Tree Per VLAN (RSTP-PV) and Spanning Tree Per VLAN (STP-PV). RSTP-PV is the IEEE 802.1w (RSTP) standard implemented per VLAN. A single instance of rapid spanning tree (RSTP) runs on each configured VLAN. Each RSTP instance on a VLAN has a root switch. STP-PV is the IEEE 802.1s (STP) standard implemented per VLAN.

Link Aggregation Features

For information about configuring link aggregation (port-channel) features, see "Link Aggregation " on page 979.

Link Aggregation

Up to eight ports can combine to form a single Link Aggregation Group (LAG). This enables fault tolerance protection from physical link disruption, higher bandwidth connections and improved bandwidth granularity. LAGs are formed from similarly configured physical links; i.e., the speed, duplex, auto-negotiation, PFC configuration, DCBX configuration, etc., must be compatible on all member links.

Per IEEE 802.1AX, only links with the identical operational characteristics, such as speed and duplex setting, may be aggregated. Dell Networking N-Series switches aggregate links only if they have the same operational speed and duplex setting, as opposed to the configured speed and duplex setting. This allows operators to aggregate links that use auto-negotiation to set values for speed and duplex or to aggregate ports with SFP+ technology operating at a lower speed, e.g., 1G. Dissimilar ports will not become active in the LAG if their operational settings do not match those of the first member of the LAG.

In practice, some ports in a LAG may auto-negotiate a different operational speed than other ports depending on the far-end settings and any link impairments. Per the above, these ports will not become active members of the LAG. On a reboot or on flapping the LAG links, a lower-speed port may be the first port selected to be aggregated into the LAG. In this case, the higher-speed ports are not aggregated. Use the **lacp port-priority** command to select one or more primary links to lead the formation of the aggregation group.

While it is a requirement of a port-channel that the link members operate at the same duplex and speed settings, administrators should be aware that copper ports have larger latencies than fiber ports. If fiber and copper ports are aggregated together, packets sent over the fiber ports would arrive significantly sooner at the destination than packets sent over the copper ports. This can cause significant issues in the receiving host (e.g., a TCP receiver) as it would be required to buffer a potentially large number of out-of-order frames. Devices unable to buffer the requisite number of frames will show excessive frame discard. Configuring copper and fiber ports together in an aggregation group is not recommended.

Link Aggregate Control Protocol (LACP)

Link Aggregate Control Protocol (LACP) uses peer exchanges across links to determine, on an ongoing basis, the aggregation capability of various links, and continuously provides the maximum level of aggregation capability achievable between a given pair of systems. LACP automatically determines, configures, binds, and monitors the binding of ports to aggregators within the system.

Multi-Switch LAG (MLAG)

NOTE: This feature is not available on Dell Networking N1500 Series switches. It is also not available on N3000 Series switches running the AGGREGATION ROUTER image.

Dell Networking N-Series switches support the MLAG feature to extend the LAG bandwidth advantage across multiple Dell Networking N-Series switches connected to a LAG partner device. The LAG partner device is unaware that it is connected to two peer Dell Networking N-Series switches; instead, the two switches appear as a single switch to the partner. When using MLAG, all links can carry data traffic across a physically diverse topology and, in the case of a link or switch failure, traffic can continue to flow with minimal disruption.

Routing Features

Address Resolution Protocol (ARP) Table Management

Static ARP entries can created and many settings for the dynamic ARP table can be managed, such as age time for entries, retries, and cache size.

For information about managing the ARP table, see "IP Routing " on page 1115.

VLAN Routing

Dell Networking N-Series switches support VLAN routing. The software can also be configured to allow traffic on a VLAN to be treated as if the VLAN were a router port.

For information about configuring VLAN routing interfaces, see "Routing Interfaces " on page 1141.

IP Configuration

The switch IP configuration settings to allow the configuration of network information for VLAN routing interfaces such as IP address and subnet mask, and ICMP redirects. Global IP configuration settings for the switch allow enabling or disabling the generation of several types of ICMP messages and enabling or disabling the routing mode.

For information about managing global IP settings, see "IP Routing " on page 1115.

Open Shortest Path First (OSPF)

NOTE: This feature is not available on Dell Networking N1500 Series switches.

Open Shortest Path First (OSPF) is a dynamic routing protocol commonly used within medium-to-large enterprise networks. OSPF is an interior gateway protocol (IGP) that operates within a single autonomous system.

For information about configuring OSPF, see "OSPF and OSPFv3 " on page 1181.

Border Gateway Protocol (BGP)

NOTE: This feature is not available on Dell Networking N1500 and N2000 Series switches. It is also not available on N3000 Series switches running the ACCESS ROUTER image.

BGP is a protocol used for exchanging reachability information between autonomous systems. BGP uses a standardized decision process, which, when used in conjunction with network policies configured by the administrator, support a robust set of capabilities for managing the distribution of routing information.

Dell Networking supports BGP4 configured as an IGP or an EGP. As an IGP, configuration as a source or client route reflector is supported. Both IPv6 and IPv4 peering sessions are supported.

For more information about configuring BGP, see "BGP " on page 1321.

Virtual Routing and Forwarding (VRF)

NOTE: This feature is not available on Dell Networking N1500 and N2000 Series switches.

VRF allows multiple independent instances of the forwarding plane to exist simultaneously. This allows segmenting the network without incurring the costs of multiple routers. Each VRF instance operates as an independent VPN. The IP addresses assigned to each VPN may overlap. Static route leaking to and from the global instance is supported. VLANs associated with a VRF may not overlap with other VRF instances.

For more information about configuring VRFs, see "VRF " on page 1273.

BOOTP/DHCP Relay Agent

The switch BootP/DHCP Relay Agent feature relays BootP and DHCP messages between DHCP clients and DHCP servers that are located in different IP subnets.

For information about configuring the BootP/DHCP Relay agent, see "Layer-2 and Layer-3 Relay Features " on page 1157.

IP Helper and UDP Relay

The IP Helper and UDP Relay features provide the ability to relay various protocols to servers on a different subnet.

For information about configuring the IP helper and UDP relay features, see "Layer-2 and Layer-3 Relay Features " on page 1157.

Routing Information Protocol

Routing Information Protocol (RIP), like OSPF, is an IGP used within an autonomous Internet system. RIP is an IGP that is designed to work with moderate-size networks.

For information about configuring RIP, see "RIP " on page 1279.

Router Discovery

For each interface, the Router Discovery Protocol (RDP) can be configured to transmit router advertisements. These advertisements inform hosts on the local network about the presence of the router.

For information about configuring router discovery, see "IP Routing " on page 1115.

Routing Table

The routing table displays information about the routes that have been dynamically learned. Static and default routes and route preferences can be configured. A separate table shows the routes that have been manually configured.

For information about viewing the routing table, see "IP Routing " on page 1115.

Virtual Router Redundancy Protocol (VRRP)

NOTE: This feature is not available on Dell Networking N2000 Series switches.

VRRP provides hosts with redundant routers in the network topology without any need for the hosts to reconfigure or know that there are multiple routers. If the primary (master) router fails, a secondary router assumes control and continues to use the virtual router IP (VRIP) address.

VRRP Route Interface Tracking extends the capability of VRRP to allow tracking of specific route/interface IP states within the router that can alter the priority level of a virtual router for a VRRP group.

For information about configuring VRRP settings, see "VRRP " on page 1295.

Tunnel and Loopback Interfaces

NOTE: This feature is not available on Dell Networking N1500 and N2000 Series switches.

Dell Networking N-Series switches support the creation, deletion, and management of tunnel and loopback interfaces. Tunnel interfaces facilitate the transition of IPv4 networks to IPv6 networks. A loopback interface is always expected to be up, so a stable IP address can be configured to enable other network devices to contact or identify the switch.

For information about configuring tunnel and loopback interfaces, see "Routing Interfaces " on page 1141.

IPv6 Routing Features

NOTE: This feature is not available on Dell Networking N1500 and N2000 Series switches.

IPv6 Configuration

The switch supports IPv6, the next generation of the Internet Protocol. IPv6 can be globally enabled on the switch and settings such as the IPv6 hop limit and ICMPv6 rate limit error interval can be configured. The administrator can also control whether IPv6 is enabled on a specific interface. The switch supports the configuration of many per-interface IPv6 settings including the IPv6 prefix and prefix length.

For information about configuring general IPv6 routing settings, see "IPv6 Routing " on page 1397.

IPv6 Routes

Because IPv4 and IPv6 can coexist on a network, the router on such a network needs to forward both traffic types. Given this coexistence, each switch maintains a separate routing table for IPv6 routes. The switch can forward IPv4 and IPv6 traffic over the same set of interfaces.

For information about configuring IPv6 routes, see "IPv6 Routing " on page 1397.

OSPFv3

OSPFv3 provides a routing protocol for IPv6 networking. OSPFv3 is a new routing component based on the OSPF version 2 component. In dual-stack IPv6, both OSPF and OSPFv3 components can be configured and used.

For information about configuring OSPFv3, see "OSPF and OSPFv3 " on page 1181.

DHCPv6

DHCPv6 incorporates the notion of the “stateless” server, where DHCPv6 is not used for IP address assignment to a client, rather it only provides other networking information such as DNS, Network Time Protocol (NTP), and/or Session Initiation Protocol (SIP) information.

For information about configuring DHCPv6 settings, see "DHCPv6 Server and Relay Settings " on page 1425.

Quality of Service (QoS) Features

NOTE: Some features that can affect QoS, such as ACLs and Voice VLAN, are described in other sections within this chapter.

Differentiated Services (DiffServ)

The QoS Differentiated Services (DiffServ) feature allows traffic to be classified into streams and given certain QoS treatment in accordance with defined per-hop behaviors. Dell Networking N-Series switches support both IPv4 and IPv6 packet classification.

For information about configuring DiffServ, see "Differentiated Services " on page 1445.

Class Of Service (CoS)

The Class Of Service (CoS) queueing feature enables directly configuring certain aspects of switch queuing. This provides the desired QoS behavior for different types of network traffic when the complexities of DiffServ are not required. CoS queue characteristics, such as minimum guaranteed bandwidth and transmission rate shaping, are configurable at the queue (or port) level.

For information about configuring CoS, see "Class-of-Service " on page 1473.

Auto Voice over IP (VoIP)

This feature provides ease of use for the user in setting up VoIP for IP phones on a switch. This is accomplished by enabling a VoIP profile that a user can select on a per port basis.

For information about configuring Auto VoIP, see "Auto VoIP " on page 1503.

This capability is not available on the N3000 Series switches when running the AGGRAGATION ROUTER image.

Internet Small Computer System Interface (iSCSI) Optimization

NOTE: This feature is not available on Dell Networking N1500 Series switches. It is also not available on N3000 Series switches running the AGGREGATION ROUTER image.

The iSCSI Optimization feature helps network administrators track iSCSI traffic between iSCSI initiator and target systems. This is accomplished by monitoring, or snooping traffic to detect packets used by iSCSI stations in establishing iSCSI sessions and connections. Data from these exchanges may optionally be used to create classification rules to assign the traffic between the stations to a configured traffic class. This affects how the packets in the flow are queued and scheduled for egress on the destination port.

For information about configuring iSCSI settings, see "iSCSI Optimization " on page 573.

Layer-2 Multicast Features

For information about configuring layer-2 multicast features, see "Layer-2 Multicast Features " on page 867.

MAC Multicast Support

Multicast service is a limited broadcast service that supports one-to-many and many-to-many forwarding behavior. In the layer-2 multicast service, a single frame addressed to a specific multicast address is received and copies of the frame to be transmitted on each relevant port are forwarded.

IGMP Snooping

Internet Group Management Protocol (IGMP) Snooping is a feature that allows a switch to forward multicast traffic intelligently on the switch. Multicast traffic is traffic that is destined to a host group. Host groups are identified by the destination MAC address, i.e. the range 01:00:5e:00:00:00 to 01:00:5e:7f:ff:ff:ff for IPv4 multicast traffic or 33:33:xx:xx:xx:xx for IPv6 multicast traffic. Based on the IGMP query and report messages, the switch forwards traffic only to the ports that request the multicast traffic. This prevents the switch from broadcasting the traffic to all ports and possibly affecting network performance.

IGMP Snooping Querier

When Protocol Independent Multicast (PIM) and IGMP are enabled in a network with IP multicast routing, an IP multicast router acts as the IGMP querier. However, if it is desirable to keep the multicast network layer-2 switched only, the IGMP Snooping Querier can perform the query functions of a layer-3 multicast router.

MLD Snooping

In IPv4, layer-2 switches can use IGMP Snooping to limit the flooding of multicast traffic by dynamically configuring layer-2 interfaces so that multicast traffic is forwarded to only those interfaces associated with IP multicast address.

In IPv6, MLD snooping performs a similar function. With MLD snooping, IPv6 multicast data is selectively forwarded to a list of ports intended to receive the data (instead of being flooded to all of the ports in a VLAN). This list is constructed by snooping IPv6 multicast control packets.

Multicast VLAN Registration

NOTE: This capability is not available on N3000 Series switches running the AGGREGATION ROUTER image.

The Multicast VLAN Registration (MVR) protocol, like IGMP Snooping, allows a layer-2 switch to listen to IGMP frames and forward the multicast traffic only to the receivers that request it. Unlike IGMP Snooping, MVR allows the switch to forward multicast frames across different VLANs. MVR uses a dedicated VLAN, which is called the multicast VLAN, to forward multicast traffic over the layer-2 network to the various VLANs that have multicast receivers as members.

Layer-3 Multicast Features

For information about configuring layer-3 (L3) multicast features, see "IPv4 and IPv6 Multicast " on page 1509.

NOTE: This feature is not available on Dell Networking N1500 and N2000 Series switches.

Distance Vector Multicast Routing Protocol

Distance Vector Multicast Routing Protocol (DVMRP) exchanges probe packets with all DVMRP-enabled routers, establishing two way neighboring relationships and building a neighbor table. It exchanges report packets and creates a unicast topology table, which is used to build the multicast routing table. This multicast route table is then used to route the multicast packets.

Internet Group Management Protocol

The Internet Group Management Protocol (IGMP) is used by IPv4 systems (hosts and routers) to report their IP multicast group memberships to any neighboring multicast routers. Dell Networking N-Series switches perform the “multicast router part” of the IGMP protocol, which means it collects the membership information needed by the active multicast router.

IGMP Proxy

The IGMP Proxy feature allows the switch to act as a proxy for hosts by sending IGMP host messages on behalf of the hosts that the switch discovered through standard IGMP router interfaces.

Protocol Independent Multicast—Dense Mode

Protocol Independent Multicast (PIM) is a standard multicast routing protocol that provides scalable inter-domain multicast routing across the Internet, independent of the mechanisms provided by any particular unicast routing protocol. The Protocol Independent Multicast-Dense Mode (PIM-DM) protocol uses an existing Unicast routing table and a Join/Prune/Graft mechanism to build a tree. PIM-DM creates source-based shortest-path distribution trees, making use of reverse path forwarding (RPF).

Protocol Independent Multicast—Sparse Mode

Protocol Independent Multicast-Sparse Mode (PIM-SM) is used to efficiently route multicast traffic to multicast groups that may span wide area networks, and where bandwidth is a constraint. PIM-SM uses shared trees by default and implements source-based trees for efficiency. This data threshold rate is used to toggle between trees.

Protocol Independent Multicast—Source Specific Multicast

Protocol Independent Multicast—Source Specific Multicast (PIM-SSM) is a subset of PIM-SM and is used for one-to-many multicast routing applications, such as audio or video broadcasts. PIM-SSM does not use shared trees.

Protocol Independent Multicast IPv6 Support

PIM-DM and PIM-SM support IPv6 routes.

MLD/MLDv2 (RFC2710/RFC3810)

MLD is used by IPv6 systems (listeners and routers) to report their IP multicast addresses memberships to any neighboring multicast routers. The implementation of MLD v2 is backward compatible with MLD v1.

MLD protocol enables the IPv6 router to discover the presence of multicast listeners, the nodes that want to receive the multicast data packets, on its directly attached interfaces. The protocol specifically discovers which multicast addresses are of interest to its neighboring nodes and provides this information to the multicast routing protocol that make the decision on the flow of the multicast data packets.

3

Hardware Overview

This section provides an overview of the switch hardware. It is organized by product type:

- Dell Networking N1500 Series Switch Hardware
- Dell Networking N2000 Series Switch Hardware
- Dell Networking N3000 Series Switch Hardware
- Dell Networking N4000 Series Switch Hardware
- Switch MAC Addresses

Dell Networking N1500 Series Switch Hardware

This section contains information about device characteristics and modular hardware configurations for the Dell Networking N1500 Series switches.

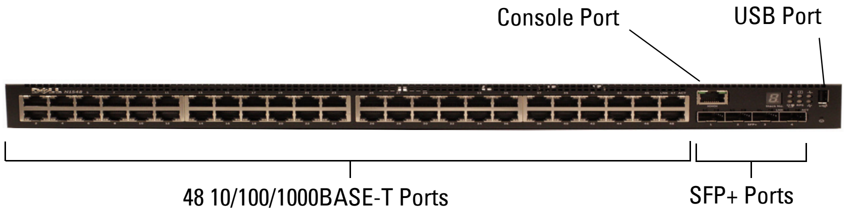
Front Panel

Then Dell Networking N1500 Series front panel includes the following features:

- Switch Ports
- Console Port
- USB Port
- Reset Button
- SFP+ Ports
- Port and System LEDs
- Stack Master LED and Stack Number Display

The following images show the front panels of the switch models in the Dell Networking N1500 Series.

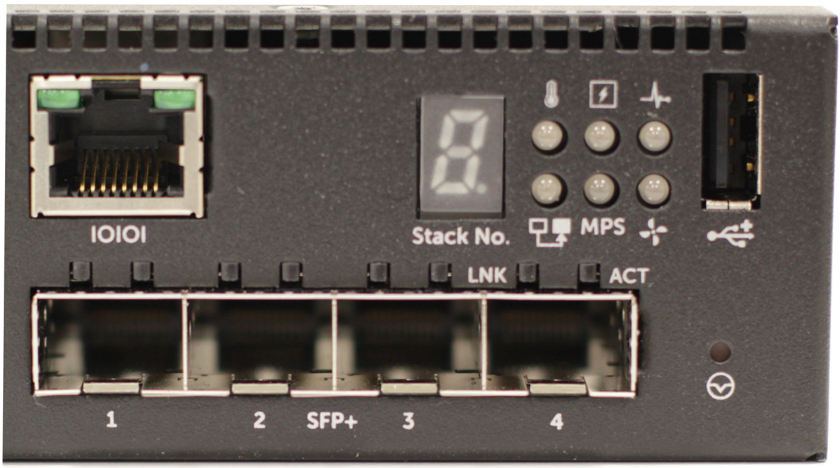
Figure 3-1. Dell Networking N1548 Front-Panel Switch with 48 10/100/1000BASE-T Ports (Front Panel)



In addition to the switch ports, the front panel of each model in the Dell Networking N1500 Series includes the following ports:

- RJ-45 Console port
- USB port for storage

Figure 3-2. Dell Networking N1524P Close-up



The Dell Networking 1524 front panel has status LEDs for over-temperature alarm (left), internal power (middle), and status (right) on the top row. The bottom row of status LEDs displays, from left to right, the Stack Master, redundant power supply (RPS) status, and fan alarm status.

The Dell Networking 1524P front panel, shown in Figure 3-2, has status LEDs for over-temperature alarm, internal power, and status on the top row. The bottom row of status LEDs displays Stack Master, modular power supply (MPS) status, and fan alarm status.

Switch Ports

The Dell Networking N1524/N1524P front panel provides 24 Gigabit Ethernet (10/100/1000BASE-T) RJ-45 ports that support auto-negotiation for speed, flow control, and duplex. The Dell Networking N1500 Series front-panel ports operate in full- or half-duplex mode. The Dell Networking N1524/N1524P models support four SFP+ 10G ports. Dell-qualified SFP+ transceivers are sold separately.

The Dell Networking N1548/N1548P front panel provides 48 Gigabit Ethernet (10BASE-T, 100BASE-TX, 1000BASE-T) RJ-45 ports that support auto-negotiation for speed, flow control, and duplex. The Dell Networking N1500 Series front-panel ports operate in full- or half-duplex mode. The Dell Networking N1548/N1548P supports four SFP+ 10G ports. Dell-qualified SFP+ transceivers are sold separately.

The front-panel switch ports have the following characteristics:

- The switch automatically detects the difference between crossed and straight-through cables on RJ-45 ports and automatically chooses the MDI or MDIX configuration to match the other end.
- SFP+ ports support Dell-qualified transceivers utilizing 10GBASE-SR, 10GBASE-LR, 10GBASE-CR, or 1000BASE-X technologies. The default behavior is to log a message and generate an SNMP trap on insertion or removal of an optic that is not qualified by Dell. The message and trap can be suppressed by using the **service unsupported-transceiver** command.
- RJ-45 front-panel ports support full- or half-duplex mode 10/100/1000 Mbps speeds on standard Category 5 UTP cable. 1000BASE-T operation requires the use of auto-negotiation.
- SFP+ ports support SFP+ transceivers and SFP+ copper twin-ax technology operating at 10G or 1G speeds in full-duplex mode. SFP transceivers are supported in SFP+ ports and operate at 1G full-duplex. SFP transceivers require auto-negotiation to be enabled.

  SFP+ post may be configured to support 16 GB stacking over Ethernet cables. These ports may be configured to support stacking in pairs, e.g., Te1/0/1 and Te1/0/2 may be configured to support stacking, or Te1/0/3 and Te1/0/4 may be configured to support stacking, or all four ports may be configured to support stacking.

- The Dell Networking N1524P/N1548P front-panel ports support PoE (15.4W) and PoE+ (34.2W) as well as legacy capacitive detection for pre-standard powered devices (PDs).

Console Port

The console port provides serial communication capabilities, which allows communication using the RS-232 protocol. The serial port provides a direct connection to the switch and allows access to the CLI from a console terminal connected to the port through the provided serial cable (with RJ45 YOST to female DB-9 connectors).

The console port is separately configurable and can be run as an asynchronous link from 1200 baud to 115,200 baud. The Dell CLI supports changing the speed only. The defaults are 9600 baud rate, 8 data bits, No Parity, 1 Stop Bit, No Flow Control.

USB Port

The Type-A, female USB port supports a USB 2.0-compliant flash memory drive. The Dell Networking N-Series switch can read or write to a flash drive with a single partition formatted as FAT-32. Use a USB flash drive to copy switch configuration files and images between the USB flash drive and the switch. The USB flash drive may be used to move and copy configuration files and images from one switch to other switches in the network. The system does not support the deletion of files on USB flash drives.

The USB port does not support any other type of USB device.

Reset Button

The reset button is accessed through the pinhole and enables performing a hard reset on the switch. To use the reset button, insert an unbent paper clip or similar tool into the pinhole. When the switch completes the boot process after the reset, it resumes operation with the most recently saved configuration. Any changes made to the running configuration that were not saved to the startup configuration prior to the reset are lost.

Port and System LEDs

The front panel contains light emitting diodes (LEDs) that indicate the status of port links, power supplies, fans, stacking, and the overall system status. See "LED Definitions " on page 104 for more information.

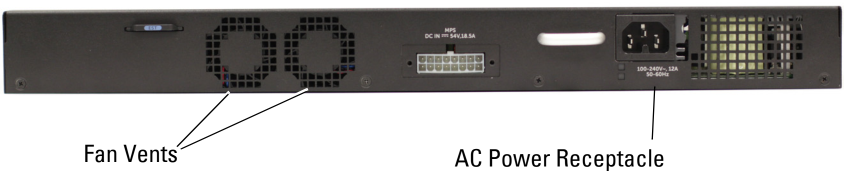
Stack Master LED and Stack Number Display

When a switch within a stack is the master unit, the Stack Master LED is solid green. If the Stack Master LED is off, the stack member is not the master unit. The Stack No. panel displays the unit number for the stack member. If a switch is not part of a stack (in other words, it is a stack of one switch), the Stack Master LED is illuminated, and the unit number is displayed.

Back Panel

The following images show the back panels of the Dell Networking N1500 Series switches.

Figure 3-3. Dell Networking N1500 Series Back Panel



Power Supplies

Dell Networking N1524 and N1548

The Dell Networking N1524 and N1548 Series switches have an internal 100-watt power supply. The additional redundant power supply (Dell Networking RPS720) provides 180 watts of power and gives full redundancy for the switch.

Dell Networking N1524P and N1548P

The Dell Networking N1524P and N1548P switches have an internal 600-watt power supply feeding up to 24 PoE devices at full PoE+ power (500W). An additional modular power supply (MPS1000) provides 1000 watts and gives full power coverage for all 48 PoE devices (1500W).

NOTE: PoE power is dynamically allocated. Not all ports will require the full PoE+ power.

CAUTION: Remove the power cable from the power supplies prior to removing the power supply module itself. Power must not be connected prior to insertion in the chassis.

Ventilation System

Two internal fans cool the Dell Networking N1500 Series switches.

Information Tag

The back panel includes a slide-out label panel that contains system information, such as the Service Tag, MAC address, and so on.

LED Definitions

This section describes the LEDs on the front and back panels of the switch.

Port LEDs

Each port on a Dell Networking N1500 Series switch includes two LEDs. One LED is on the left side of the port, and the second LED is on the right side of the port. This section describes the LEDs on the switch ports.

100/1000/10000Base-T Port LEDs

Each 100/1000/10000Base-T port has two LEDs. Figure 3-4 illustrates the 100/1000/10000Base-T port LEDs.

Figure 3-4. 100/1000/10000Base-T Port LEDs

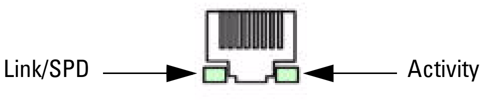


Table 3-1 shows the 100/1000/10000Base-T port LED definitions.

Table 3-1. 100/1000/10000Base-T Port Definitions

LED	Color	Definition
Link/SPD LED	Off	There is no link.
	Solid yellow	The port is operating at 10/100 Mbps.
	Solid green	The port is operating at 1000 Mbps.
Activity LED (on non-PoE switches)	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.
Activity/PoE LED (on PoE switches)	Off	There is no current transmit/receive activity and PoE power is off.
	Blinking green	The port is actively transmitting/receiving and PoE power is off.
	Blinking yellow	The port is actively transmitting/receiving and PoE power is on.
	Solid yellow	There is no current transmit/receive activity and PoE power is on.

Stacking Port LEDs

Table 3-2. Stacking Port LED Definitions

LED	Color	Definition
Link LED	Off	There is no link.
	Solid green	The port is actively transmitting/receiving.
Activity LED	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.

Table 3-3. Console Port LED Definitions

LED	Color	Definition
Link/SPD LED	Off	There is no link.
	Solid green	A link is present.

System LEDs

The system LEDs, located on the back panel, provide information about the power supplies, thermal conditions, and diagnostics.

Table 3-4 shows the System LED definitions for the Dell Networking N1500 Series switches.

Table 3-4. System LED Definitions

LED	Color	Definition
Status	Solid green	Normal operation.
	Blinking green	The switch is booting
	Solid red	A critical system error has occurred.
	Blinking red	A noncritical system error occurred (fan or power supply failure).
Power	Off	There is no power or the switch has experienced a power failure.
	Solid green	Power to the switch is on.
	Blinking green	The switch locator function is enabled.

Table 3-4. System LED Definitions (Continued)

LED	Color	Definition
RPS (on non-PoE switches)	Off	There is no redundant power supply (RPS).
	Solid green	Power to the RPS is on.
	Solid red	An RPS is detected but it is not receiving power.
EPS (on PoE switches)	Off	There is no external power supply (EPS).
	Solid green	Power to the EPS is on.
	Solid red	An EPS is detected but it is not receiving power.
Fan	Solid green	The fan is powered and is operating at the expected RPM.
	Solid red	A fan failure has occurred.
Stack Master	Off	The switch is not stack master.
	Solid green	The switch is master for the stack.
Temp	Solid green	The switch is operating below the threshold temperature.
	Solid red	The switch temperature exceeds the threshold of 75°C.
Stack No.	–	Switch ID within the stack.

Power Consumption for PoE Switches

Table 3-5 shows power consumption data for the PoE-enabled switches.

Table 3-5. Power Consumption

Model	Input Voltage	Power Supply Configuration	Max Steady Current Consumption (A)	Max Steady Power (W)
Dell Networking N1524P	100V	Main PSU+EPS PSU	8.8	876.0
	110V	Main PSU+EPS PSU	7.9	871.0
	120V	Main PSU+EPS PSU	7.2	865.0
	220V	Main PSU+EPS PSU	3.8	844.0
	240V	Main PSU+EPS PSU	3.5	840.0

Table 3-5. Power Consumption

Model	Input Voltage	Power Supply Configuration	Max Steady Current Consumption (A)	Max Steady Power (W)
Dell Networking N1548P	100V	Main PSU+EPS PSU	17.1	1719.0
	110V	Main PSU+EPS PSU	15.5	1704.0
	120V	Main PSU+EPS PSU	14.1	1690.0
	220V	Main PSU+EPS PSU	7.5	1642.4
	240V	Main PSU+EPS PSU	6.9	1647.0

The PoE power budget for each interface is controlled by the switch firmware. The administrator can limit the power supplied on a port or prioritize power to some ports over others. Table 3-6 shows power budget data.

Table 3-6. Dell Networking N1500 Series PoE Power Budget Limit

		Internal Only PSU		MPS Only		Two PSUs	
Model Name	**System Pwr. Max. Dissipation**	**Max. PSU Output Ability**	**PoE+ Power Turn-on Limitation**	**Max. PSU Output Ability**	**PoE+ Power Turn-on Limitation**	**Max. PSUs Output Ability**	**PoE+ Power Turn-on Limitation**
Dell Networking N1524P	40W	600W	Power budget is 500W: The total PoE supplied power must not exceed 500W.	1000W	Power budget is 900W: The total PoE supplied power must not exceed 900W.	1600W	Power budget is 1350W: All PoE+ ports can supply maximum power.
Dell Networking N1548P	62W	600W	Power budget is 500W: The total PoE supplied power must not exceed 500W.	1000W	Power budget is 900W: The total PoE supplied power must not exceed 900W.	1600W	Power budget is 1350W: The total PoE supplied power must not exceed 1350W.

Dell Networking N2000 Series Switch Hardware

This section contains information about device characteristics and modular hardware configurations for the Dell Networking N2000 Series switches.

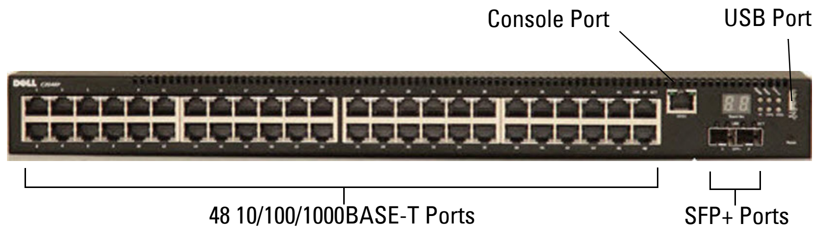
Front Panel

Then Dell Networking N2000 Series front panel includes the following features:

- Switch Ports
- Console Port
- USB Port
- Reset Button
- SFP+ Ports
- Port and System LEDs
- Stack Master LED and Stack Number Display

The following images show the front panels of the switch models in the Dell Networking N2000 Series.

Figure 3-5. Dell Networking N2048 Switch with 48 10/100/1000BASE-T Ports (Front Panel)



In addition to the switch ports, the front panel of each model in the Dell Networking N2000 Series includes the following ports:

- RJ-45 Console port
- USB port for storage

Figure 3-6. Dell Networking N2024 Close-up

The Dell Networking N2024 front panel, shown in Figure 3-6, has status LEDs for over-temperature alarm (left), internal power (middle), and status (right) on the top row. The bottom row of status LEDs displays, from left to right, the Stack Master, redundant power supply (RPS) status, and fan alarm status.

The Dell Networking N2024P front panel has status LEDs for over-temperature alarm, internal power and status on the top row. The bottom row of status LEDs displays Stack Master, modular power supply (MPS), status and fan alarm status.

Switch Ports

The Dell Networking N2024/N2024P front panel provides 24 Gigabit Ethernet (10/100/1000BASE-T) RJ-45 ports that support auto-negotiation for speed, flow control, and duplex. The Dell Networking N2024/N2024P models support two SFP+ 10G ports. Dell-qualified SFP+ transceivers are sold separately. Dell Networking N2000 Series switches operate in full-duplex mode only.

The Dell Networking N2048/N2048P front panel provides 48 Gigabit Ethernet (10BASE-T, 100BASE-TX, 1000BASE-T) RJ-45 ports that support auto-negotiation for speed, flow control, and duplex. The Dell Networking N2048/N2048P supports two SFP+ 10G ports. Dell-qualified SFP+ transceivers are sold separately.

The front-panel switch ports have the following characteristics:

- The switch automatically detects the difference between crossed and straight-through cables on RJ-45 ports and automatically chooses the MDI or MDIX configuration to match the other end.
- SFP+ ports support Dell-qualified transceivers. The default behavior is to log a message and generate an SNMP trap on insertion or removal of an optic that is not qualified by Dell. The message and trap can be suppressed by using the **service unsupported-transceiver** command.
- RJ-45 ports support full-duplex mode 10/100/1000 Mbps speeds on standard Category 5 UTP cable. 1000BASE-T operation requires the use of auto-negotiation.
- SFP+ ports support SFP+ transceivers and SFP+ copper twin-ax technology operating at 10G or 1G speeds in full-duplex mode. SFP transceivers are supported in SFP+ ports and operate at 1G full-duplex. SFP transceivers require auto-negotiation to be enabled.
- The Dell Networking N2024P/N2048P front-panel ports support PoE (15.4W) and PoE+ (34.2W) as well as legacy capacitive detection for pre-standard powered devices (PDs). Additionally, ports 1–12 support UPOE (60W) power when configured in high-power mode.

Console Port

The console port provides serial communication capabilities, which allows communication using RS-232 protocol. The serial port provides a direct connection to the switch and allows access to the CLI from a console terminal connected to the port through the provided serial cable (with RJ45 YOST to female DB-9 connectors).

The console port is separately configurable and can be run as an asynchronous link from 1200 baud to 115,200 baud.

The Dell CLI only supports changing the speed. The defaults are 9600 baud rate, 8 data bits, No Parity, 1 Stop Bit, No Flow Control.

USB Port

The Type-A, female USB port supports a USB 2.0-compliant flash memory drive. The Dell Networking N-Series switch can read or write to a flash drive with a single partition formatted as FAT-32. Use a USB flash drive to copy switch configuration files and images between the USB flash drive and the switch. The USB flash drive may be used to move and copy configuration files and images from one switch to other switches in the network. The system does not support the deletion of files on USB flash drives.

The USB port does not support any other type of USB device.

Reset Button

The reset button is accessed through the pinhole and enables performing a hard reset on the switch. To use the reset button, insert an unbent paper clip or similar tool into the pinhole. When the switch completes the boot process after the reset, it resumes operation with the most recently saved configuration. Any changes made to the running configuration that were not saved to the startup configuration prior to the reset are lost.

Port and System LEDs

The front panel contains light emitting diodes (LEDs) that indicate the status of port links, power supplies, fans, stacking, and the overall system status. See "LED Definitions " on page 114 for more information.

Stack Master LED and Stack Number Display

When a switch within a stack is the master unit, the Stack Master LED is solid green. If the Stack Master LED is off, the stack member is not the master unit. The Stack No. panel displays the unit number for the stack member. If a switch is not part of a stack (in other words, it is a stack of one switch), the Stack Master LED is illuminated, and the unit number is displayed.

Back Panel

The following images show the back panels of the Dell Networking N2000 Series switches.

Figure 3-7. Dell Networking N2000 Series Back Panel

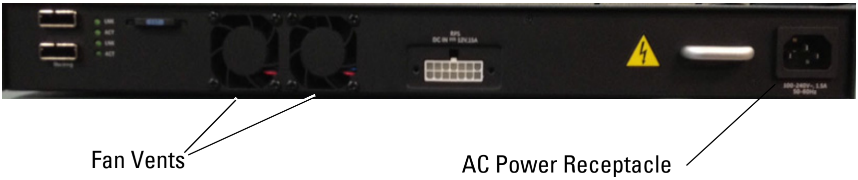
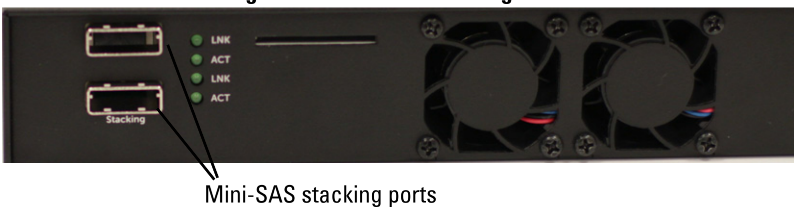


Figure 3-8. Dell Networking N2024P/N2048P Back Panel

The term mini-SAS refers to the stacking port cable connections shown in Figure 3-9. See "Stacking " on page 193 for information on using the mini-SAS ports to connect switches.

Figure 3-9. Dell Networking N2048 Mini-SAS Stacking Ports and Fans



Power Supplies

Dell Networking N2024 and N2048

The Dell Networking N2024 and N2048 Series switches have an internal 100-watt power supply. The additional redundant power supply (Dell Networking RPS720) provides 180 watts of power and gives full redundancy for the switch.

Dell Networking N2024P and N2048P

The Dell Networking N2024P and N2048P switches have an internal 1000-watt power supply feeding up to 24 PoE devices at full PoE+ power (850W). An additional modular power supply (MPS1000) provides 1000 watts and gives full power coverage for all 48 PoE devices (1800W).

NOTE: PoE power is dynamically allocated. Not all ports will require the full PoE+ power.

CAUTION: Remove the power cable from the power supplies prior to removing the power supply module itself. Power must not be connected prior to insertion in the chassis.

Ventilation System

Two internal fans cool the Dell Networking N2000 Series switches.

Information Tag

The back panel includes a slide-out label panel that contains system information, such as the Service Tag, MAC address, and so on.

LED Definitions

This section describes the LEDs on the front and back panels of the switch.

Port LEDs

Each port on a Dell Networking N2000 Series switch includes two LEDs. One LED is on the left side of the port, and the second LED is on the right side of the port. This section describes the LEDs on the switch ports.

100/1000/10000Base-T Port LEDs

Each 100/1000/10000Base-T port has two LEDs. Figure 3-10 illustrates the 100/1000/10000Base-T port LEDs.

Figure 3-10. 100/1000/10000Base-T Port LEDs

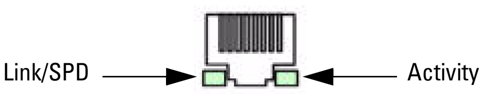


Table 3-7 shows the 100/1000/10000Base-T port LED definitions.

Table 3-7. 100/1000/10000Base-T Port Definitions

LED	Color	Definition
Link/SPD LED	Off	There is no link.
	Solid yellow	The port is operating at 10/100 Mbps.
	Solid green	The port is operating at 1000 Mbps.
Activity LED (on non-PoE switches)	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.
Activity/PoE LED (on PoE switches)	Off	There is no current transmit/receive activity and PoE power is off.
	Blinking green	The port is actively transmitting/receiving and PoE power is off.
	Blinking yellow	The port is actively transmitting/receiving and PoE power is on.
	Solid yellow	There is no current transmit/receive activity and PoE power is on.

Stacking Port LEDs

Table 3-8. Stacking Port LED Definitions

LED	Color	Definition
Link LED	Off	There is no link.
	Solid green	The port is actively transmitting/receiving.
Activity LED	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.

Table 3-9. Console Port LED Definitions

LED	Color	Definition
Link/SPD LED	Off	There is no link.
	Solid green	A link is present.

System LEDs

The system LEDs, located on the back panel, provide information about the power supplies, thermal conditions, and diagnostics.

Table 3-10 shows the System LED definitions for the Dell Networking N2000 Series switches.

Table 3-10. System LED Definitions

LED	Color	Definition
Status	Solid green	Normal operation.
	Blinking green	The switch is booting
	Solid red	A critical system error has occurred.
	Blinking red	A noncritical system error occurred (fan or power supply failure).
Power	Off	There is no power or the switch has experienced a power failure.
	Solid green	Power to the switch is on.
	Blinking green	The switch locator function is enabled.
RPS (on non-PoE switches)	Off	There is no redundant power supply (RPS).
	Solid green	Power to the RPS is on.
	Solid red	An RPS is detected but it is not receiving power.
EPS (on PoE switches)	Off	There is no external power supply (EPS).
	Solid green	Power to the EPS is on.
	Solid red	An EPS is detected but it is not receiving power.
Fan	Solid green	The fan is powered and is operating at the expected RPM.
	Solid red	A fan failure has occurred.

Table 3-10. System LED Definitions (Continued)

LED	Color	Definition
Stack Master	Off	The switch is not stack master.
	Solid green	The switch is master for the stack.
Temp	Solid green	The switch is operating below the threshold temperature.
	Solid red	The switch temperature exceeds the threshold of 75°C.
Stack No.	–	Switch ID within the stack.

Power Consumption for PoE Switches

Table 3-11 shows power consumption data for the PoE-enabled switches.

Table 3-11. Power Consumption

Model	Input Voltage	Power Supply Configuration	Max Steady Current Consumption (A)	Max Steady Power (W)
Dell Networking N2024P	100V	Main PSU+EPS PSU	8.9	890.0
	110V	Main PSU+EPS PSU	8.3	913.0
	120V	Main PSU+EPS PSU	7.6	912.0
	220V	Main PSU+EPS PSU	4.0	880.0
	240V	Main PSU+EPS PSU	3.6	873.6
Dell Networking N2048P	100V	Main PSU+EPS PSU	17.8	1780.0
	110V	Main PSU+EPS PSU	15.8	1740.2
	120V	Main PSU+EPS PSU	14.5	1740.0
	220V	Main PSU+EPS PSU	7.7	1687.4
	240V	Main PSU+EPS PSU	7.1	1704.0

The PoE power budget for each interface is controlled by the switch firmware. The administrator can limit the power supplied on a port or prioritize power to some ports over others. Table 3-12 shows power budget data.

Table 3-12. Dell Networking N2000 Series PoE Power Budget Limit

		One PSU		Two PSUs	
Model Name	**System Power Max. Dissipation**	**Max. PSU Output Ability**	**PoE+ Power Turn-on Limitation**	**Max. PSUs Output Ability**	**PoE+ Power Turn-on Limitation**
Dell Networking N2024P	90W	1000W	Power budget is 850W: The total PoE supplied power must not exceed 850W.	2000W	Power budget is 1700W: All PoE+ ports can supply maximum power.
Dell Networking N2048P	110W	1000W	Power budget is 850W: The total PoE supplied power must not exceed 850W.	2000W	Power budget is 1700W: All PoE+ ports can supply maximum power.

Dell Networking N3000 Series Switch Hardware

This section contains information about device characteristics and modular hardware configurations for the Dell Networking N3000 Series switches.

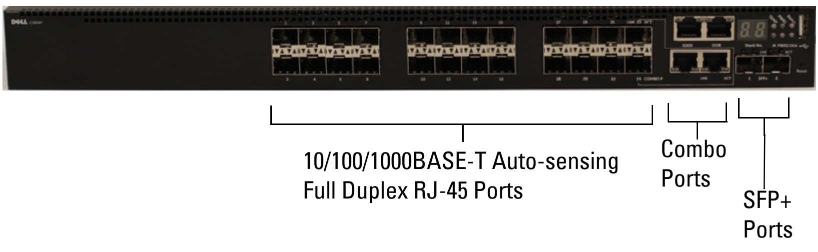
Front Panel

The Dell Networking N3000 Series front panel includes the following features:

- Switch Ports
- Console Port
- Out-of-Band Management Port
- USB Port
- SFP+ Ports
- Reset Button
- Port and System LEDs
- Stack Master LED and Stack Number Display

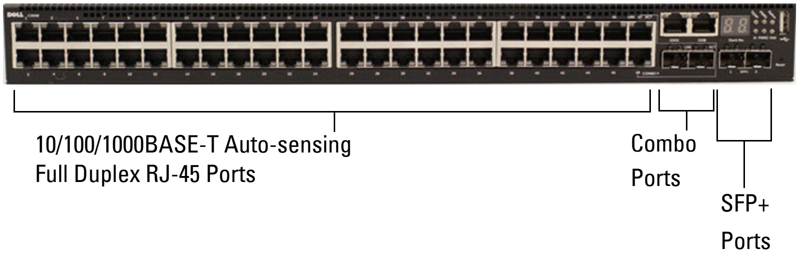
The following images show the front panels of the switch models in the Dell Networking N3000 Series.

Figure 3-11. Dell Networking N3024F with 24 10/100/1000BASE-T Ports (Front Panel)



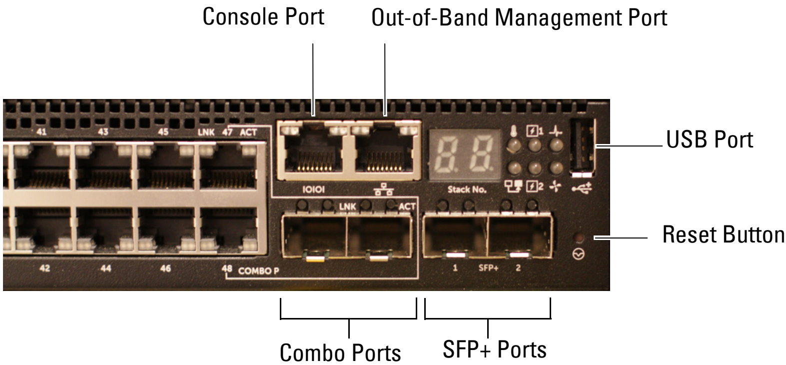
The Dell Networking N3000 Series switch includes two combo ports. The combo ports are SFP on the Dell Networking N3000 Series and 1000BaseT on the Dell Networking N3024F switch.

Figure 3-12. Dell Networking N3048 with 48 10/100/1000BASE-T Ports (Front Panel)



The additional ports are on the right side of the front panel, as shown in Figure 3-12 and Figure 3-13.

Figure 3-13. Additional Dell Networking N3000 Series Ports



The Dell Networking N3000 Series front panel above also contains a reset button (pinhole) and several status LEDs. See Figure 3-13.

The Dell Networking N3000 Series front panel displays, from left to right, status LEDs for over-temperature alarm, internal power supply 1, and switch status on the top row. The bottom row of status LEDs displays, from left to right, the Stack Master, internal power supply 2, and fan alarm status.

Switch Ports

The Dell Networking N3024/N3024P front panel provides 24 Gigabit Ethernet (10/100/1000BASE-T) RJ-45 ports that support auto-negotiation for speed, flow control, and duplex. The Dell Networking N3024P models support two SFP+ 10G ports. Dell-qualified SFP+ transceivers are sold separately. The Dell Networking N3000 Series switches operate in full-duplex mode only.

The Dell Networking N3024F front panel provides 24 Gigabit Ethernet 100BASE-FX/1000BASE-X SFP ports plus 2 1000BASE-T combo ports. Dell-qualified SFP transceivers are sold separately. 1000BASE-T operation requires the use of auto-negotiation.

The Dell Networking N3048/N3048P front panel provides 48 Gigabit Ethernet (10BASE-T, 100BASE-TX, 1000BASE-T) RJ-45 ports that support auto-negotiation for speed, flow control, and duplex. The Dell Networking N3048/N3048P supports two SFP+ 10G ports. Dell-qualified SFP+ transceivers are sold separately.

The front-panel switch ports have the following characteristics:

- The switch automatically detects the difference between crossed and straight-through cables on RJ-45 ports and automatically chooses the MDI or MDIX configuration to match the other end.
- SFP+ ports support Dell-qualified transceivers. The default behavior is to log a message and generate an SNMP trap on insertion or removal of an optic that is not qualified by Dell. The message and trap can be suppressed by using the **service unsupported-transceiver** command.
- RJ-45 ports support full-duplex mode 10/100/1000 Mbps speeds on standard Category 5 UTP cable.
- SFP+ ports support SFP+ transceivers and SFP+ copper twin-ax technology operating at 10G/1G speeds in full-duplex mode. SFP transceivers are supported in SFP+ ports and operate at 1G full-duplex. SFP transceivers require auto-negotiation to be enabled.
- The Dell Networking N3024P/N3048P front-panel ports support PoE (15.4W) and PoE+ (34.2W) as well as legacy capacitive detection for pre-standard PDs.
- Additionally, ports 1–12 support UPOE (60W) power when configured in high-power mode.

Combo Ports

Combo ports automatically select the active media and always choose fiber media if both copper and fiber are active. Copper combo ports do not support 10 Mbps forced mode.

Console Port

The console port provides serial communication capabilities, which allows communication using RS-232 protocol. The serial port provides a direct connection to the switch and allows access to the CLI from a console terminal connected to the port through the provided serial cable (with RJ45 YOST to female DB-9 connectors).

The console port is separately configurable and can be run as an asynchronous link from 1200 baud to 115,200 baud.

The Dell CLI only supports changing the speed.

The defaults are 9600 baud rate, 8 data bits, No Parity, 1 Stop Bit, No Flow Control.

Out-of-Band Management Port

The Out-of-Band (OOB) management port is a 10/100/1000BASE-T Ethernet port connected directly to the switch CPU and dedicated to switch management. Traffic on this port is segregated from operational network traffic on the switch ports and cannot be switched or routed to or from the operational network. In addition, ACLs (including management ACLS), do not operate on the out-of-band port. Connect the out-of-band port only to a physically secure network.

USB Port

The Type-A, female USB port supports a USB 2.0-compliant flash memory drive. The Dell Networking N-Series switch can read or write to a flash drive with a single partition formatted as FAT-32. Use a USB flash drive to copy switch configuration files and images between the USB flash drive and the switch. It is also possible to use the USB flash drive to move and copy configuration files and images from one switch to other switches in the network. The system does not support the deletion of files on attached USB flash drives.

The USB port does not support any other type of USB device.

Reset Button

The reset button is accessed through the pinhole and enables performing a hard reset on the switch. To use the reset button, insert an unbent paper clip or similar tool into the pinhole. When the switch completes the boot process after the reset, it resumes operation with the most recently saved configuration. Any changes made to the running configuration that were not saved to the startup configuration prior to the reset are lost.

Port and System LEDs

The front panel contains light emitting diodes (LEDs) that indicate the status of port links, power supplies, fans, stacking, and the overall system status.

For information about the status that the LEDs indicate, see the *User's Configuration Guide.*

Stack Master LED and Stack Number Display

When a switch within a stack is the master unit, the Stack Master LED is solid green. If the Stack Master LED is off, the stack member is not the master unit. The Stack No. panel displays the unit number for the stack member. If a switch is not part of a stack (in other words, it is a stack of one switch), the Stack Master LED is illuminated and the unit number is displayed.

Back Panel

The following images show the back panels of the Dell Networking N3000 Series switches.

Figure 3-14. Dell Networking N3000 Series Back Panel

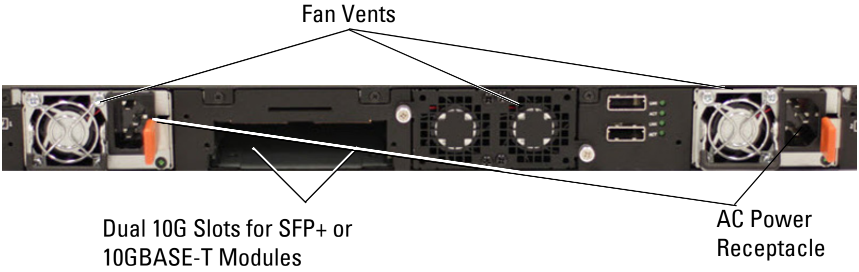


Figure 3-15. Dell Networking N3024P/N3048P Back Panel

Figure 3-16. Dell Networking N3048 Mini-SAS Stacking Ports Close-up

The term mini-SAS refers to the stacking port cable connections shown in Figure 3-16. See "Stacking " on page 193 for information on using the mini-SAS ports to connect switches.

Expansion Slots for Plug-in Modules

One expansion slot is located on the back of the Dell Networking N3000 Series models and can support the following modules:

- 10GBASE-T module
- SFP+ module

Each plug-in module has two ports. The plug-in modules include hot-swap support, so a switch reboot is not needed after a new module is installed.

Power Supplies

Dell Networking N3024, N3024F and N3048

Dell Networking N3024, N3024F and N3048 switches support two 200-watt Field Replaceable Unit (FRU) power supplies which give full power redundancy for the switch. The Dell Networking N3024, N3024F, and N3048 switches offer the V-lock feature for users desiring the need to eliminate accidental power disconnection. The V-lock receptacle on the Power Supply Unit (PSU) allows for the use of a power cord that has the V-lock feature to create an integral secure locking connection.

Dell Networking N3024P and N3048P

Dell Networking N3024P and N3048P switches support one or two 1100-watt FRU power supplies. The Dell Networking N3024P switch is supplied with a single 715-watt power supply (the default configuration) and supports an additional 1100-watt supply. For the Dell Networking N3048P switch, a single 1100-watt power supply is supplied and another 1100 watt power supply can be added.

A single 1100-watt power supply can feed up to 24 PoE devices at full PoE+ power (950W). Dual-equipped switches will feed up to 48 PoE devices at full PoE+ power (1800W), as well as provide power supply redundancy.

NOTE: PoE power is dynamically allocated by default. Not all ports will require the full PoE+ power.

CAUTION: Remove the power cable from the power supplies prior to removing the power supply module itself. Power must not be connected prior to insertion in the chassis.

Ventilation System

Two fans cool the Dell Networking N3000 Series switches. The Dell Networking N3000 Series switches additionally have a fan in each internal power supply. The Dell Networking N3000 Series fan is field-replaceable.

Information Tag

The back panel includes a slide-out label panel that contains system information, such as the Service Tag, MAC address, and other information.

LED Definitions

This section describes the LEDs on the front and back panels of the switch.

Port LEDs

Each port on a Dell Networking N3000 Series switch includes two LEDs. One LED is on the left side of the port, and the second LED is on the right side of the port. This section describes the LEDs on the switch ports.

100/1000/10000Base-T Port LEDs

Each 100/1000/10000Base-T port has two LEDs. Figure 3-17 illustrates the 100/1000/10000Base-T port LEDs.

Figure 3-17. 100/1000/10000Base-T Port LEDs

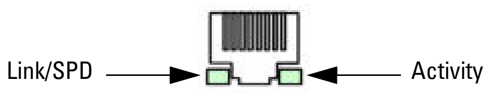


Table 3-13 shows the 100/1000/10000Base-T port LED definitions.

Table 3-13. 100/1000/10000Base-T Port Definitions

LED	Color	Definition
Link/SPD LED	Off	There is no link.
	Solid yellow	The port is operating at 10/100 Mbps.
	Solid green	The port is operating at 1000 Mbps.
Activity LED (on non-PoE switches)	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.
Activity/PoE LED (on PoE switches)	Off	There is no current transmit/receive activity and PoE power is off.
	Blinking green	The port is actively transmitting/receiving and PoE power is off.
	Blinking yellow	The port is actively transmitting/receiving and PoE power is on.
	Solid yellow	There is no current transmit/receive activity and PoE power is on.

Module Bay LEDs

The following tables describe the purpose of each of the module bay LEDs when SFP+ and 10GBaseT modules are used.

Table 3-14. SFP+ Module LED Definitions

LED	Color	Definition
Link/SPD LED	Off	There is no link.
	Solid green	The port is operating at 10 Gbps.
	Solid amber	The port is operating at 1000 Mbps.
Activity LED	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.

Table 3-15. 10GBase-T Module LED Definitions

LED	Color	Definition
Link/SPD LED	Off	There is no link.
	Solid green	The port is operating at 10 Gbps.
	Solid amber	The port is operating at 100/1000 Mbps.
Activity LED	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.

Table 3-16. Stacking Port LED Definitions

LED	Color	Definition
Link LED	Off	There is no link.
	Solid green	The port is actively transmitting/receiving.
Activity LED	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.

Table 3-17. OOB Port LED Definitions

LED	Color	Definition
Link/SPD LED	Off	There is no link.
	Solid green	The port is actively transmitting/receiving at 1000 Mbps.
	Solid amber	The port is actively transmitting/receiving at 10/100 Mbps.
Activity LED	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.

Table 3-18. Console Port LED Definitions

LED	Color	Definition
Link/SPD LED	Off	There is no link.
	Solid green	A link is present.

System LEDs

The system LEDs, located on the back panel, provide information about the power supplies, thermal conditions, and diagnostics.

Table 3-19 shows the System LED definitions for the Dell Networking N3000 Series switches.

Table 3-19. System LED Definitions

LED	Color	Definition
Status	Solid green	Normal operation.
	Blinking green	The switch is booting
	Solid red	A critical system error has occurred.
	Blinking red	A noncritical system error occurred (fan or power supply failure).
Power 1, Power 2	Off	There is no power or the switch has experienced a power failure.
	Solid green	Power to the switch is on.
	Blinking green	The switch locator function is enabled.
Fan	Solid green	The fan is powered and is operating at the expected RPM.
	Solid red	A fan failure has occurred.
Stack Master	Off	The switch is in stand-alone mode.
	Solid green	The switch is master for the stack.
Temp	Solid green	The switch is operating below the threshold temperature.
	Solid red	The switch temperature exceeds the threshold of 75°C.
Stack No.	–	Switch ID within the stack.

Power Consumption for PoE Switches

Table 3-20 shows power consumption data for the PoE-enabled switches.

Table 3-20. Dell Networking N3000 Series Power Consumption

Model	Input Voltage	Power Supply Configuration	Max Steady Current Consumption (A)	Max Steady Power (W)
Dell Networking N3024P	100V	PSU1+PSU2	13.1	1310.0
	110V	PSU1+PSU2	11.7	1287.0
	120V	PSU1+PSU2	10.6	1272.0
	220V	PSU1+PSU2	5.6	1232.0
	240V	PSU1+PSU2	5.2	1240.8
Dell Networking N3048P	100V	PSU1+PSU2	21.8	2180.0
	110V	PSU1+PSU2	19.5	2145.0
	120V	PSU1+PSU2	17.8	2136.0
	220V	PSU1+PSU2	9.31	2048.2
	240V	PSU1+PSU2	8.6	2064.0

The PoE power budget for each interface is controlled by the switch firmware. The administrator can limit the power supplied on a port or prioritize power to some ports over others. Table 3-21 shows the power budget data.

Table 3-21. Dell Networking N3000 Series PoE Power Budget Limit

		One PSU		Two PSUs	
Model Name	**System Power Dissipation**	**Max. PSU Output Ability**	**PoE+ Power Turn-on Limitation**	**Max. PSUs Output Ability**	**PoE+ Power Turn-on Limitation**
Dell Networking N3024P	110W	715W	Power budget is 550W: The total PoE supplied power must not exceed 550W.	715W	Power budget is 1100W: All PoE+ ports can supply maximum power.
Dell Networking N3048P	140W	1100W	Power budget is 950W: The total PoE supplied power must not exceed 950W.	2200W	Power budget is 1900W: All PoE+ ports can supply maximum power.

Dell Networking N4000 Series Switch Hardware

NOTE: Both the Dell Networking PC8100 and N4000 Series switches can run firmware versions 6.0.0.8 and beyond. The Dell Networking N4000 Series switches cannot run firmware prior to version 6.0.0.8.

This section contains information about device characteristics and modular hardware configurations for the Dell Networking N4000 Series switches.

Front Panel

The Dell Networking N4000 Series front panel includes the following features:

- Switch ports
- Module bay that supports the following modules:
  - 2 x 40 Gig QSFP (each QSFP may be configured as 4 x 10 Gig ports)
  - 4 x SFP+ module
  - 4 x 10GBaseT module

  See "Hot-Pluggable Interface Modules " on page 134 for more information.
- USB port
- Reset button
- Port and system LEDs
- Stack LED

The Dell Networking N4032 front panel provides 24 x 10GbE copper ports that support up to 100M of CAT-6A UTP cabling. The Dell Networking N4032F provides 24 SFP+ ports supporting SFP+ and SFP transceivers.

Figure 3-18.   Dell Networking N4032 Front Panel

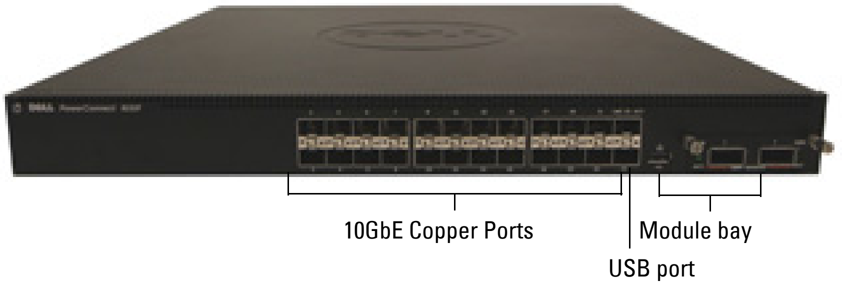
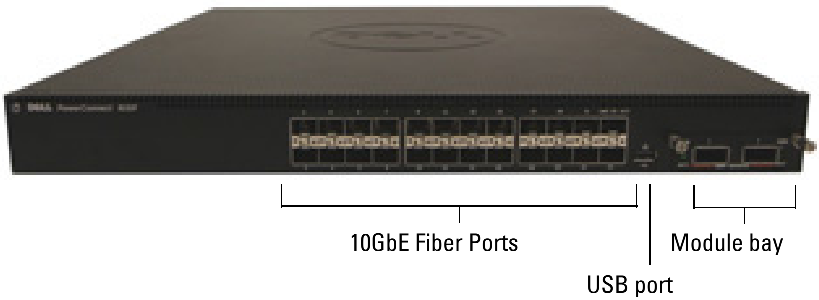


Figure 3-19.   Dell Networking N4032F Front Panel



Dell Networking N4032 and N4032F switches can be stacked with other Dell Networking N4000 Series switches using 10G or 40G SFP+ or QSFP modules in the module bay.

The Dell Networking N4064 front panel provides 48 x 10GbE copper ports and two fixed QSFP ports, each supporting 4 x 10G or 1 x 40G connections. The Dell Networking N4064F front panel provides 48 SFP+ ports supporting SFP+ and SFP transceivers plus two fixed QSFP ports, each supporting 4 x 10G or 1 x 40G connections.

All Dell Networking N4000 Series ports operate in full-duplex mode only. The copper ports require that auto-negotiation be enabled. Auto-negotiation should be disabled for Dell Networking N4000 Series SFP+ ports and enabled for SFP+ ports with SFP transceivers installed.

Figure 3-20. Dell Networking N4064 Front Panel

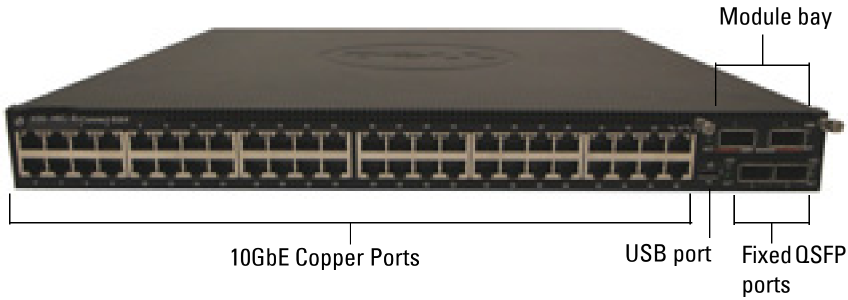
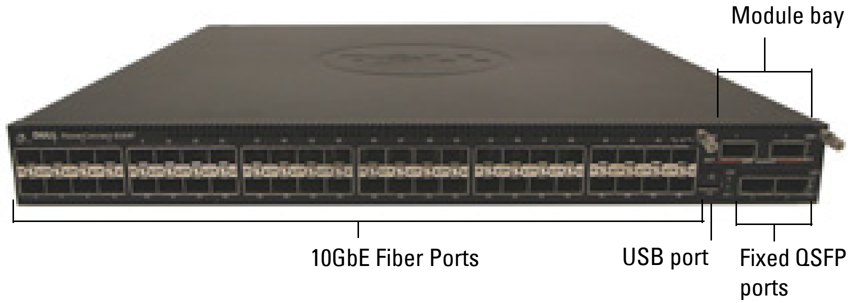


Figure 3-21. Dell Networking N4064F Front Panel



The Dell Networking N4064 and N4064F switches can be stacked with other Dell Networking N4000 Series switches using the 10G or 40G SFP+ or QSFP modules in the module bay or fixed QSFP ports.

Hot-Pluggable Interface Modules

The Dell Networking N4032, N4032F, N4064, and N4064F switches support the following hot-pluggable interface modules:

- N4000-QSFP — 2 x 40G QSFP port module - defaults to 2 x 40G
- N4000-SFP+ — 4 x SFP+ port module - defaults to 4 x 10G mode
- N4000-10GBT — 4 x 10GBase-T ports module - defaults to 4 x 10G mode
- Blank module — defaults to 10G mode

A reboot is not necessary when a hot-pluggable module is replaced with a module of different type. Plug-in modules with any port configured as a stacking port are not hot-swappable. Remove the stack-port configuration from a slot before plugging in a module.

A **no slot** or **clear config** command must be executed prior to inserting the new module. Note that changing the role of a port from stacking to Ethernet or vice-versa also requires a switch reboot.

If a **no slot** command is not issued prior to inserting a module, a message such as the following will appear:

```
Card Mismatch: Unit:1 Slot:1 Inserted-Card: Dell 2 Port QSFP
Expansion Card Config-Card: Dell 4 Port 10GBase-T Expansion Card
```

The following sections provides details on each module.

Quad-Port SFP (QSFP) Uplink Module

The QSFP module supports features four ports that support 10G SFP+ transceivers. The QSFP module supports the following features:

- Four 10G ports with quad-breakout/QBO cable or one 40G port
- Front-panel port status LEDs

The QSFP interfaces can be used for stacking. Stacking is supported at distances of up to 100M.

Quad-Port SFP+ Uplink Module

The N4000-SFP+ module features four SFP+ ports, each providing the following features:

- SFP+ optical interfaces
- SFP+ copper twinax interface
- Front-panel port status LEDs

The SFP+ connections can be used for stacking. Stacking is supported at distances of up to 100M.

10GBase-T Copper Uplink Module

The 10GBase-T copper module features four copper ports that can support 10GbE/1GbE/100MbE switching and provides following features:

- Complies with IEEE802.3z, IEEE 802.3, IEEE802.3u, IEEE802.3ab, IEEE802.3az, IEEE802.3an
- Four 10GBase-T/1GBase-T/100MBase-T copper ports.
- front-panel port status LEDs

USB Port

The Type-A, female USB port supports a USB 2.0-compliant flash memory drive. The Dell Networking N4000 Series switch can read or write to a flash drive with a single partition formatted as FAT-32. Use a USB flash drive to copy switch configuration files and images between the USB flash drive and the switch. The USB flash drive may be used to move and copy configuration files and images from one switch to other switches in the network. Deletion of files on the USB drive is not supported.

The USB port does not support any other type of USB device.

Port and System LEDs

The front panel contains light emitting diodes (LEDs) to indicate port status.

For information about the status that the LEDs indicate, see "LED Definitions " on page 138.

SFP+ and QSFP+ Ports

SFP+ and QSFP+ ports support Dell-qualified transceivers. The default behavior is to log a message and generate an SNMP trap on insertion or removal of an optic that is not qualified by Dell. This message and trap can be suppressed by using the **service unsupported-transceiver** command.

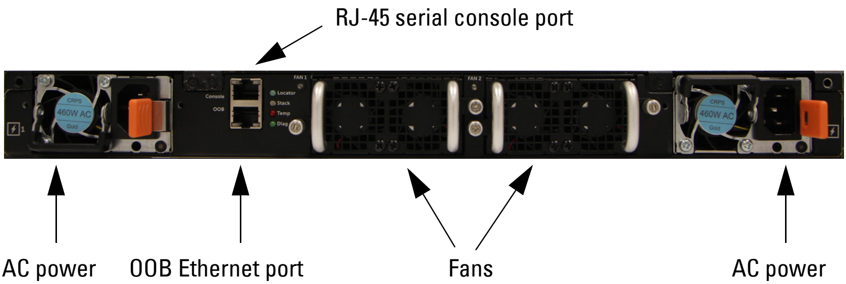
Back Panel

The Dell Networking N4000 Series back panel has the following features:

- Console port
- Out-of-band management port
- Power Supplies
- Ventilation System

The following image show the back panel of the Dell Networking N4000 Series switches.

Figure 3-22. Dell Networking N4000 Series Back Panel



Console Port

The console port is for management through a serial interface. This port provides a direct connection to the switch and provides access to the CLI from a console terminal connected to the port through the provided serial cable (RJ-45 to female DB-9 connectors).

The console port supports asynchronous data of eight data bits, one stop bit, no parity bit, and no flow control. The default baud rate is 9600 bps.

Out-of-Band Management Port

The Out-of-Band (OOB) management port is a 10/100/1000BASE-T Ethernet port connected directly to the switch CPU and dedicated to switch management. Traffic on this port is segregated from operational network traffic on the switch ports and cannot be switched or routed to or from the operational network. In addition, ACLs (including management ACLS), do not operate on the out-of-band port. Only connect the out-of-band port to a physically secure network.

Power Supplies

Each Dell Networking N4000 Series switch has two power supplies for redundant or loadsharing operation. Each power supply can support 300W.

CAUTION: Remove the power cable from the modules prior to removing the module itself. Power must not be connected prior to insertion in the chassis.

Ventilation System

The Dell Networking N4000 Series switches have two fans. Each switch also has four thermal sensors and a fan speed controller, which can be used to control FAN speeds. Verify operation by observing the LEDs.

LED Definitions

This section describes the LEDs on the front and back panels of the switch.

Port LEDs

Each port on a Dell Networking N4000 Series switch includes two LEDs. One LED is on the left side of the port, and the second LED is on the right side of the port. This section describes the LEDs on the switch ports.

100/1000/10000Base-T Port LEDs

Each 100/1000/10000Base-T port has two LEDs. Figure 3-23 illustrates the 100/1000/10000Base-T port LEDs.

Figure 3-23. 100/1000/10000Base-T Port LEDs

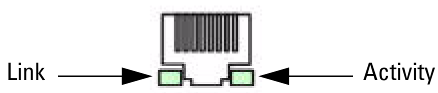


Table 3-22 shows the 100/1000/10000Base-T port LED definitions.

Table 3-22. 100/1000/10000Base-T Port Definitions

LED	Color	Definition
Link LED	Off	There is no link.
	Solid green	The port is operating at 10 Gbps.
	Solid amber	The port is operating at 100/1000 Mbps.
Activity LED	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.

Module Bay LEDs

The following tables describe the purpose of each of the module bay LEDs when SFP+, 10GBase-T, and QSFP modules are used.

Table 3-23. SFP+ Module LED Definitions

LED	Color	Definition
Link LED	Off	There is no link.
	Solid green	The port is operating at 10 Gbps.
	Solid amber	The port is operating at 100/1000 Mbps.
Activity LED	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.

Table 3-24. 10GBase-T Module LED Definitions

LED	Color	Definition
Link LED	Off	There is no link.
	Solid green	The port is operating at 10 Gbps.
	Solid amber	The port is operating at 100/1000 Mbps.
Activity LED	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.

Table 3-25. QSFP Module LED Definitions

LED	Color	Definition
Link LED	Off	There is no link.
	Solid green	The port is operating at 40 Gbps.
	Solid amber	The port is operating at other speeds.
Activity LED	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.

Out-of-Band Ethernet Management Port LEDs

Table 3-26 shows the LED definitions for the OOB Ethernet management port.

Table 3-26. OOB Ethernet Management Port LED Definitions

LED	Color	Definition
Link LED	Off	There is no link.
	Solid green	The port is operating at 1000 Mbps.
	Solid amber	The port is operating at 10/100 Mbps.
Activity LED	Off	There is no current transmit/receive activity.
	Blinking green	The port is actively transmitting/receiving.

System LEDs

The system LEDs, located on the back panel, provide information about the power supplies, thermal conditions, and diagnostics.

Table 3-27 shows the System LED definitions for the Dell Networking N4000 Series switches.

Table 3-27. System LED Definitions—Dell Networking N4000 Series Switches

LED	Color	Definition
System	Blinking blue	The switch is booting
	Solid red	A critical system error has occurred.
	Blinking red	A noncritical system error occurred (fan or power supply failure).
Temp	Off	The switch is operating at normal temperature.
	Solid amber	The thermal sensor's system temperature threshold of 75°C has been exceeded.
Diag	Off	The switch is operating normally
	Blinking green	A diagnostic test is running.
Fan	Solid green	The fan is powered and is operating at the expected RPM.
	Solid red	A fan failure has occurred.
Stack	Solid blue	The switch is in stacking master mode.
	Solid amber	The switch is in stacking slave mode.
	Off	The switch is in stand-alone mode.
Locator	Blinking green	The locator function is enabled.
	Solid green	The locator function is disabled.

Switch MAC Addresses

The switch allocates MAC addresses from the Vital Product Data information stored locally in flash. MAC addresses are used as follows:

Table 3-28. MAC Address Use

Base	Switch address, layer 2
Base + 1	Out-of-band port (not available on Dell Networking N1500/N2000 Series switches)
Base + 3	Layer 3

Shown below are three commands that display the MAC addresses used by the switch:

```
console#show system

System Description: Dell Ethernet Switch
System Up Time: 0 days, 00h:05m:11s
System Contact:
System Name:
System Location:
Burned In MAC Address: 001E.C9F0.004D
System Object ID: 1.3.6.1.4.1.674.10895.3042
System Model ID: N4032
Machine Type: N4032
Temperature Sensors:

Unit     Description        Temperature    Status
                            (Celsius)
----     -----------        -----------    ------
1        MAC                32             Good
1        CPU                31             Good
1        PHY (left side)    26             Good
1        PHY (right side)   29             Good

Fans:

Unit     Description    Status
----     -----------    ------
 1       Fan 1          OK
 1       Fan 2          OK
 1       Fan 3          OK
 1       Fan 4          OK
 1       Fan 5          OK
 1       Fan 6          No Power

Power Supplies:

Unit  Description     Status      Average      Current          Since
                                   Power        Power          Date/Time
                                  (Watts)      (Watts)
```

```
---- ----------- ----------- ---------- -------- -------------------
1     System      OK          42.0       43.4
1     Main        OK          N/A        N/A      04/06/2001 16:36:16
1     Secondary   No Power    N/A        N/A      01/01/1970 00:00:00

USB Port Power Status:
----------------------
Device Not Present
console#show ip interface out-of-band

IP Address..................................... 10.27.21.29
Subnet Mask.................................... 255.255.252.0
Default Gateway................................ 10.27.20.1
Configured IPv4 Protocol....................... DHCP
Burned In MAC Address.......................... 001E.C9F0.004E
```

```
console#show ip interface vlan 1

Routing Interface Status....................... Down
Primary IP Address............................. 1.1.1.2/255.255.255.0
Method......................................... Manual
Routing Mode................................... Enable
Administrative Mode............................ Enable
Forward Net Directed Broadcasts................ Disable
Proxy ARP...................................... Enable
Local Proxy ARP................................ Disable
Active State................................... Inactive
MAC Address.................................... 001E.C9F0.0050
Encapsulation Type............................. Ethernet
IP MTU......................................... 1500
Bandwidth...................................... 10000 kbps
Destination Unreachables....................... Enabled
ICMP Redirects................................. Enabled
```

4

Using Dell OpenManage Switch Administrator

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This section describes how to use the Dell OpenManage Switch Administrator application. The topics covered in this section include:

- About Dell OpenManage Switch Administrator
- Starting the Application
- Understanding the Interface
- Using the Switch Administrator Buttons and Links
- Defining Fields

About Dell OpenManage Switch Administrator

Dell OpenManage Switch Administrator is a web-based tool for managing and monitoring Dell Networking N1500, N2000, N3000, and N4000 Series switches. Table 4-1 lists the web browsers that are compatible with Dell OpenManage Switch Administrator. The browsers have been tested on a PC running the Microsoft Windows operating system.

Table 4-1. Compatible Browsers

Browser	Version
Internet Explorer	v9
Mozilla Firefox	v14
Safari	v5.0
Chrome	v21

NOTE: Additional operating systems and browsers might be compatible but have not been explicitly tested with Dell OpenManage Switch Administrator.

Starting the Application

To access the Dell OpenManage Switch Administrator and log on to the switch:

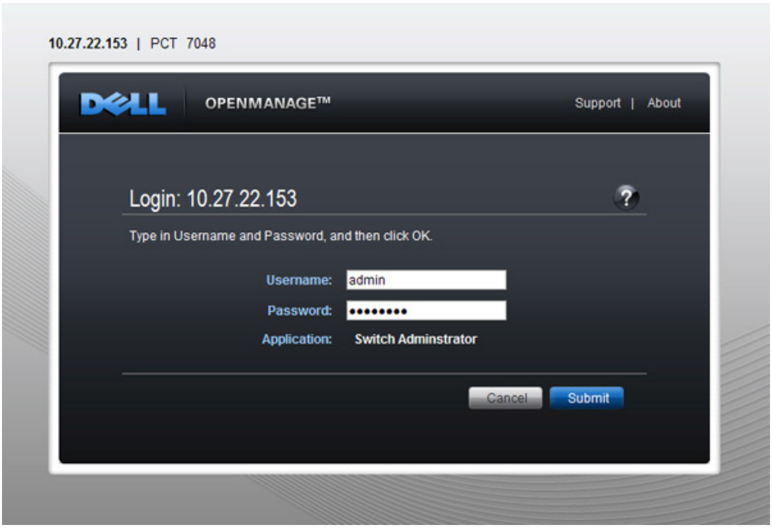
1 Open a web browser.

2 Enter the IP address of the switch in the address bar and press <Enter>.

   For information about assigning an IP address to a switch, see "Setting the IP Address and Other Basic Network Information " on page 165.

3 When the **Login** window displays, enter a user name and password. Passwords are both case sensitive and alpha-numeric.

Figure 4-1. Login Screen



NOTE: The switch is not configured with a default user name or password. The administrator must connect to the CLI by using the console port to configure the initial user name and password. For information about connecting to the console, see "Console Connection " on page 153. For information about creating a user and password, see "Authentication, Authorization, and Accounting " on page 229.

4 Click **Submit**.

5 The **Dell OpenManage Switch Administrator** home page displays.

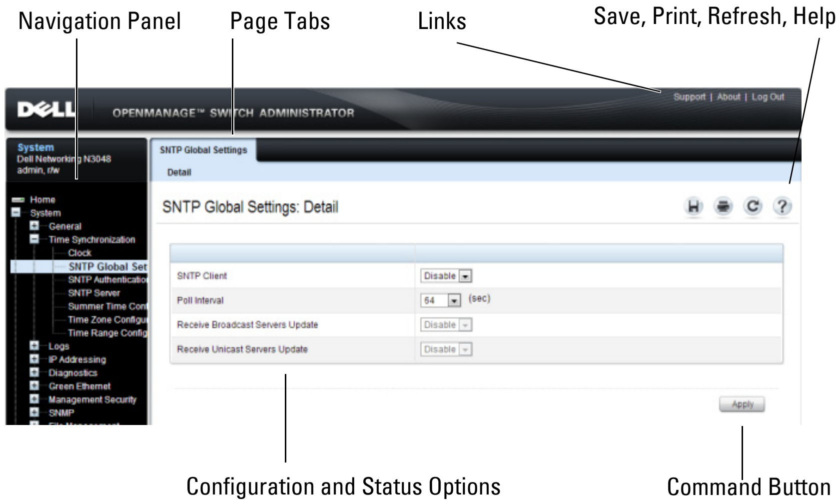
The home page is the **Device Information** page, which contains a graphical representation of the front panel of the switch. For more information about the home page, see "Device Information " on page 359.

Understanding the Interface

The Dell OpenManage Switch Administrator interface contains the following components:

- Navigation panel — Located on the left side of the page, the navigation pane provides an expandable view of features and their components.
- Configuration and status options — The main panel contains the fields used to configure and monitor the switch.
- Page tabs — Some pages contain tabs that allow the administrator to access additional pages related to the feature.
- Command buttons — Command buttons are located at the bottom of the page. Use the command buttons to submit changes, perform queries, or clear lists.
- Save, Print, Refresh, and Help buttons — These buttons appear on the top-right side of the main panel and are on every page.
- Support, About, and Logout links — These links appear at the top of every page.

Figure 4-2. Switch Administrator Components



Using the Switch Administrator Buttons and Links

Table 4-2 describes the buttons and links available from the **Dell OpenManage Switch Administrator** interface.

Table 4-2. Button and Link Descriptions

Button or Link	Description
Support	Opens the Dell Support page at www.dell.com/support.
About	Contains the version and build number and Dell copyright information.
Log Out	Logs out of the application and returns to the login screen.
Save	Saves the running configuration to the startup configuration. When a user clicks **Apply**, changes are saved to the running configuration. When the system boots, it loads the startup configuration. Any changes to the running configuration that were not saved to the startup configuration are lost across a power cycle.
Print	Opens the printer dialog box that enables printing the current page. Only the main panel prints.
Refresh	Refreshes the screen with the current information.
Help	Online help that contains information to assist in configuring and managing the switch. The online help pages are context sensitive. For example, if the **IP Addressing** page is open, the help topic for that page displays if the user clicks **Help**.
Apply	Updates the running configuration on the switch with the changes. Configuration changes take effect immediately.
Clear	Resets statistic counters and log files to the default configuration.
Query	Queries tables.
Left arrow and Right arrow	Moves information between lists.

NOTE: A few pages contain a button that occurs only on that page. Page-specific buttons are described in the sections that pertain to those pages.

Defining Fields

User-defined fields can contain 1–159 characters, unless otherwise noted on the Dell OpenManage Switch Administrator web page.

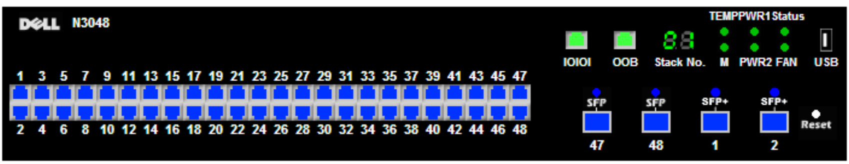
All characters may be used except for the following:

- \
- /
- :
- *
- ?
- <
- >
- |

Understanding the Device View

The Device View shows various information about switch. This graphic appears on the OpenManage Switch Administrator **Home** page, which is the page that displays after a successful login. The graphic provides information about switch ports and system health.

Figure 4-3. Dell Networking N3048 Device View



Using the Device View Port Features

The switching-port coloring indicates if a port is currently active. Green indicates that the port has a link, red indicates that an error has occurred on the port, and blue indicates that the link is down. Each port image is a hyperlink to the **Port Configuration** page for the specific port.

Using the Device View Switch Locator Feature

The Device View graphic includes a **Locate** button and a drop-down menu of timer settings. When the user clicks **Locate**, the switch locator LED on the back panel of the switch blinks for the number of seconds selected from the timer menu. The green, blinking LED on the back of the switch can help the administrator or a technician near the switch identify the physical location of the switch within a room or rack full of switches. After the user clicks the **Locate** button, it turns green and remains green while the LED is blinking.

NOTE: The **locate** command in the CLI can be used to enable the locator LED.

5

Using the Command-Line Interface

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This section describes how to use the Command-Line Interface (CLI) on a Dell Networking N1500, N2000, N3000, and N4000 Series switches.

The topics covered in this section include:

- Accessing the Switch Through the CLI
- Understanding Command Modes
- Entering CLI Commands

Accessing the Switch Through the CLI

The CLI provides a text-based way to manage and monitor the Dell Networking N1500, N2000, N3000, and N4000 Series switches. The CLI can be accessed using a direct connection to the console port or by using a Telnet or SSH client.

To access the switch by using Telnet or Secure Shell (SSH), the switch must have an IP address, and the management station used to access the device must be able to ping the switch IP address.

For information about assigning an IP address to a switch, see "Setting the IP Address and Other Basic Network Information " on page 165.

Console Connection

Use the following procedures to connect to the CLI by connecting to the console port. For more information about creating a serial connection, see the *Getting Started Guide* available at www.dell.com/support.

1. Connect the DB-9 connector of the supplied serial cable to a management station, and connect the RJ-45 connector to the switch console port.

   On Dell Networking N1500, N2000, and N3000 Series switches, the console port is located on the right side of the front panel and is labeled with the |O|O| symbol. On the Dell Networking N4000 Series switches, it is located on the back panel above the OOB Ethernet port.

NOTE: For a stack of switches, be sure to connect to the console port on the Master switch. The Master LED is illuminated on the stack Master.

2 Start the terminal emulator, such as Microsoft HyperTerminal, and select the appropriate serial port (for example, COM 1) to connect to the console.

3 Configure the management station serial port with the following settings:

- Data rate — 9600 baud.
- Data format — 8 data bits
- Parity — None
- Stop bits — 1
- Flow control — None

4 Power on the switch (or stack).

After the boot process completes, the `console>` prompt displays, and CLI commands can be entered.

NOTE: By default, no authentication is required for console access. However, if an authentication method has been configured for console port access, the User: login prompt displays.

Telnet Connection

Telnet is a terminal emulation TCP/IP protocol. ASCII terminals can be virtually connected to the local device through a TCP/IP protocol network.

Telnet connections are enabled by default, and the Telnet port number is 23. The switch supports up to four simultaneous Telnet sessions. All CLI commands can be used over a Telnet session.

NOTE: SSH, which is more secure than Telnet, is disabled by default.

To connect to the switch using Telnet, the switch must have an IP address, and the switch and management station must have network connectivity. Any Telnet client on the management station can be used to connect to the switch.

A Telnet session can also be initiated from the OpenManage Switch Administrator. For more information, see "Initiating a Telnet Session from the Web Interface " on page 399.

Understanding Command Modes

The CLI groups commands into modes according to the command function. Each of the command modes supports specific software commands. The commands in one mode are not available until the user switches to that particular mode, with the exception of the User EXEC mode commands. The User EXEC mode commands can be executed in the Privileged EXEC mode.

To display the commands available in the current mode, enter a question mark (?) at the command prompt. In each mode, a specific command is used to navigate from one command mode to another.

The main command modes include the following:

- User EXEC — Commands in this mode permit connecting to remote devices, changing terminal settings on a temporary basis, performing basic tests, and listing system information.
- Privileged EXEC — Commands in this mode enable viewing all switch settings and entering the global configuration mode.
- Global Configuration — Commands in this mode manage the device configuration on a global level and apply to system features, rather than to a specific protocol or interface.
- Interface Configuration — Commands in this mode configure the settings for a specific interface or range of interfaces.
- VLAN Configuration — Commands in this mode create and remove VLANs and configure IGMP/MLD Snooping parameters for VLANs.

The CLI includes several additional command modes. For more information about the CLI command modes, including details about all modes, see the *CLI Reference Guide*.

Table 5-1 describes how to navigate between CLI Command Mode and lists the prompt that displays in each mode.

Table 5-1. Command Mode Overview

Command Mode	Access Method	Command Prompt	Exit or Access Previous Mode
User EXEC	The user is automatically in User EXEC mode unless the user is defined as a privileged user.	`console>`	logout
Privileged EXEC	From User EXEC mode, enter the **enable** command	`console#`	Use the **exit** command, or press Ctrl-Z to return to User EXEC mode.
Global Configuration	From Privileged EXEC mode, use the **configure** command.	`console(config)#`	Use the **exit** command, or press Ctrl-Z to return to Privileged EXEC mode.
Interface Configuration	From Global Configuration mode, use the **interface** command and specify the interface type and ID.	`console(config-if)#`	To exit to Global Configuration mode, use the **exit** command, or press Ctrl-Z to return to Privileged EXEC mode.

Entering CLI Commands

The switch CLI provides several techniques to help users enter commands.

Using the Question Mark to Get Help

Enter a question mark (?) at the command prompt to display the commands available in the current mode.

```
console(config-vlan)#?

exit          To exit from the mode.
help          Display help for various special keys.
ip            Configure IP parameters.
ipv6          Configure IPv6 parameters.
protocol      Configure the Protocols associated with
              particular Group Ids.
vlan          Create a new VLAN or delete an existing
              VLAN.
```

Enter a question mark (?) after entering each word to display available command keywords or parameters.

```
console(config)#vlan ?

<vlan-list>    <1-4093> - separate non-consecutive IDs with ',' and
               no spaces; Use '-' for range.
makestatic     Change the VLAN type from 'Dynamic' to 'Static'.
protocol       Configure the protocol based VLAN settings.
```

If there are no additional command keywords or parameters, or if additional parameters are optional, the following message appears in the output:

```
<cr>          Press enter to execute the command.
```

Typing a question mark (?) after one or more characters of a word shows the available command or parameters that begin with the characters, as shown in the following example:

```
console#show po?

policy-map               port                     ports
```

Using Command Completion

The CLI can complete partially entered commands when the <Tab> or <Space> key are pressed.

```
console#show run<Tab>
console#show running-config
```

If the characters entered are not enough for the switch to identify a single matching command, continue entering characters until the switch can uniquely identify the command. Use the question mark (?) to display the available commands matching the characters already entered.

Entering Abbreviated Commands

To execute a command, enter enough characters so that the switch can uniquely identify a command. For example, to enter Global Configuration mode from Privileged EXEC mode, enter **conf** instead of **configure**.

```
console#conf

console(config)#
```

Negating Commands

For many commands, the prefix keyword **no** is entered to cancel the effect of a command or reset the configuration to the default value. Many configuration commands have this capability.

Command Output Paging

Lines are printed on the screen up to the configured terminal length limit (default 24). Use the space bar to show the next page of output or the carriage return to show the next line of output. Setting the terminal length to zero disables paging. Command output displays until no more output is available.

Understanding Error Messages

If a command is entered and the system is unable to execute it, an error message appears. Table 5-2 describes the most common CLI error messages.

Table 5-2. CLI Error Messages

Message Text	Description
`% Invalid input detected at '^' marker.`	Indicates that an incorrect or unavailable command was entered. The carat (^) shows where the invalid text is detected. This message also appears if any of the parameters or values are not recognized.
`Command not found / Incomplete command. Use ? to list commands.`	Indicates that the required keywords or values were not entered.
`Ambiguous command`	Indicates that not enter enough letters were entered to uniquely identify the command.

If you attempt to execute a command and receive an error message, use the question mark (?) to help determine the possible keywords or parameters that are available.

Recalling Commands from the History Buffer

Every time a command is entered in the CLI, it is recorded in an internally managed Command History buffer. By default, the history buffer is enabled and stores the last 10 commands entered. These commands can be recalled, reviewed, modified, and reissued. This buffer is not preserved after switch resets.

Table 5-3. History Buffer Navigation

Keyword	Source or Destination
Up-arrow key <Ctrl>+<P>	Recalls commands in the history buffer, beginning with the most recent command. Repeats the key sequence to recall successively older commands.
Down-arrow key <Ctrl>+<N>	Returns to more recent commands in the history buffer after recalling commands with the up-arrow key. Repeating the key sequence recalls more recent commands in succession.

6

Default Settings

This section describes the default settings for many of the software features on the Dell Networking N-Series switches.

Table 6-1. Default Settings

Feature	Default
IP address	DHCP on OOB interface, if equipped. DHCP on VLAN1, if no OOB interface
Subnet mask	None
Default gateway	None
DHCP client	Enabled on out-of-band (OOB) interface.
VLAN 1 Members	All switch ports
SDM template	Dual IPv4 and IPv6 routing
Users	None
Minimum password length	8 characters
IPv6 management mode	Enabled
SNTP client	Disabled
Global logging	Enabled
Switch auditing	Disabled
CLI command logging	Disabled
Web logging	Disabled
SNMP logging	Disabled
Console logging	Enabled (Severity level: warnings and above)
Monitor logging:	Disabled
Buffer (In-memory) logging	Enabled (Severity level: informational and above)

Table 6-1. Default Settings (Continued)

Feature	Default
Persistent (flash) logging	Enabled (Severity level: Emergencies and above)
DNS	Enabled (No servers configured)
SNMP	Enabled (SNMPv1)
SNMP Traps	Enabled
Auto Configuration	Enabled
Auto Save	Disabled
Stacking	Enabled
Nonstop Forwarding on the Stack	Enabled
sFlow	Disabled
ISDP	Enabled (Versions 1 and 2)
RMON	Enabled
TACACS+	Not configured
RADIUS	Not configured
SSH/SSL	Disabled
Telnet	Enabled
Denial of Service Protection	Disabled
Captive Portal	Disabled
IEEE 802.1X Authentication	Disabled
MAC-Based Port Security	All ports are unlocked
Access Control Lists (ACL)	None configured
IP Source Guard (IPSG)	Disabled
DHCP Snooping	Disabled
Dynamic ARP Inspection	Disabled
Protected Ports (Private VLAN Edge)	None
Energy Detect Mode	Disabled
EEE Lower Power Mode	Disabled

Table 6-1. Default Settings (Continued)

Feature	Default
PoE Plus (POE switches)	Auto
Flow Control Support (IEEE 802.3x)	Enabled
Maximum Frame Size	1518 bytes
Auto-MDI/MDIX Support	Enabled
Auto-negotiation	Enabled
Advertised Port Speed	Maximum Capacity
Broadcast Storm Control	Disabled
Port Mirroring	Disabled
LLDP	Enabled
LLDP-MED	Disabled
MAC Table Address Aging	300 seconds (Dynamic Addresses)
Cisco Protocol Filtering (LLPF)	No protocols are blocked
DHCP Layer-2 Relay	Disabled
Default VLAN ID	1
Default VLAN Name	Default
GVRP	Disabled
GARP Timers	Leave: 60 centiseconds Leave All: 1000 centiseconds Join: 20 centiseconds
Voice VLAN	Disabled
Guest VLAN	Disabled
RADIUS-assigned VLANs	Disabled
Double VLANs	Disabled
Spanning Tree Protocol (STP)	Enabled
STP Operation Mode	IEEE 802.1w Rapid Spanning Tree
Optional STP Features	Disabled
STP Bridge Priority	32768

Table 6-1. Default Settings (Continued)

Feature	Default
Multiple Spanning Tree	Disabled
Link Aggregation	No LAGs configured
LACP System Priority	1
Routing Mode	Disabled
OSPF Admin Mode	Disabled
OSPF Router ID	0.0.0.0
IP Helper and UDP Relay	Disabled
RIP	Disabled
VRRP	Disabled
Tunnel and Loopback Interfaces	None
IPv6 Routing	Disabled
DHCPv6	Disabled
OSPFv3	Disabled
DiffServ	Enabled
Auto VoIP	Disabled
Auto VoIP Traffic Class	6
PFC	Disabled; no classifications configured.
DCBx version	Auto detect
iSCSI	Enabled
MLD Snooping	Enabled
IGMP Snooping	Enabled
IGMP Snooping Querier	Disabled
GMRP	Disabled
IPv4 Multicast	Disabled
IPv6 Multicast	Disabled

7

Setting the IP Address and Other Basic Network Information

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure basic network information for the switch, such as the IP address, subnet mask, and default gateway. The topics in this chapter include:

- IP Address and Network Information Overview
- Default Network Information
- Configuring Basic Network Information (Web)
- Configuring Basic Network Information (CLI)
- Basic Network Information Configuration Examples

IP Address and Network Information Overview

What Is the Basic Network Information?

The basic network information includes settings that define the Dell Networking N1500, N2000, N3000, and N4000 Series switches in relation to the network. Table 7-1 provides an overview of the settings this chapter describes.

Table 7-1. Basic Network Information

Feature	Description
IP Address	On an IPv4 network, the a 32-bit number that uniquely identifies a host on the network. The address is expressed in dotted-decimal format, for example 192.168.10.1.
Subnet Mask	Determines which bits in the IP address identify the network, and which bits identify the host. Subnet masks are also expressed in dotted-decimal format, for example 255.255.255.0.

Table 7-1. Basic Network Information (Continued)

Feature	Description
Default Gateway	Typically a router interface that is directly connected to the switch and is in the same subnet. The switch sends IP packets to the default gateway when it does not recognize the destination IP address in a packet.
DHCP Client	Requests network information from a DHCP server on the network.
Domain Name System (DNS) Server	Translates hostnames into IP addresses. The server maintains a domain name databases and their corresponding IP addresses.
Default Domain Name	Identifies your network, such as dell.com. If a hostname is entered without the domain name information, the default domain name is automatically appended to the hostname.
Host Name Mapping	Allows statically mapping an IP address to a hostname.

Additionally, this chapter describes how to view host name-to-IP address mappings that have been dynamically learned by the system.

Why Is Basic Network Information Needed?

Dell Networking N-Series switches are layer-2/3 managed switches. To manage the switch remotely by using a web browser or Telnet client, the switch must have an IP address, subnet mask, and default gateway. A username and password is required to be able to log into the switch from a remote host. For information about configuring users, see "Authentication, Authorization, and Accounting " on page 229. If managing the switch by using the console connection only, configuring an IP address and user is not required. In this case, disabling the Telnet server using the **no ip telnet** command is recommended.

NOTE: The configuration example in this chapter includes commands to create an administrative user with read/write access.

Configuring the DNS information, default domain name, and host name mapping help the switch identify and locate other devices on the network and on the Internet. For example, to upgrade the switch software by using a TFTP server on the network, the TFTP server must be identified. If configuring the switch to use a DNS server to resolve hostnames into IP addresses, it is possible to enter the hostname of the TFTP server instead of the IP address. It is often easier to remember a hostname than an IP address, and if the IP address is dynamically assigned, it might change from time-to-time.

How Is Basic Network Information Configured?

A console-port connection is required to perform the initial switch configuration. When booting the switch for the first time, if there is no startup configuration file, the Dell Easy Setup Wizard starts. The Dell Easy Setup Wizard is a CLI-based tool to help the administrator perform the initial switch configuration. If no response to the Dell Easy Setup Wizard prompt is received within 60 seconds, the `console>` prompt appears, and the switch enters User Configuration mode.

For more information about performing the initial switch configuration by using the wizard, see the *Getting Started Guide* at www.dell.com/support.

If the wizard is not used to supply the initial configuration information, the administrator can manually enable the DHCP client on the switch to obtain network information from a DHCP server via the in-band ports or the out-of-band port. Alternatively, the network configuration can be statically configured.

After configuring the switch with an IP address and creating a user account, continue to use the console connection to configure basic network information, or log on to the switch by using a Telnet client or a web browser. It is possible at this point to change the IP address information and configure additional network information from the remote system.

What Is Out-of-Band Management and In-Band Management?

The Dell Networking N3000 and N4000 Series switches have an external port intended solely for management of the switch. This port is the out-of-band (OOB) management port. Traffic received on the OOB port is never switched or routed to any in-band port and is not rate limited. Likewise, traffic received on any in-band port is never forwarded or routed over the OOB port. The only applications available on the OOB port are protocols required to manage the switch, for example Telnet, SSH, DHCP client, and TFTP. If using the out-of-band management port, it is strongly recommended that the port be connected only to a physically isolated secure management network. The OOB port is a layer-3 interface that uses an internal non-user-configurable VLAN.

The out-of-band port is a logical management interface. The IP stack's routing table contains both IPv4/IPv6 routes associated with these management interfaces and IPv4/IPv6 routes associated with routing interfaces. If routes to the same destination (such as a default route) are learned or configured on both the OOB interface and a routing interface, the routing interface route is preferred. If a directly connected subnet is configured on an out-of-band interface, it cannot also be configured on an in-band interface. If a default gateway is configured on routing interfaces (front-panel ports), then IP addresses not in the OOB port subnet will not be reachable via the OOB port.

Dell recommends that, if used, the OOB port be used for remote management on a physically independent management network and be assigned an IP address from the non-routable private IP address space. The following list highlights some advantages of using OOB management instead of in-band management:

- Traffic on the OOB port is passed directly to the switch CPU, bypassing the switching silicon. The OOB port is implemented as an independent NIC, which allows direct access to the switch CPU from the management network.
- If the production network is experiencing problems, administrators can still access the switch management interface and troubleshoot issues.
- Because the OOB port is intended to be physically isolated from the production network or deployed behind a firewall, configuration options are limited to just those protocols needed to manage the switch. Limiting the configuration options makes it difficult to accidentally cut off management access to the switch.

Alternatively, network administrators may choose to manage their network via the production network. This is in-band management. Because in-band management traffic is mixed in with production network traffic, it is subject to all of the filtering rules usually applied on a switched/routed port, such as ACLs and VLAN tagging, and may be rate limited to protect against DoS attacks.

The administrator can assign an IPv4 address or IPv6 addresses to the OOB management port and to any VLAN. By default, all ports (other than the OOB port) are members of VLAN 1. If an IP address is assigned to VLAN 1, it is possible to connect to the switch management interface by using any of the front-panel switch ports. This is required to manage the Dell Networking N1500 and N2000 Series switches. The use of VLAN 1 for switch administration presents some security risks. Alternatively, a management VLAN can be assigned as the native VLAN for a limited set of front-panel ports and an IP address can be assigned to that VLAN. The use of ACLs to restrict access to switch management is strongly recommended.

DHCP can be enabled on the OOB interface and VLAN interfaces simultaneously, or they can be configured with static information. To configure static address information on the default VLAN (or the management VLAN), set the IP address and subnet mask on the VLAN interface and configure a global default gateway for the switch. If a default gateway is configured on routing interfaces (front-panel ports), then IP addresses not in the OOB port subnet will not be reachable via the OOB port.

Adjusting the Management Interface MTU

When logging in to the Dell Networking N-Series switch using TCP, the switch negotiates the TCP Maximum Segment Size (MSS) using the minimum of the requested MSS or the MTU setting of the port. TCP packets are transmitted from the switch with the DF (Don't Fragment) bit set in order to receive notification of fragmentation from any transit routers. Upon receiving an ICMP *Destination Unreachable, Fragmentation needed but DF set* notification, the switch will reduce the MSS. However, many firewalls block ICMP Destination Unreachable messages, which causes the destination to request the packet again until the connection times out.

To resolve this issue, reduce the TCP MSS setting to a more appropriate value on the local host or alternatively, set the system MTU to a smaller value.

Default Network Information

NOTE: Dell Networking N1500 and N2000 Series switches do not have an out-of-band interface.

By default, no network information is configured. The DHCP client is enabled on the OOB interface by default on Dell Networking N3000 and N4000 Series switches. The DHCP client is enabled on VLAN 1 by default on the Dell Networking N1500 and N2000 Series switches. DNS is enabled, but no DNS servers are configured. VLAN 1 does not have an IP address, subnet mask, or default gateway configured on Dell Networking N3000 and N4000 Series switches.

Configuring Basic Network Information (Web)

This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring basic network information on the Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

Out-of-Band Interface

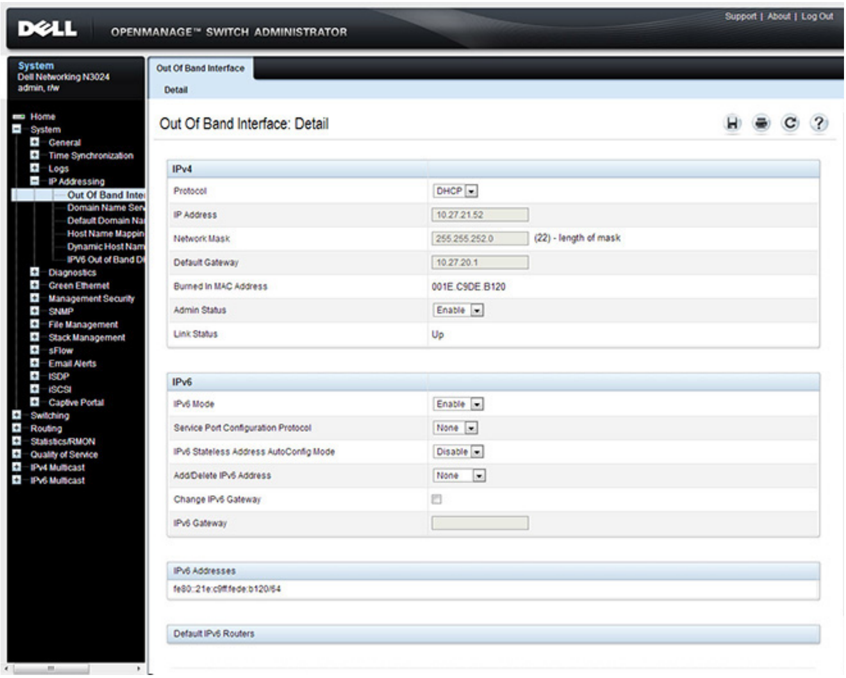
NOTE: Dell Networking N1500 and N2000 Series switches do not have an out-of-band interface.

Use the **Out of Band Interface** page to assign the out-of-band interface IP address and subnet mask or to enable/disable the DHCP client for address information assignment. DHCP is enabled by default on the OOB interface. The OOB interface must be configured on a subnet separate from the front-panel port routing interfaces. The system default gateway must not share an address range/subnet with the OOB interface.

The out-of-band interface may also be assigned an IPv6 address, either statically or via DHCP. In addition, the out-of-band port may be assigned an IPv6 address via the IPv6 auto-configuration process.

To display the **Out of Band Interface** page, click **System → IP Addressing → Out of Band Interface** in the navigation panel.

Figure 7-1. Out of Band Interface



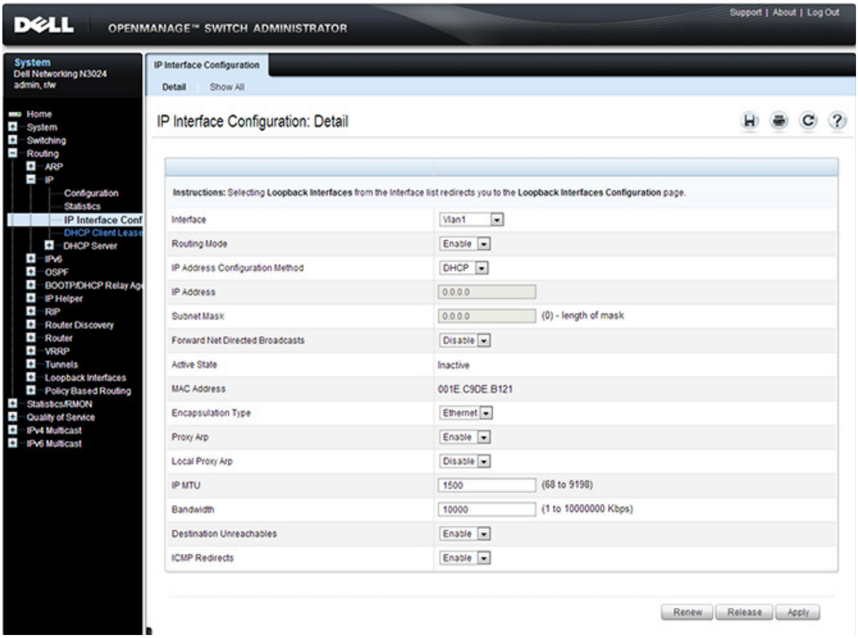
To enable the DHCP client and allow a DHCP server on your network to automatically assign the network information to the OOB interface, select DHCP from the **Protocol** menu. If the network information is statically assigned, ensure that the **Protocol** menu is set to None.

IP Interface Configuration (Default VLAN IP Address)

Use the **IP Interface Configuration** page to assign the default VLAN IP address and subnet mask, the default gateway IP address, and to assign the boot protocol.

To display the **IP Interface Configuration** page, click **Routing** → **IP** → **IP Interface Configuration** in the navigation panel.

Figure 7-2. IP Interface Configuration (Default VLAN)



Assigning Network Information to the Default VLAN

To assign an IP Address and subnet mask to the default VLAN:

1. From the **Interface** menu, select VLAN 1.
2. From the **Routing Mode** field, select **Enable**.
3. From the **IP Address Configuration Method** field specify whether to assign a static IP address (Manual) or use DHCP for automatic address assignment.
4. If **Manual** is selected for the configuration method, then the **IP Address** and **Subnet Mask** can be entered in the appropriate fields.
5. Click **Apply**.

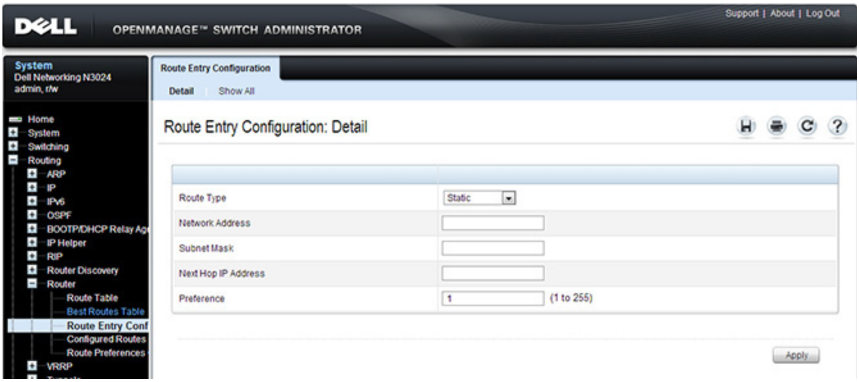
NOTE: No additional fields on the page must be configured. For information about VLAN routing interfaces, see "Routing Interfaces " on page 1141.

Route Entry Configuration (Switch Default Gateway)

Use the **Route Entry Configuration** page to configure the default gateway for the switch. The default VLAN uses the switch default gateway as its default gateway. The switch default gateway must not be on the same subnet as the OOB management port, as the OOB management port cannot route packets received on the front-panel ports.

To display the **Route Entry Configuration** page, click **Routing** → **Router** → **Route Entry Configuration** in the navigation panel.

Figure 7-3. Route Configuration (Default VLAN)



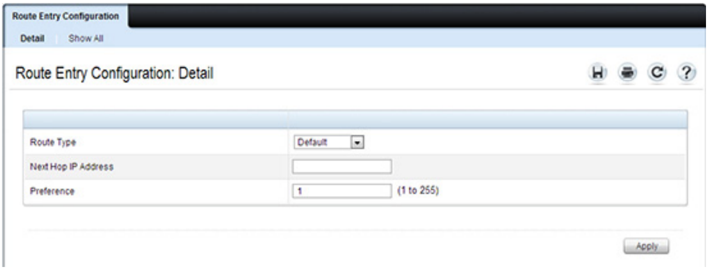
Configuring a Default Gateway for the Switch:

To configure the switch default gateway:

1 Open the **Route Entry Configuration** page.

2 From the **Route Type** field, select Default.

Figure 7-4. Default Route Configuration (Default VLAN)



3 In the **Next Hop IP Address** field, enter the IP address of the default gateway.

4 Click **Apply**.

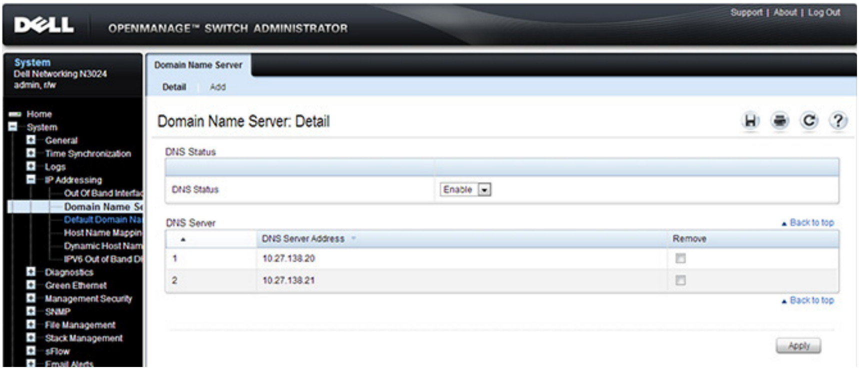
For more information about configuring routes, see "IP Routing " on page 1115.

Domain Name Server

Use the **Domain Name Server** page to configure the IP address of the DNS server. The switch uses the DNS server to translate hostnames into IP addresses.

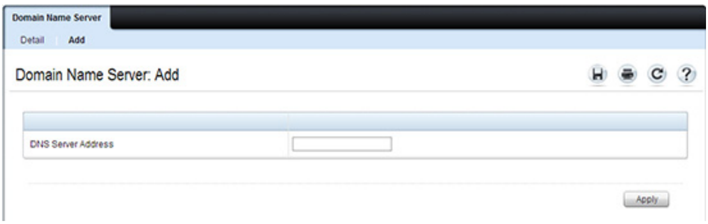
To display the **Domain Name Server** page, click **System** → **IP Addressing** → **Domain Name Server** in the navigation panel.

Figure 7-5. DNS Server



To configure DNS server information, click the **Add** link and enter the IP address of the DNS server in the available field.

Figure 7-6. Add DNS Server

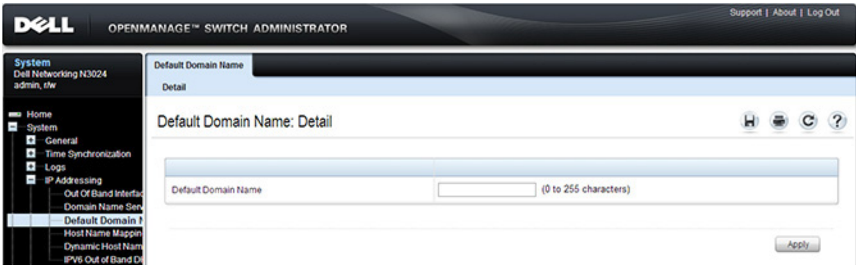


Default Domain Name

Use the **Default Domain Name** page to configure the domain name the switch adds to a local (unqualified) hostname.

To display the **Default Domain Name** page, click **System** → **IP Addressing** → **Default Domain Name** in the navigation panel.

Figure 7-7. Default Domain Name

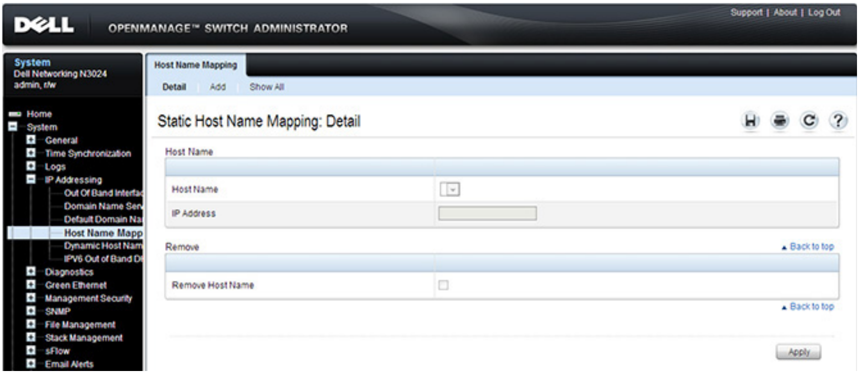


Host Name Mapping

Use the **Host Name Mapping** page to assign an IP address to a static host name. The **Host Name Mapping** page provides one IP address per host.

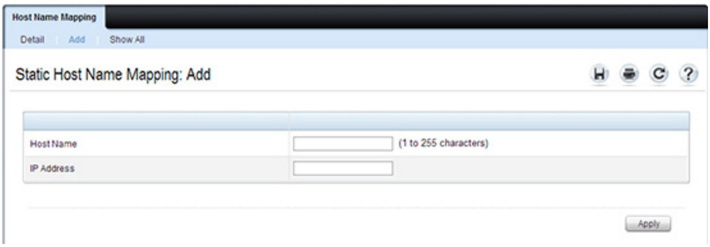
To display the **Host Name Mapping** page, click **System** → **IP Addressing** → **Host Name Mapping.**

Figure 7-8. Host Name Mapping



To map a host name to an IP address, click the **Add** link, type the name of the host and its IP address in the appropriate fields, and then click **Apply**.

Figure 7-9. Add Static Host Name Mapping



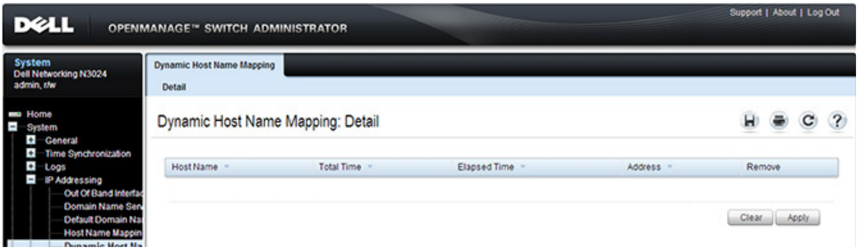
Use the **Show All** link to view all configured host name-to-IP address mappings.

Dynamic Host Name Mapping

Use the **Dynamic Host Name Mapping** page to view dynamic host entries the switch has learned. The switch learns hosts dynamically by using the configured DNS server to resolve a hostname. For example, if you ping **www.dell.com** from the CLI, the switch uses the DNS server to lookup the IP address of **dell.com** and adds the entry to the Dynamic Host Name Mapping table.

To display the **Dynamic Host Name Mapping** page, click **System** → **IP Addressing** → **Dynamic Host Name Mapping** in the navigation panel.

Figure 7-10. View Dynamic Host Name Mapping



Configuring Basic Network Information (CLI)

This section provides information about the commands used for configuring basic network information on the Dell Networking N1500, N2000, N3000, and N4000 Series switches. For more information about these commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Enabling the DHCP Client on the OOB Port

NOTE: Dell Networking N1500 and N2000 Series switches do not have an out-of-band interface.

Beginning in Privileged EXEC mode, use the following commands to enable the DHCP client on the OOB port.

Command	Purpose
configure	Enter Global Configuration mode.
interface out-of-band	Enter Interface Configuration mode for the OOB port.
ip address dhcp	Enable the DHCP client.
CTRL + Z	Exit to Privileged EXEC mode.
show ip interface out-of-band	Display network information for the OOB port.

Enabling the DHCP Client on the Default VLAN

Beginning in Privileged EXEC mode, use the following commands to enable the DHCP client on the default VLAN, which is VLAN 1. As a best practice, it is recommended that a separate VLAN other than one used for client traffic be used for in-band management of the switch. In general, using VLAN 1, or any other VLAN carrying client traffic, for in-band management introduces a security vulnerability.

Command	Purpose
configure	Enter Global Configuration mode.
interface vlan 1	Enter Interface Configuration mode for VLAN 1.

Command	Purpose
ip address dhcp	Enable the DHCP client.
ipv6 address dhcp	Enable the DHCPv6 client.
CTRL + Z	Exit to Privileged EXEC mode.
show ip interface vlan 1	Display network information for VLAN 1.

Managing DHCP Leases

Beginning in Privileged EXEC mode, use the following commands to manage and troubleshoot DHCP leases on the switch.

Command	Purpose
show dhcp lease interface [*interface*]	Display IPv4 addresses leased from a DHCP server.
show ipv6 dhcp interface vlan [*interface*]	Display information about the IPv6 DHCP information for all interfaces or for the specified interface.
debug dhcp packet	Display debug information about DHCPv4 client activities and to trace DHCPv4 packets to and from the local DHCPv4 client.
debug ipv6 dhcp	Display debug information about DHCPv6 client activities and to trace DHCPv6 packets to and from the local DHCPv6 client.
ipv6 address {[*prefix/prefixlen*] | **autoconfig** | **dhcp**}	Set the IPv6 address of the management interface or enables auto-configuration or DHCP.
ipv6 gateway *ipv6-address*	Set the IPv6 default gateway address.
ipv6 enable	Enable IPv6 functionality on the interface.
show ipv6 interface out-of-band	Show settings for the interface.

Configuring Static Network Information on the OOB Port

NOTE: Dell Networking N1500 and N2000 Series switches do not have an out-of-band interface.

Beginning in Privileged EXEC mode, use the following commands to configure a static IP address, subnet mask, and default gateway on the OOB port. If no default gateway is configured, then the zero subnet (0.0.0.0) is used. In this configuration, the OOB port can reach hosts in the local subnet only, because the OOB port will not be able to issue ARP requests to the default gateway. Configuring a default gateway address on the OOB port allows the OOB port to issue ARPs and address traffic to hosts on other subnets. The OOB port subnet may not overlap with any in-band VLAN subnet.

Command	Purpose
configure	Enter Global Configuration mode.
interface out-of-band	Enter Interface Configuration mode for the OOB port.
ip address *ip_address subnet_mask* [*gateway_ip*]	Configure a static IP address and subnet mask. Optionally, a default gateway can also be configured.
ipv6 address *prefix/prefix-length*	Configure an IPv6 prefix for the OOB port
ipv6 address enable	Enable IPv6 addressing on the OOB port
ipv6 address autoconfig	Enable IPv6 auto-configuration for the OOB port
ipv6 address dhcp	Enable DHCP address assignment for the OOB port.
CTRL + Z	Exit to Privileged EXEC mode.
show ip interface out-of-band	Verify the network information for the OOB port.

NOTE: The out-of-band port also supports IPv6 address assignment, including IPv6 auto-configuration and an IPv6 DHCP client.

Configuring Static Network Information on the Default VLAN

Beginning in Privileged EXEC mode, use the following commands to configure a static IP address, subnet mask, and default gateway on the default VLAN. Alternatively, a DHCP server may be used to obtain a network address. The switch also supports IPv6 address auto-configuration.

IP subnets on in-band ports (configured on switch VLANs) may not overlap with the OOB port subnet. If configuring management access on the front-panel ports, it is recommended that:

- A VLAN other than the default VLAN be used to avoid attack vectors enabled by incorrect cabling.
- Both ACLs and Management ACLs be utilized on front-panel ports to reduce the possibility of DoS attacks or intruders gaining access to the switch management console. Management ACLs provide software filtering with deep inspection of packets, whereas ACLs provide hardware filtering with a more limited set of capabilities.

Command	Purpose
configure	Enter Global Configuration mode.
vlan 10	Create a management VLAN and enter VLAN Configuration mode.
exit	Exit VLAN Configuration mode
interface vlan 10	Enter Interface Configuration mode for VLAN 10. VLAN 10 is the management VLAN.
ip address *ip_address subnet_mask*	Enter the IP address and subnet mask.
ipv6 address *prefix/prefix-length* [**eui64**]	Enter the IPv6 address and prefix.
ipv6 enable	Enable IPv6 on the interface.
exit	Exit to Global Configuration mode
ip default-gateway *ip_address*	Configure the default gateway.
ipv6 gateway *ip_address*	Configure the default gateway for IPv6.
exit	Exit to Privileged Exec mode.

Command	Purpose
show ip interface vlan 10	Verify the network information for VLAN 10.
show ipv6 interface vlan 10	Verify IPv6 network information for VLAN 10.
interface Gi1/0/24	Enter physical Interface Configuration mode for the specified interface.
switchport access vlan 10	Allow access to the management VLAN over this port.
exit	Exit Interface Configuration mode.

Configuring and Viewing Additional Network Information

Beginning in Privileged EXEC mode, use the following commands to configure a DNS server, the default domain name, and a static host name-to-address entry. Use the **show** commands to verify configured information and to view dynamic host name mappings. Remember to assign VLANs to interfaces.

Command	Purpose
configure	Enter Global Configuration mode.
ip domain-lookup	Enable IP DNS-based host name-to-address translation.
ip name-server *ip_address*	Enter the IP address of an available name server to use to resolve host names and IP addresses. Up to eight DNS servers may be configured.
ip domain-name *name*	Define a default domain name to complete unqualified host names.
ip host *name ip_address*	Use to configure static host name-to-address mapping in the host cache.
ip address-conflict-detect run	Trigger the switch to run active address conflict detection by sending gratuitous ARP packets for IPv4 addresses on the switch.
CTRL + Z	Exit to Privileged EXEC mode.
show ip interface vlan 1	Verify the network information for VLAN 1.
show hosts	Verify the configured network information and view the dynamic host mappings.

Command	Purpose
show ip address-conflict	View the status information corresponding to the last detected address conflict.
clear ip address-conflict-detect	Clear the address conflict detection status in the switch.

Basic Network Information Configuration Examples

Configuring Network Information Using the OOB Port

In this example, an administrator at a Dell office in California decides not to use the Dell Easy Setup Wizard to perform the initial switch configuration. The administrator configures Dell Networking N3000 and N4000 Series switches to obtain information from a DHCP server on the management network and creates the administrative user with read/write access. The administrator also configures the following information:

- Primary DNS server: 10.27.138.20
- Secondary DNS server: 10.27.138.21
- Default domain name: sunny.dell.com

The administrator also maps the administrative laptop host name to its IP address. The administrator uses the OOB port to manage the switch.

To configure the switch:

NOTE: Dell Networking N1500 and N2000 Series switches do not have an out-of-band interface.

1 Connect the OOB port to the management network. DHCP is enabled by on the switch OOB interface by default on Dell Networking N3000 and N4000 Series switches. DHCP is enabled on VLAN 1 on the Dell Networking N1500/N2000 Series switches, as they do not support an OOB interface. If the DHCP client on the switch has been disabled, use the following commands to enable the DHCP client on the OOB port.

```
console#configure
console(config)#interface out-of-band
console(config-if)#ip address dhcp
console(config-if)#exit
```

2 Configure the administrative user.

```
console(config)#username admin password secret123 privilege 15
```

3 Configure the DNS servers, default domain name, and static host mapping.

```
console(config)#ip name-server 10.27.138.20 10.27.138.21
```

```
console(config)#ip domain-name sunny.dell.com
console(config)#ip host admin-laptop 10.27.65.103
console(config)#exit
```

4 View the network information that the DHCP server on the network dynamically assigned to the switch.

```
console#show ip interface out-of-band

IP Address....................... 10.27.22.153
Subnet Mask..................... 255.255.255.0
Default Gateway................. 10.27.22.1
Protocol Current................ DHCP
Burned In MAC Address........... 001E.C9AA.AA08
```

5 View additional network information.

```
console#show hosts

Host name:
Default domain: sunny.dell.com dell.com
Name/address lookup is enabled
Name servers (Preference order): 10.27.138.20, 10.27.138.21
Configured host name-to-address mapping:
Host                      Addresses
-----------  ------------------------------------
admin-laptop              10.27.65.103

cache: TTL (Hours)

Host          Total   Elapsed Type      Addresses
--------- ------- ------- --------- ---------
No hostname is mapped to an IP address
```

6 Verify that the static hostname is correctly mapped.

```
console#ping admin-laptop
 Pinging admin-laptop with 0 bytes of data:

Reply From 10.27.65.103: icmp_seq = 0. time <10 msec.
Reply From 10.27.65.103: icmp_seq = 1. time <10 msec.
```

Configuring Network Information Using the Serial Interface

In this example, the administrator configures a Dell Networking N1500/N2000 Series switch via the serial interface while using the same DHCP server and address configuration as given in the previous example.

1 Connect a front-panel port (e.g., gi1/0/24) to the management network. Use the following commands to create a management VLAN, disable DHCP on VLAN 1, and disable L3 addressing on VLAN 1, and enable the DHCP client on the management VLAN.

```
console#configure
console(config)#vlan 4093
console(config-vlan4093)#interface vlan 1
console(config-if-vlan1)#no ip address
console(config-if-vlan1)#exit
console(config)#no interface vlan 1
console(config-)#interface vlan 4093
console(config-if-vlan4093)#ip address dhcp
```

2 Assign the management VLAN to an interface connected to the management network.

```
console(config-if-vlan4093)#interface gi1/0/24
console(config-if-Gi1/0/24)#switchport access vlan 4093
console(config-if-Gi1/0/24)#exit
```

3 Configure the administrative user.

```
console(config)#username admin password secret123 privilege 15
```

4 Configure the DNS servers, default domain name, and static host mapping.

```
console(config)#ip name-server 10.27.138.20 10.27.138.21
console(config)#ip domain-name sunny.dell.com
console(config)#ip host admin-laptop 10.27.65.103
console(config)#exit
```

5 View the network information that the DHCP server on the network dynamically assigned to the switch.

```
console#show ip interface vlan 4093

Routing interface status....................... Up
Primary IP Address..... 10.27.22.150/255.255.252.0
Method.................................... DHCP
Routing Mode.............................. Enable
Administrative Mode....................... Enable
Forward Net Directed Broadcasts.......... Disable
Proxy ARP................................. Enable
Local Proxy ARP.......................... Disable
Active State.............................. Active
MAC Address...................... 001E.C9DE.B77A
Encapsulation Type..................... Ethernet
IP MTU.................................... 1500
```

```
Bandwidth.............................. 10000 kbps
Destination Unreachables.................. Enabled
ICMP Redirects............................ Enabled
```

Refer to the Access Control Lists section for information on restricting access to the switch management interface.

8

Managing QSFP Ports

Dell Networking N4000 Series Switches

QSFP ports available on Dell Networking N4000 Series switches can operate in 1 x 40G mode or in 4 x 10G mode. Appropriate cables must be used that match the selected mode. When changing from one mode to another, a switch reboot is required. The QSFP ports also support stacking over the interfaces in either 1 x 40G or 4 x 10G mode. Changing from Ethernet mode to stacking mode and vice-versa requires a reboot as well.

The ports on a QSFP plugin module are named Fo1/1/1-2 in 40-gigabit mode and Te1/1/1-8 in 10-gigabit mode. On the N4064, the fixed QSFP ports are named Fo1/0/1-2 in 40-gigabit mode and Te1/0/49-56 in 10-gigabit mode. All of the possible populated or configured interfaces will show in the **show interfaces status** command regardless of the port mode, i.e. 40-gigabit or 10-gigabit. Unpopulated or unconfigured interfaces for plug in modules do not show in the **show interfaces status** command.

The default setting for a 40-gigabit Ethernet interface is nonstacking, 40-gigabit Ethernet (1 x 40G).

The commands to change 1 x 40G and 4 x 10G modes are always entered on the 40-gigabit interfaces.

The commands to change the Ethernet/stack mode are entered on the appropriate interface (tengigabitethernet or fortygigabitethernet). It is possible to configure some of the 10G ports in a 40G interface as stacking and not others.

To reconfigure a QSFP port, select the 40-gigabit port to change in Interface Config mode and enter the **hardware profile portmode** command with the selected mode. For example, to change a 1 x 40G port to 4 x 10G mode, enter the following commands on the forty-gigabit interface:

```
console(config)#interface fo1/1/1
console(config-if-Fo1/1/2)#hardware profile portmode 4x10g

This command will not take effect until the switch is rebooted.

console(config-if-Fo1/1/2)#do reload
```

```
Are you sure you want to reload the stack? (y/n)
```

To change a 4 x 10G port to 1 x 40G mode, enter the following commands on the 40-gigabit interface:

```
console(config)#interface Fo2/1/1
console(config-if-Fo2/1/1)#hardware profile portmode 1x40g

This command will not take effect until the switch is rebooted.

console(config-if-Fo1/1/2)#do reload

Are you sure you want to reload the stack? (y/n)
```

Attempting to change the port mode on the tengigabit interface will give the error “An invalid interface has been used for this function.”

9

Stacking

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure and manage a stack of switches.

The topics covered in this chapter include:

- Stacking Overview
- Default Stacking Values
- Managing and Monitoring the Stack (Web)
- Managing the Stack (CLI)
- Stacking and NSF Usage Scenarios

Stacking Overview

The Dell Networking N2000, N3000, and N4000 Series switches include a stacking feature that allows up to 12 switches to operate as a single unit. The Dell Networking N1500 Series switches stack up to four units using 10GB Ethernet links configured as stacking. Dell Networking N2000 and N3000 Series switches have two fixed mini-SAS stacking connectors at the rear. Dell Networking N2000 Series switches only stack with other Dell Networking N2000 Series switches, and Dell Networking N3000 Series switches stack with other Dell Networking N3000 Series switches.

Dell Networking N4000 Series switches stack with other Dell Networking N4000 Series switches over front-panel ports configured for stacking. Dell Networking N1500 Series switches stack using the 10G SFP+ front-panel ports.

Dell Networking N1500 Series switches support high-performance stacking over the 10G front-panel ports, allowing increased capacity to be added as needed, without affecting network performance and providing a single point of management. Up to four Dell Networking N1500 Series switches can be stacked using any 10G port as long as the link bandwidth for parallel stacking links is the same. Note that configuring a 10G port for stacking also configures the adjacent partner 10G port for stacking.

A stack of four 48-port Dell Networking N1500 Series switches has an aggregate throughput capacity of 192 Gbps. Dell Networking N1500 Series stacking links operate at 10 Gbps or 5.2% of total aggregate throughput capacity of a full stack; therefore, it is recommended that operators provision large stacking topologies such that it is unlikely that a significant portion of the stack capacity will transit stacking links. One technique for achieving this is to distribute uplinks evenly across the stack vs. connecting all uplinks to a single stack unit or to adjacent stacking units.

A stack of twelve 48-port Dell Networking N2000 or Dell Networking N3000 Series switches has an aggregate throughput capacity of 576 Gbps. Dell Networking N2000/N3000 Series stacking links operate at 21 Gbps or 3.6% of total aggregate throughput capacity of a full stack; therefore, it is recommended that operators provision large stacking topologies such that it is unlikely that a significant portion of the stack capacity will transit stacking links. One technique for achieving this is to distribute uplinks evenly across the stack vs. connecting all uplinks to a single stack unit or to adjacent stacking units.

Dell Networking N4000 Series switches support high performance stacking over front-panel ports, allowing increased capacity to be added as needed, without affecting network performance and providing a single point of management. Up to twelve Dell Networking N4000 Series switches can be stacked using any port as long as the link bandwidth for parallel stacking links is the same. In other words, all the port types on the Dell Networking N4000 Series switches can be used for stacking. Additional stacking connections can be made between adjacent switch units to increase the stacking bandwidth provided that all redundant stacking links have the same port speed. It is strongly recommended that the stacking bandwidth be kept equal across all stacking connections; that is, avoid mixing single and double stacking connections within a stack. Up to eight redundant stacking links operating at the same speed can be configured on a Dell Networking N4000 Series stack unit (four in each direction).

A stack of twelve Dell Networking N4000 Series switches has an aggregate front panel capacity of 5.760 terabits (not including the 40G ports). Provisioning for 5% inter-stack capacity requires 280 gigabits of bandwidth dedicated to stacking or all four 40G ports plus another twelve 10G ports. Therefore, it is recommended that operators provision large stacking topologies such that it is unlikely that a significant portion of the stack capacity will transit stacking links. One technique for achieving this is to distribute downlinks and transit links evenly across the stack vs. connecting all downlinks/transit links to a single stack unit or to adjacent stacking units.

If Priority Flow Control (PFC) is enabled on any port in a Dell Networking N4000 Series stack, stacking is supported at distances up to 100 meters on the stacking ports. If PFC is not enabled, stacking is supported up to the maximum distance supported by the transceiver on the stack links. Note that PFC cannot be enabled on stacking ports — the system handles the buffering and flow control automatically.

A single switch in the stack manages all the units in the stack (the stack master), and the stack is managed by using a single IP address. The IP address of the stack does not change, even if the stack master changes.

A stack is created by daisy-chaining stacking links on adjacent units. If available, up to eight links per stack unit can be used for stacking (four in each direction). A stack of units is manageable as a single entity when the units are connected together. If a unit cannot detect a stacking partner on any port enabled for stacking, the unit automatically operates as a standalone unit. If a stacking partner is detected, the switch always operates in stacking mode. One unit in the stack is designated as the stack master. The master manages all the units in the stack. The stack master runs the user interface and switch software, and propagates changes to the member units. To manage a stack using the serial interface, the administrator must connect to the stack master via the **connect** command or by physically connecting the cable to the stack master.

A second switch is designated as the standby unit, which becomes the master if the stack master is unavailable. The unit to be selected as the standby can be manually configured, or the system can select the standby automatically.

When units are in a stack, the following activities occur:

- All units are checked for software version consistency.
- The switch Control Plane is active only on the master. The Control Plane is a software layer that manages system and hardware configuration and runs the network control protocols to set system configuration and state.
- The switch Data Plane is active on all units in the stack, including the master. The Data Plane is the set of hardware components that forward data packets without intervention from a control CPU.

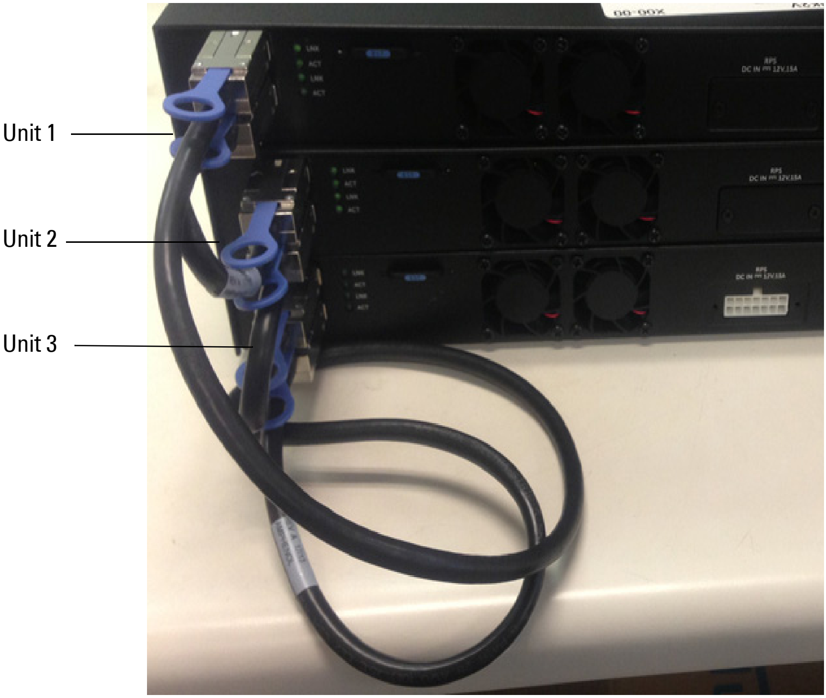
- The running configuration is propagated to all units and the application state is synchronized between the master and standby during normal stacking operation. The startup configuration and backup configuration on the stack members are not overwritten with the master switch configuration.

Dell strongly recommends connecting the stack in a ring topology so that each switch is connected to two other switches. Connecting switches in a ring topology allows the stack to utilize the redundant communication path to each switch. If a switch in a ring topology fails, the stack can automatically establish a new communications path to the other switches. Switches not stacked in a ring topology may split into multiple independent stacks upon the failure of a single switch or stacking link.

Additional stacking connections can be made between adjacent switch units to increase the stacking bandwidth, provided that all redundant stacking links have the same bandwidth. It is strongly recommended that the stacking bandwidth be kept equal across of all stacking connections; that is, avoid mixing single and double stacking connections within a stack. Up to eight redundant stacking links can be configured on a stacking unit (four in each direction).

Figure 9-1 shows a stack with three switches as stack members connected in a ring topology.

Figure 9-1. Connecting a Stack of Switches



The stack in Figure 9-1 has the following physical connections between the switches:

- The lower stacking port on Unit 1 is connected to the upper stacking port on Unit 2.
- The lower stacking port on Unit 2 is connected to the upper stacking port on Unit 3.
- The lower stacking port on Unit 3 is connected to the upper stacking port on Unit 1.

Dell Networking N1500, N2000, N3000, and N4000 Stacking Compatibility

Dell Networking N1500, N2000, N3000, and N4000 Series switches do not stack with different Dell Networking Series switches or other Dell switches. Dell Networking N1500 Series switches only stack with other Dell Networking

N1500 Series switches. Dell Networking N2000 Series switches only stack with other Dell Networking N2000 Series switches. Likewise, Dell Networking N3000 Series switches only stack with other Dell Networking N3000 Series switches. Dell Networking N4000 Series switches only stack with other Dell Networking N4000 Series switches.

How is the Stack Master Selected?

A stack master is elected or re-elected based on the following considerations, in order:

1. The switch is currently the stack master.
2. The switch has the higher MAC address.
3. A unit is selected as standby by the administrator, and a fail over action is manually initiated or occurs due to stack master failure.

In most cases, a switch that is added to an existing stack will become a stack member, and not the stack master. When a switch is added to the stack, one of the following scenarios takes place regarding the management status of the new switch:

- If the switch has the stack master function enabled but another stack master is already active, then the switch changes its configured stack master value to disabled.
- If the stack master function is unassigned and there is another stack master in the system then the switch changes its configured stack master value to disabled.
- If the stack master function is enabled or unassigned and there is no other stack master in the system, then the switch becomes stack master.
- If the stack master function is disabled, the unit remains a non-stack master.

If the entire stack is powered OFF and ON again, the unit that was the stack master before the reboot will remain the stack master after the stack resumes operation.

The unit number for the switch can be manually configured. To avoid unit-number conflicts, one of the following scenarios takes place when a new member is added to the stack:

- If the switch has a unit number that is already in use, then the unit that is added to the stack changes its configured unit number to the lowest unassigned unit number.
- If the added switch does not have an assigned unit number, then the switch sets its configured unit number to the lowest unassigned unit number.
- If the unit number is configured and there are no other devices using the unit number, then the switch starts using the configured unit number.
- If the switch detects that the maximum number of units already exist in the stack making it unable to assign a unit number, then the switch sets its unit number to *unassigned* and does not participate in the stack.

Adding a Switch to the Stack

When adding a new member to a stack, make sure that only the stack cables, and no network cables, are connected before powering up the new unit. Stack port configuration is stored on the member units. If stacking over Ethernet ports (Dell Networking N4000 Series only), configure the ports on the unit to be added to the stack as stacking ports and power the unit off prior to connecting the stacking cables. Make sure the links are not already connected to any ports of that unit. This is important because if STP is enabled and any links are UP, the STP reconvergence will take place as soon as the link is detected.

After the stack cables on the new member are connected to the stack, the units can be powered up, beginning with the unit directly attached to the currently powered-up unit. Always power up new stack units closest to an existing powered unit first. Do not connect a new member to the stack after it is powered up. Also, do not connect two functional, powered-up stacks together. Hot insertion of units into a stack is not supported.

If a new switch is added to a stack of switches that are powered and running and already have an elected stack master, the newly added switch becomes a stack member rather than the stack master. Use the **boot auto-copy-sw** command on the stack master to enable automatic firmware upgrade of newly added switches. If a firmware mismatch is detected, the newly added switch does not fully join the stack and holds until it is upgraded to the same firmware version as the master switch. After firmware synchronization finishes, the running configuration of the newly added unit is overwritten with the stack master configuration. Stack port configuration is always stored on the local unit and may be updated with preconfiguration information from the stack master when the unit joins the stack.

Information about a stack member and its ports can be pre-configured before the unit is added to the stack. The preconfiguration takes place on the stack master. If there is saved configuration information on the stack master for the newly added unit, the stack master applies the configuration to the new unit; otherwise, the stack master applies the default configuration to the new unit.

Removing a Switch from the Stack

Prior to removing a member from a stack, check that other members of the stack will not become isolated from the stack due to the removal. Check the stack-port error counters to ensure that a stack configured in a ring topology can establish a communication path around the member to be removed.

The main point to remember when removing a unit from the stack is to disconnect all the links on the stack member to be removed. Also, be sure to take the following actions:

- Remove all the STP participating ports and wait to stabilize the STP.
- Remove all the member ports of any Port-Channels (LAGs) so there will not be any control traffic destined to those ports connected to this member.
- Statically re-route any traffic going through this unit.

When a unit in the stack fails, the stack master removes the failed unit from the stack. The failed unit reboots with its original running-config. If the stack is configured in a ring topology, then the stack automatically routes around the failed unit. If the stack is not configured in a ring topology, then the stack may split, and the isolated members will reboot and re-elect a new stack master. No changes or configuration are applied to the other stack members; however, the dynamic protocols will try to reconverge as the topology could change because of the failed unit.

If you remove a unit and plan to renumber the stack, issue a **no member *unit*** command in Stack Configuration mode to delete the removed switch from the configured stack member information.

How is the Firmware Updated on the Stack?

When adding a new switch to a stack, the Stack Firmware Synchronization feature, if enabled, automatically synchronizes the firmware version with the version running on the stack master per the configuration on the master switch. The synchronization operation may result in either upgrade or downgrade of firmware on the mismatched stack member. Use the **boot auto-copy-sw** command to enable stack firmware synchronization.

Upgrading the firmware on a stack of switches is the same as upgrading the firmware on a single switch. After downloading a new image by using the File Download page or **copy** command, the downloaded image is distributed to all the connected units of the stack. For more information about downloading and installing images, see "Images and File Management " on page 469.

What is Stacking Standby?

The standby unit may be pre-configured or automatically selected. If the current stack master fails, the standby unit becomes the stack master. If no switch is pre-configured as the standby unit, the software automatically selects a standby unit from among the existing stack units.

When the failed master resumes normal operation, it joins the stack as a member (not as the master) if the new stack master has already been elected.

The stack master copies its running configuration to the standby unit whenever it changes (subject to some restrictions to reduce overhead). This enables the standby unit to take over the stack operation with minimal interruption if the stack master becomes unavailable.

Operational state synchronization also occurs:

- when the running configuration is saved to the startup configuration on the stack master.
- when the standby unit changes.

What is Nonstop Forwarding?

Networking devices, such as the Dell Networking N-Series switches, are often described in terms of three semi-independent functions called the forwarding plane, the control plane, and the management plane. The forwarding plane forwards data packets and is implemented in hardware. The control plane is the set of protocols that determine how the forwarding plane should forward packets, deciding which data packets are allowed to be forwarded and where they should go. Application software on the stack master acts as the control plane. The management plane is application software running on the stack master that provides interfaces allowing a network administrator to configure the device.

The Nonstop Forwarding (NSF) feature allows the forwarding plane of stack units to continue to forward packets while the control and management planes restart as a result of a power failure, hardware failure, or software fault on the stack master. This type of operation is called nonstop forwarding (NSF). When the stack master fails, only the switch ASICs and processor on the stack master need to be restarted.

To prevent adjacent networking devices from rerouting traffic around the restarting device, the NSF feature uses the following three techniques:

1. A protocol can distribute a part of its control plane across stack units so that the protocol can give the appearance that it is still functional during the restart.
2. A protocol may enlist the cooperation of its neighbors through a technique known as graceful restart.
3. A protocol may simply restart after the failover if neighbors react slowly enough that they will not normally detect the outage.

The NSF feature enables the stack master unit to synchronize the running-config within 60 seconds after a configuration change has been made. However, if a lot of configuration changes happen concurrently, NSF uses a back-off mechanism to reduce the load on the switch. In this case, the stack master will attempt resynchronization no more often than once every 120 seconds.

The **show nsf** command output includes information about when the next running-config synchronization will occur.

Initiating a Failover

The NSF feature allows the administrator to initiate a failover using the **initiate failover** command. This method is preferred over the **reload** *unit* command as it ensures synchronization of the stack master and standby unit.

Initiating a failover reloads the stack master, triggering the standby unit to take over. Before the failover, the stack master pushes application data and other important information to the standby unit. Although the handoff is controlled and causes minimal network disruption, some ephemeral application state is lost, such as pending timers and other pending internal events. Use the **show nsf** command to view the stack checkpoint status prior to reloading a stack member. Do not reload while a checkpoint operation is in progress.

Always check the stack health before failing over to the standby unit. Use the **show switch stack-ports counters** command to verify that the stack ports are up and no errors are present. Resolve any error conditions prior to failing over a stack master. Use the **show switch stack-ports stack-path** command to verify the reachability of all stack units. If any units are not reachable, the stack may split during a failover.

Checkpointing

Switch applications (features) that build up a list of data such as neighbors or clients can significantly improve their restart behavior by remembering this data across a warm restart. This data can either be stored persistently, as in the case of configuration data, or the stack master can checkpoint this data directly to the standby unit active processes, as in the case of operational data.

Use the **show nsf** command to view the stack checkpoint status prior to reloading a stack member. Do not reload while a checkpoint operation is in progress.

The NSF checkpoint service allows the stack master to communicate startup configuration data to the standby unit in the stack. When the stack selects a standby unit, the checkpoint service notifies applications to start a complete checkpoint. After the initial checkpoint is done, applications checkpoint changes to their data every 120 seconds.

NOTE: The switch cannot guarantee that a standby unit has exactly the same data that the stack master has when it fails. For example, the stack master might fail before the checkpoint service gets data to the standby if an event occurs shortly before a failover.

Table 9-1 lists the applications on the switch that checkpoint data and describes the type of data that is checkpointed.

Table 9-1. Applications that Checkpoint Data

Application	Checkpointed Data
ARP	Dynamic ARP entries
Auto VOIP	Calls in progress
Captive Portal	Authenticated clients
DHCP server	Address bindings (persistent)
DHCP snooping	DHCP bindings database
DOT1Q	Internal VLAN assignments
DOT1S	Spanning tree port roles, port states, root bridge, etc.
802.1X	Authenticated clients
DOT3ad	Port states
IGMP/MLD Snooping	Multicast groups, list of router ports, last query data for each VLAN
IPv6 NDP	Neighbor cache entries
iSCSI	Connections
LLDP	List of interfaces with MED devices attached
OSPFv2	Neighbors and designated routers
OSPFv3	Neighbors and designated routers
Route Table Manager	IPv4 and IPv6 dynamic routes
SIM	The system's MAC addresses. System up time. IP address, network mask, default gateway on each management interface, DHCPv6 acquired IPv6 address.
Voice VLAN	VoIP phones identified by CDP or DHCP (not LLDP)

Switch Stack MAC Addressing and Stack Design Considerations

The switch stack uses the MAC addresses assigned to the stack master.

NOTE: Each switch is assigned four consecutive MAC addresses. A stack of switches uses the MAC addresses assigned to the stack master.

If the backup unit assumes control due to a stack master failure or warm restart, the backup unit continues to use the original stack master's MAC addresses. This reduces the amount of disruption to the network because ARP and other layer-2 entries in neighbor tables remain valid after the failover to the backup unit.

Stack units should always be connected with a ring topology (or other redundant topology), so that the loss of a single stack link does not divide the stack into multiple stacks. If a stack is partitioned such that some units lose all connectivity to other units, then both parts of the stack start using the same MAC addresses. This can cause severe problems in the network.

If removing the stack master from a stack for use in a different place in the network, make sure to power down the whole stack before redeploying the stack master so that the stack members do not continue to use the MAC address of the redeployed master switch.

NSF Network Design Considerations

A network can be designed to take maximum advantage of NSF. For example, by distributing a LAG's member ports across multiple units, the stack can quickly switch traffic from a port on a failed unit to a port on a surviving unit. When a unit fails, the forwarding plane of surviving units removes LAG members on the failed unit so that it only forwards traffic onto LAG members that remain up. If a LAG is left with no active members, the LAG goes down. To prevent a LAG from going down, configure LAGs with members on multiple units within the stack, when possible. If a stack unit fails, the system can continue to forward on the remaining members of the stack.

If the switch stack performs VLAN routing, another way to take advantage of NSF is to configure multiple “best paths” to the same destination on different stack members. If a unit fails, the forwarding plane removes Equal Cost Multipath (ECMP) next hops on the failed unit from all unicast forwarding table entries. If the cleanup leaves a route without any next hops, the route is deleted. The forwarding plane only selects ECMP next hops on surviving units. For this reason, try to distribute links providing ECMP paths across multiple stack units.

Why is Stacking Needed?

Stacking increases port count without requiring additional configuration. If you have multiple Dell Networking N-Series switches, stacking them helps make management of the switches easier because you configure the stack as a single unit and do not need to configure individual switches.

Default Stacking Values

Stacking is always enabled on Dell Networking N-Series switches.

On the Dell Networking N1500/N4000 Series switches, by default, the 10G SFP+ ports are in Ethernet mode and must be configured to be used as stacking ports. Ports that are configured in stacking mode show as "detached" in the output of the **show interfaces status** command.

NOTE: N1500 10G SFP+ ports may only be configured as stacking in pairs.

Configuring an Ethernet port as a stacking port changes the default configuration of the port. The port stacking configuration does not show in the running-config. To determine the stacking configuration of a port, use the **show switch stack-ports** command. On the Dell Networking N2000/N3000 Series switches, there are two fixed stacking ports in the rear of the switch. Stacking on Ethernet ports is not supported. The fixed stacking ports show as TwentygigabitStacking and are abbreviated Tw.

NSF is enabled by default. NSF can be disabled to redirect the CPU resources consumed by data checkpointing; however, this is ill-advised, as checkpointing consumes almost no switch resources. Checkpointing only occurs when a backup unit is elected, so there is no need to disable the NSF feature on a standalone switch. When a new unit is added to the stack, the new unit is given the configuration of the stack, including the NSF setting. OSPF implements a separate graceful restart control that enables NSF for OSPF. OSPF graceful restart is not enabled by default.

Managing and Monitoring the Stack (Web)

This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring stacking on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click at the top of the page.

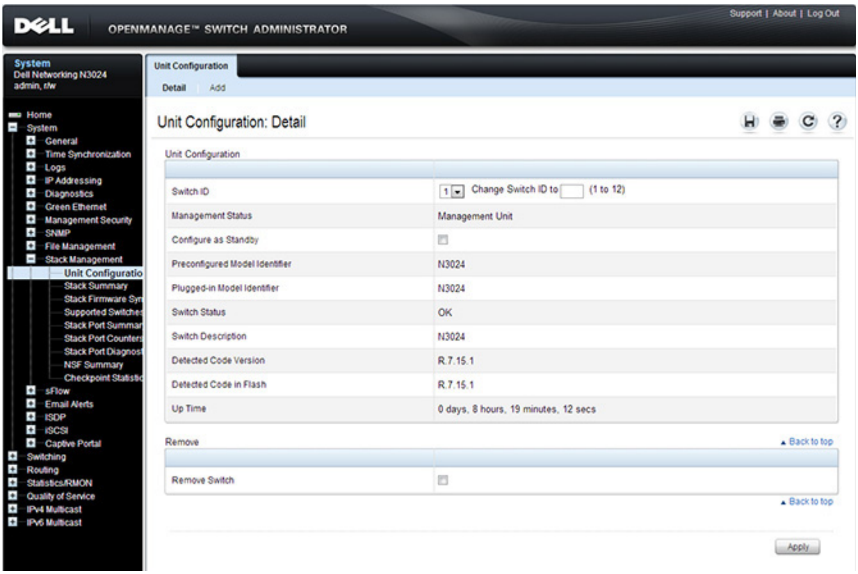
NOTE: Changes made on the Stacking configuration pages take effect only after the device is reset.

Unit Configuration

Use the **Unit Configuration** page to change the unit number and unit type (Management, Member, or Standby).

To display the **Unit Configuration** page, click **System** → **Stack Management** → **Unit Configuration** in the navigation panel.

Figure 9-2. Stack Unit Configuration



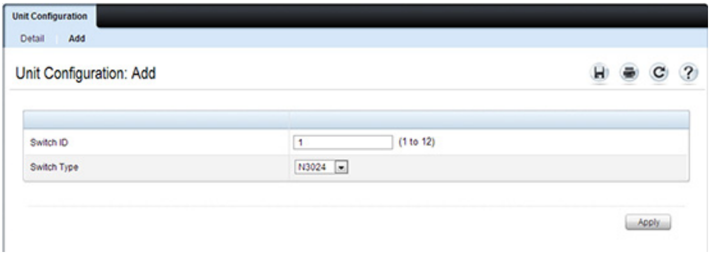
Changing the ID or Switch Type for a Stack Member

To change the switch ID or type:

1 Open the **Unit Configuration** page.

2 Click **Add** to display the **Add Unit** page.

Figure 9-3. Add Remote Log Server Settings



3 Specify the switch ID, and select the model number of the switch.

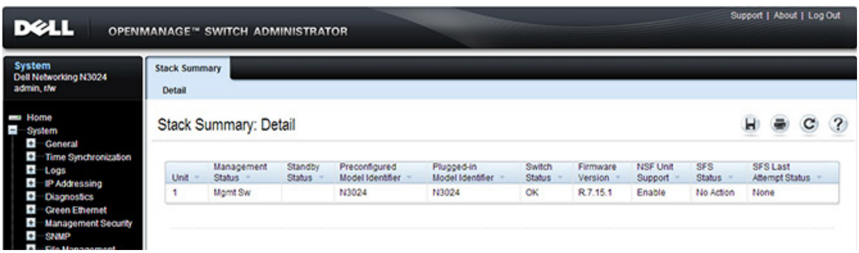
4 Click **Apply**.

Stack Summary

Use the **Stack Summary** page to view a summary of switches participating in the stack.

To display the **Stack Summary** page, click **System** → **Stack Management** → **Stack Summary** in the navigation panel.

Figure 9-4. Stack Summary

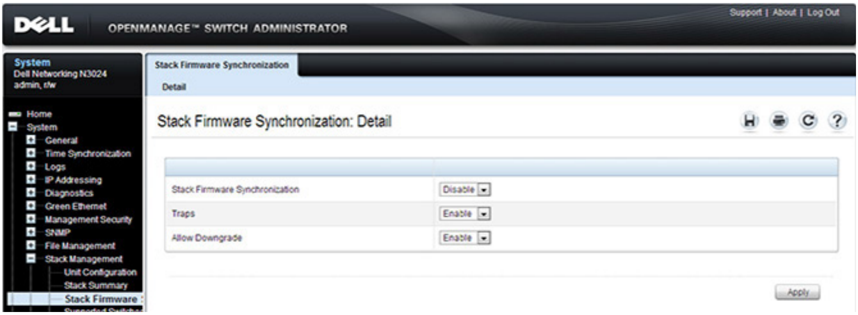


Stack Firmware Synchronization

Use the **Stack Firmware Synchronization** page to control whether the firmware image on a new stack member can be automatically upgraded or downgraded to match the firmware image of the stack master.

To display the **Stack Firmware Synchronization** page, click **System** → **Stack Management** → **Stack Firmware Synchronization** in the navigation panel.

Figure 9-5. Stack Firmware Synchronization

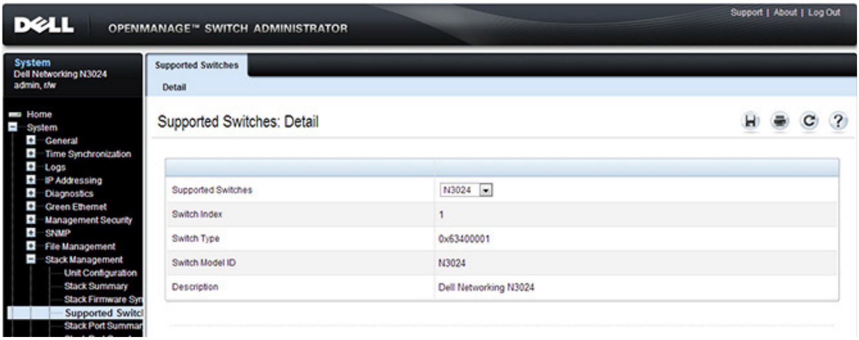


Supported Switches

Use the **Supported Switches** page to view information regarding each type of supported switch for stacking, and information regarding the supported switches.

To display the **Supported Switches** page, click **System** → **Stack Management** → **Supported Switches** in the navigation panel.

Figure 9-6. Supported Switches



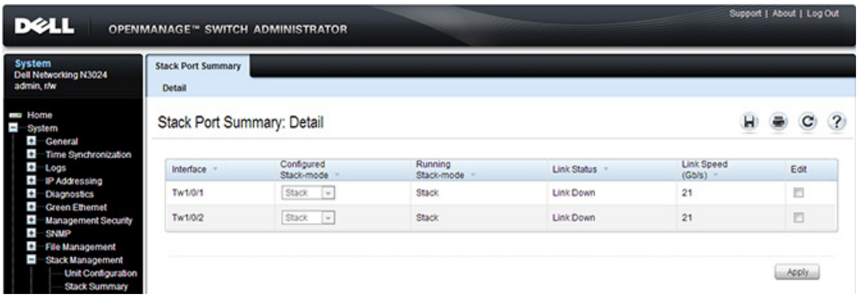
Stack Port Summary

Use the **Stack Port Summary** page to configure the stack-port mode and to view information about the stackable ports. This screen displays the unit, the stackable interface, the configured mode of the interface, the running mode as well as the link status and link speed of the stackable port.

NOTE: By default the ports are configured to operate as Ethernet ports. To configure a port as a stack port, the Configured Stack Mode setting must be changed from Ethernet to Stack.

To display the **Stack Port Summary** page, click **System** → **Stack Management** → **Stack Port Summary** in the navigation panel.

Figure 9-7. Stack Port Summary

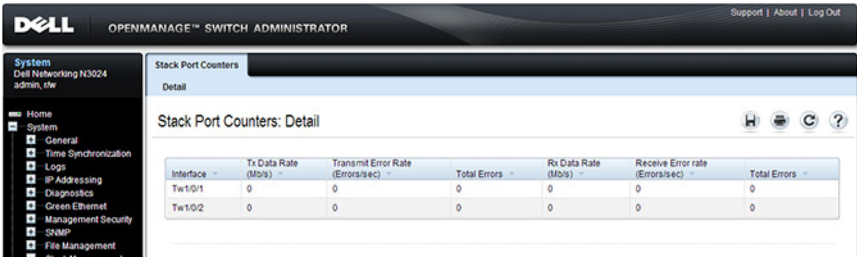


Stack Port Counters

Use the **Stack Port Counters** page to view the transmitted and received statistics, including data rate and error rate.

To display the **Stack Port Counters** page, click **System** → **Stack Management** → **Stack Point Counters** in the navigation panel.

Figure 9-8. Stack Port Counters



Stack Port Diagnostics

The **Stack Port Diagnostics** page is intended for Field Application Engineers (FAEs) and developers only.

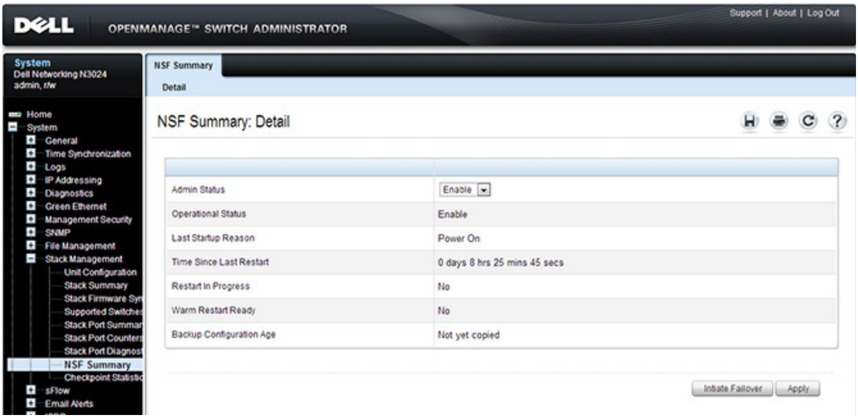
NSF Summary

Use the **NSF Summary** page to change the administrative status of the NSF feature and to view NSF information.

NOTE: The OSPF feature uses NSF to enable the hardware to continue forwarding IPv4 packets using OSPF routes while a backup unit takes over stack master responsibility. To configure NSF on a stack that uses OSPF or OSPFv3, see "NSF OSPF Configuration " on page 1207 and "NSF OSPFv3 Configuration " on page 1224.

To display the **NSF Summary** page, click **System** → **Stack Management** → **NSF Summary** in the navigation panel.

Figure 9-9. NSF Summary



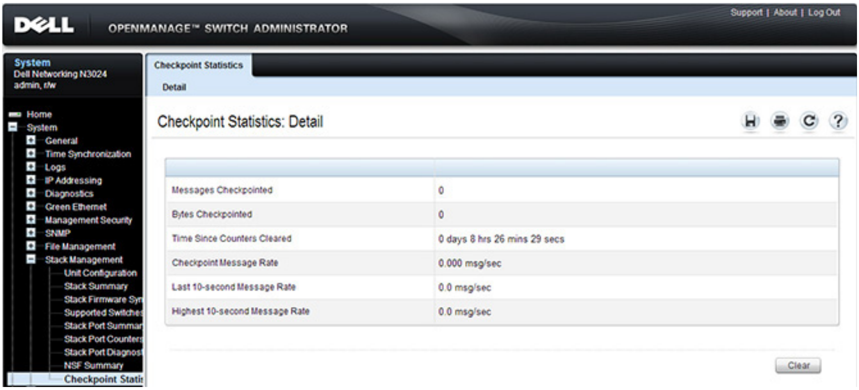
To cause the maser unit to failover to the standby unit, click **Initiate Failover**. The failover results in a warm restart of the stack master. Initiating a failover reloads the stack master, triggering the backup unit to take over.

Checkpoint Statistics

Use the **Checkpoint Statistics** page to view information about checkpoint messages generated by the stack master.

To display the **Checkpoint Statistics** page, click **System** → **Stack Management** → **Checkpoint Statistics** in the navigation panel.

Figure 9-10. Checkpoint Statistics



Managing the Stack (CLI)

This section provides information about the commands for managing the stack and viewing information about the switch stack. For more information about these commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Stack Member, Stack Port, and NSF Settings

Beginning in Privileged EXEC mode, use the following commands to configure stacking and NSF settings.

Command	Purpose
configure	Enter Global Configuration mode.
switch *current_ID* **renumber** *new_ID*	Change the switch ID number. The valid range is 1-10. **NOTE:** Changing the ID number causes all switches in the stack to be reset to perform stack master renumbering. The running configuration is cleared when the units reset.
stack	Enter Global Stack Configuration mode.
initiate failover	Move the management switch functionality from the master switch to the standby switch.
standby *unit*	Specify the stack member that will come up as the master if a stack failover occurs.
set description *unit*	Configure a description for the specified stack member.

Command	Purpose
member *unit SID*	Add a switch to the stack and specify the model of the new stack member. • *unit* - The switch unit ID • *SID* - The index into the database of the supported switch types, indicating the type of the switch being pre-configured. **Note**: Member configuration displayed in the running config may be learned from the physical stack. Member configuration is not automatically saved in the startup configuration. Save the configuration to retain the current member settings. To view the SID associated with the supported switch types, use the **show supported switchtype** command in Privileged EXEC mode.
stack-port tengigabitethernet *unit/slot/port* {**ethernet** | **stack**}	Set the mode of the port to either Ethernet or stacking (Dell Networking N4000 Series only).
nsf	Enable nonstop forwarding on the stack. (Enabled by default.)
exit	Exit to Global Config mode.
boot auto-copy-sw	Enable the Stack Firmware Synchronization feature.
boot auto-copy-sw allow-downgrade	Allow the firmware version on the newly added stack member to be downgraded if the firmware version on manager is older. Config migration is not assured for firmware downgrade.
exit	Exit to Privileged EXEC mode.
show auto-copy-sw	View the Stack Firmware Synchronization settings for the stack.
reload *unit*	If necessary, reload the specified stack member.

NOTE: The OSPF feature uses NSF to enable the hardware to continue forwarding IPv4 packets using OSPF routes while a backup unit takes over stack master responsibility. Additional NSF commands are available in OSPF and OSPFv3 command modes. For more information, see "NSF OSPF Configuration " on page 1207 and "NSF OSPFv3 Configuration " on page 1224

Viewing and Clearing Stacking and NSF Information

Beginning in Privileged EXEC mode, use the following commands to view stacking information and to clear NSF statistics.

Command	Purpose
show switch [*stack-member-number]*	View information about all stack members or the specified member.
show switch stack-standby	View the ID of the switch that will assume the role of the stack master if it goes down.
show switch stack-ports	View information about the stacking ports.
show switch stack-ports counters	View the statistics about the data the stacking ports have transmitted and received.
show switch stack-ports stack-path	View the path that packets take from one stack member to another.
show supported switchtype	View the Dell Networking models that are supported in the stack and the switch index (SID) associated with each model.
show nsf	View summary information about the NSF state of the master and standby switches.
show checkpoint statistics	View information about checkpoint messages generated by the stack master.
clear checkpoint statistics	Reset the checkpoint statistics counters to zero.

Connecting to the Management Console from a Stack Member

From the CLI Unavailable prompt, use the following command to connect the console session to the local unit.

Command	Purpose
connect [*unit*]	Connect the console on the remote unit to the local unit

Stacking and NSF Usage Scenarios

Only a few settings are available to control the stacking configuration, such as the designation of the standby unit or enabling/disabling NSF. The examples in this section describe how the stacking and NSF feature act in various environments.

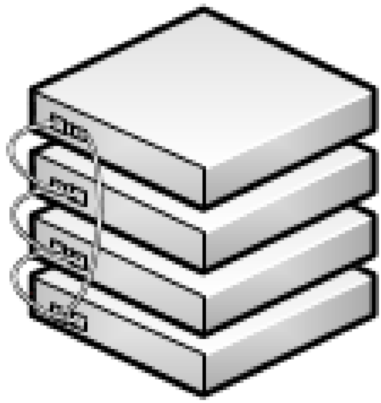
This section contains the following examples:

- Basic Failover
- Preconfiguring a Stack Member
- NSF in the Data Center
- NSF and VoIP
- NSF and DHCP Snooping
- NSF and the Storage Access Network
- NSF and Routed Access

Basic Failover

In this example, the stack has four members that are connected in a ring topology, as Figure 9-11 shows.

Figure 9-11. Basic Stack Failover



When all four units are up and running, the **show switch** CLI command gives the following output:

```
console#show switch

SW   Management      Standby  Preconfig Plugged-  Switch  Code
     Status          Status   Model ID  in Model  Status  Version
                                        ID
---  ---------       -------  --------  --------- ------- --------
1    Stack Member             N3048     N3048     OK      6.0.0.0
2    Stack Member             N3048     N3048     OK      6.0.0.0
3    Mgmt Switch              N3048     N3048     OK      6.0.0.0
4    Stack Member             N3048     N3048     OK      6.0.0.0
```

At this point, if Unit 2 is powered off or rebooted due to an unexpected failure, **show switch** gives the following output:

```
console#show switch

SW  Management     Standby Preconfig Plugged- Switch       Code
    Status         Status  Model ID  in Model Status       Version
                                     ID
--- ---------      ------- --------  ------------------    --------
1   Stack Member           N3048     N3048    OK           6.0.0.0
2   Unassigned             N3048              Not Present  0.0.0.0
3   Mgmt Switch            N3048     N3048    OK           6.0.0.0
```

```
4   Stack Member         N3048     N3048    OK          6.0.0.0
```

When the failed unit resumes normal operation, the previous configuration that exists for that unit is reapplied by the stack master.

To permanently remove the unit from the stack, enter into Stack Config Mode and use the member command, as the following example shows.

```
console#configure
console(config)#stack
console(config-stack)#no member 2
console(config-stack)#exit
console(config)#exit
console#show switch

SW  Management    Standby Preconfig Plugged- Switch      Code
    Status        Status  Model ID  in Model Status      Version
                                    ID
--- ---------     ------- --------  -------------------  --------
1   Stack Member          N3048     N3048    OK          6.0.0.0
3   Mgmt Switch           N3048     N3048    OK          6.0.0.0
4   Stack Member          N3048     N3048    OK          6.0.0.0
```

Preconfiguring a Stack Member

To preconfigure a stack member before connecting the physical unit to the stack, use the **show supported switchtype** command to obtain the SID of the unit to be added.

The example in this section demonstrates pre-configuring a stand-alone Dell Networking N-Series switch.

To configure the switch:

1 View the list of SIDs to determine which SID identifies the switch to preconfigure. The following is the output on Dell Networking N3000 and N2000 Series switches. The supported switch types vary by model.

```
console#show supported switchtype

SID         Switch Model ID
--- --------------------------------
1   N3024
2   N3024F
3   N3024P
4   N3048
5   N3048P
```

```
6   N2024
7   N2024P
8   N2048
9   N2048P
```

The following is the output on Dell Networking N1500 Series switches:

```
console#show supported switchtype


SID Switch Model ID
--- -------------------------------
1   N3024
2   N3024F
3   N3024P
4   N3048
5   N3048P
6   N2024
7   N2024P
8   N2048
9   N2048P
```

NOTE: Dell Networking N1500, N2000, and N3000 Series switches cannot be stacked together.

2 Preconfigure the switch (SID = 2) as member number 2 in the stack.

```
console#configure
console(config)#stack
console(config-stack)#member 2 2
console(config-stack)#exit
console(config)#exit
```

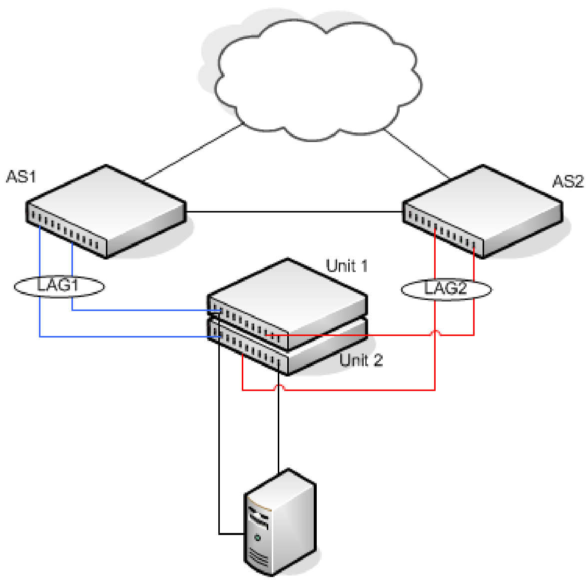
3 Confirm the stack configuration. Some of the fields have been omitted from the following output due to space limitations.

```
console#show switch

Management Standby  Preconfig    Plugged-in   Switch   Code
SW Status  Status   Model ID     Model ID     Status   Version
-- ------  ------   --------     --------     ------   -------
1  Mgmt Sw          N3048        N3048        OK       6.0.0.0
```

NSF in the Data Center

Figure 9-12 illustrates a data center scenario, where the stack of two Dell Networking N-Series switches acts as an access switch. The access switch is connected to two aggregation switches, AS1 and AS2. The stack has a link from two different units to each aggregation switch, with each pair of links grouped together in a LAG. The two LAGs and link between AS1 and AS2 are members of the same VLAN. Spanning tree is enabled on the VLAN. Assume spanning tree selects AS1 as the root bridge. Assume the LAG to AS1 is the root port on the stack and the LAG to AS2 is discarding. Unit 1 is the stack master. If unit 1 fails, the stack removes the Unit 1 link to AS1 from its LAG. The stack forwards outgoing packets through the Unit 2 link to AS1 during the failover. During the failover, the stack continues to send BPDUs and LAG PDUs on its links on Unit 2. The LAGs stay up (with one remaining link in each), and spanning tree on the aggregation switches does not see a topology change.

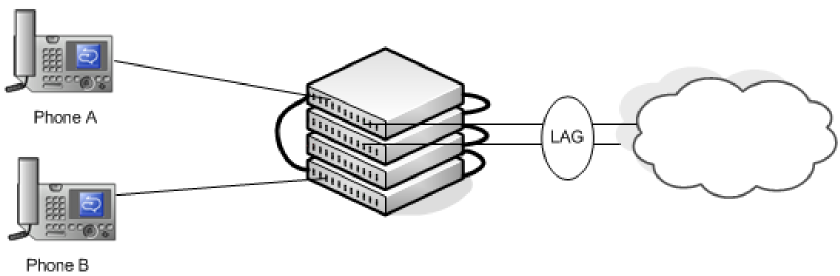
Figure 9-12. Data Center Stack Topology



NSF and VoIP

Figure 9-13 shows how NSF maintains existing voice calls during a stack master failure. Assume the top unit is the stack master. When the stack master fails, the call from phone A is immediately disconnected. The call from phone B continues. On the uplink, the forwarding plane removes the failed LAG member and continues using the remaining LAG member. If phone B has learned VLAN or priority parameters through LLDP-MED, it continues to use those parameters. The stack resumes sending LLDPDUs with MED TLVs once the control plane restarts. Phone B may miss an LLDPDU from the stack, but should not miss enough PDUs to revert its VLAN or priority, assuming the administrator has not reduced the LLDPDU interval or hold count. If phone B is receiving quality of service from policies installed in the hardware, those policies are retained across the stack master restart.

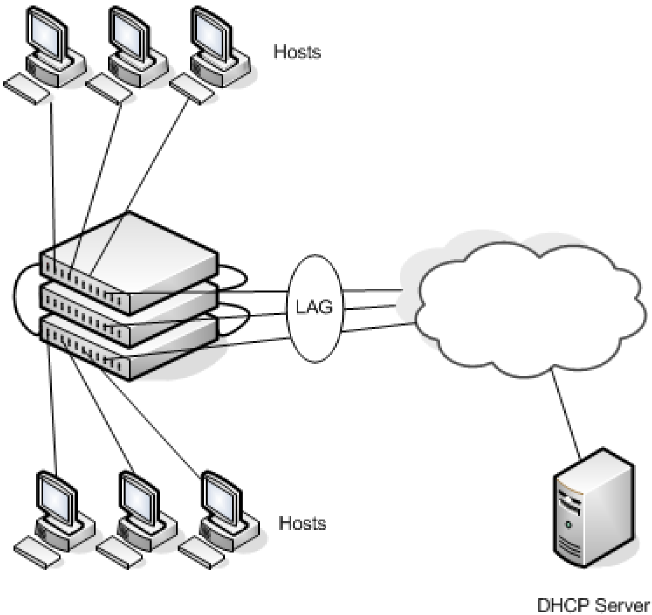
Figure 9-13. NSF and VoIP



NSF and DHCP Snooping

Figure 9-14 illustrates a layer-2 access switch running DHCP snooping. DHCP snooping only accepts DHCP server messages on ports configured as *trusted* ports. DHCP snooping listens to DHCP messages to build a bindings database that lists the IP address the DHCP server has assigned to each host. IP Source Guard (IPSG) uses the bindings database to filter data traffic in hardware based on source IP address and source MAC address. Dynamic ARP Inspection (DAI) uses the bindings database to verify that ARP messages contain a valid sender IP address and sender MAC address. DHCP snooping checkpoints its bindings database.

Figure 9-14. NSF and DHCP Snooping



If the stack master fails, all hosts connected to that unit lose network access until that unit reboots. The hardware on surviving units continues to enforce source filters IPSG installed prior to the failover. Valid hosts continue to communicate normally. During the failover, the hardware continues to drop data packets from unauthorized hosts so that security is not compromised.

If a host is in the middle of an exchange with the DHCP server when the failover occurs, the exchange is interrupted while the control plane restarts. When DHCP snooping is enabled, the hardware traps all DHCP packets to the CPU. The control plane drops these packets during the restart. The DHCP client and server retransmit their DHCP messages until the control plane has resumed operation and messages get through. Thus, DHCP snooping does not miss any new bindings during a failover.

As DHCP snooping applies its checkpointed DHCP bindings, IPSG confirms the existence of the bindings with the hardware by reinstalling its source IP address filters.

If Dynamic ARP Inspection is enabled on the access switch, the hardware traps ARP packets to the CPU on untrusted ports. During a restart, the control plane drops ARP packets. Thus, new traffic sessions may be briefly delayed until after the control plane restarts.

If IPSG is enabled and a DHCP binding is not checkpointed to the backup unit before the failover, that host will not be able to send data packets until it renews its IP address lease with the DHCP server.

NSF and the Storage Access Network

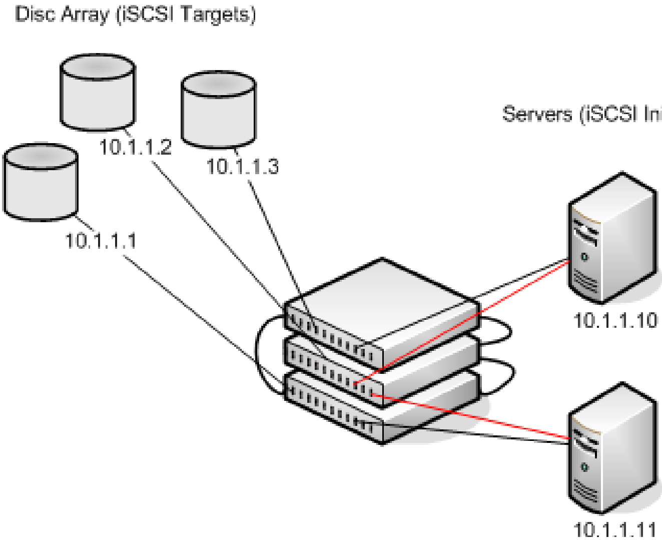
Figure 9-15 illustrates a stack of three Dell Networking N-Series switches connecting two servers (iSCSI initiators) to a disk array (iSCSI targets). There are two iSCSI connections as follows:

Session A: 10.1.1.10 to 10.1.1.3

Session B: 10.1.1.11 to 10.1.1.1

An iSCSI application running on the stack master (the top unit in the diagram) has installed priority filters to ensure that iSCSI traffic that is part of these two sessions receives priority treatment when forwarded in hardware.

Figure 9-15. NSF and a Storage Area Network



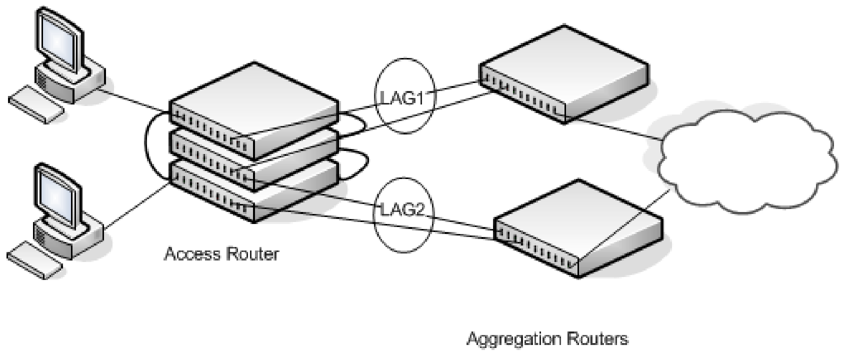
When the stack master fails, session A drops. The initiator at 10.1.1.10 detects a link down on its primary NIC and attempts to reestablish the session on its backup NIC to a different IP address on the disk array. The hardware forwards the packets to establish this new session, but assuming the session is established before the control plane is restarted on the backup unit, the new session receives no priority treatment in the hardware.

Session B remains established and fully functional throughout the restart and continues to receive priority treatment in the hardware.

NSF and Routed Access

Figure 9-16 shows a stack of three units serving as an access router for a set of hosts. Two LAGs connect the stack to two aggregation routers. Each LAG is a member of a VLAN routing interface. The stack has OSPF and PIM adjacencies with each of the aggregation routers. The top unit in the stack is the stack master.

Figure 9-16. NSF and Routed Access



If the stack master fails, its link to the aggregation router is removed from the LAG. When the control plane restarts, both routing interfaces come back up by virtue of the LAGs coming up. OSPF sends grace LSAs to inform its OSPF neighbors (the aggregation routers) that it is going through a graceful restart.

NOTE: The graceful restart feature for OSPF is disabled by default. For information about the web pages and commands to configure NSF for OSPF or OSPFv3, see "OSPF and OSPFv3 " on page 1181.

The grace LSAs reach the neighbors before they drop their adjacencies with the access router. PIM starts sending hello messages to its neighbors on the aggregation routers using a new generation ID to prompt the neighbors to quickly resend multicast routing information. PIM neighbors recognize the new generation ID and immediately relay the group state back to the restarting router. IGMP sends queries to relearn the hosts' interest in multicast groups. IGMP tells PIM the group membership, and PIM sends

JOIN messages upstream. The control plane updates the driver with checkpointed unicast routes. The forwarding plane reconciles layer-3 hardware tables.

The OSPF graceful restart finishes, and the control plane deletes any stale unicast routes not relearned at this point. The forwarding plane reconciles layer-3 multicast hardware tables. Throughout the process, the hosts continue to receive their multicast streams, possibly with a short interruption as the top aggregation router learns that one of its LAG members is down. The hosts see no more than a 50 ms interruption in unicast connectivity.

10

Authentication, Authorization, and Accounting

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to control access to the switch management interface using authentication and authorization. These services can also be used to restrict or allow network access when used in conjunction with IEEE 802.1x. It also describes how to record this access using accounting. Together the three services are referred to by the acronym AAA.

The topics covered in this chapter include:

- AAA Introduction
- Authentication
- Authorization
- Accounting
- IEEE 802.1X
- Captive Portal

AAA Introduction

AAA is a framework for configuring management security in a consistent way. Three services make up AAA:

- Authentication—Validates the user identity. Authentication takes place before the user is allowed access to switch services.
- Authorization—Determines which services the user is allowed to access. Examples of services are access to the switch management console and access to network services.
- Accounting—Collects and sends security information about switch management console users and switch management commands

Each service is configured using method lists. Method lists define how each service is to be performed by specifying the methods available to perform the service. The first method in a list is tried first. If the first method returns an error, the next method in the list is tried. This continues until all methods in the list have been attempted. If no method can perform the service, then the service fails. A method may return an error due to lack of network access, misconfiguration of a server, and other reasons. If there is no error, the method returns success if the user is allowed access to the service and failure if the user is not.

AAA gives the user flexibility in configuration by allowing different method lists to be assigned to different access lines. In this way, it is possible to configure different security requirements for the serial console than for Telnet, for example.

Methods

A method performs authentication or authorization for the configured service. Not every method is available for every service. Some methods require a username and password and other methods only require a password. Table 10-1 summarizes the various methods:

Table 10-1. AAA Methods

Method	Username?	Password?	Can Return an Error?
enable	no	yes	yes
ias	yes	yes	no
line	no	yes	yes
local	yes	yes	yes
none	no	no	no
radius	yes	yes	yes
tacacs	yes	yes	yes

Methods that never return an error cannot be followed by any other methods in a method list.

- The **enable** method uses the enable password. If there is no enable password defined, then the enable method will return an error.
- The **ias** method is a special method that is only used for 802.1X. It uses an internal database (separate from the local user database) that acts like an 802.1X authentication server. This method never returns an error. It will always pass or deny a user.
- The **line** method uses the password for the access line on which the user is accessing the switch. If there is no line password defined for the access line, then the line method will return an error.
- The **local** method uses the local user database. If the user password does not match, then access is denied. This method returns an error if the user name is not present in the local user database.
- The **none** method does not perform any service, but instead always returns a result as if the service had succeeded. This method never returns an error. If none is configured as a method, the user will always be authenticated and allowed to access the switch.
- The **radius** and **tacacs** methods communicate with servers running the RADIUS and TACACS+ protocols, respectively. These methods can return an error if the switch is unable to contact the server.

Method Lists

The method lists shown in Table 10-2 are defined by default. They cannot be deleted, but they can be modified. Using the "no" command on these lists will return them to their default configuration.

Table 10-2. Default Method Lists

AAA Service (type)	List Name	List Methods
Authentication (login)	defaultList	none
Authentication (login)	networkList	local
Authentication (enable)	enableList	enable none
Authentication (enable)	enableNetList	enable
Authorization (exec)	dfltExecAuthList	none

Table 10-2. Default Method Lists (Continued)

AAA Service (type)	List Name	List Methods
Authorization (commands)	dfltCmdAuthList	none
Accounting (exec)	dfltExecList	tacacs (start-stop)
Accounting (commands)	dfltCmdList	tacacs (stop-only)

Access Lines

There are five access lines: console, Telnet, SSH, HTTP, and HTTPS. HTTP and HTTPS are not configured using AAA method lists. Instead, the authentication list for HTTP and HTTPS is configured directly (authorization and accounting are not supported). The default method lists for both the HTTP and HTTPS access lines consist of only the local method. Each of the other access lines may be assigned method lists independently for the AAA services.

The SSH line has built-in authentication beyond that configured by the administrator.

In the SSH protocol itself, there are multiple methods for authentication. These are not the authentication methods configured in AAA, but are internal to SSH itself. When an SSH connection is attempted, the challenge-response method is specified in the connection request.

The methods available for authentication are: host-based authentication, public key authentication, challenge-response authentication, and password authentication. Authentication methods are tried in the order specified above, although SSH-2 has a configuration option to change the default order.

Host-based authentication operates as follows:

If the host from which the user logs in is listed in a specific file (/etc/hosts.equiv or /etc/ssh/shosts.equiv) on the remote host, and the user names are the same on both hosts, or if the files ~/.rhosts or ~/.shosts exist in the user's home directory on the remote host and contain a line containing the name of the client machine and the name of the user on that machine, the user is considered for login. Additionally, the server must be able to verify the client's host key for login to be permitted. This authentication method closes security holes due to IP spoofing, DNS spoofing, and routing spoofing.

This authentication method is not implemented by Dell Networking N-Series switches. Use the Management ACL capability to perform the equivalent function.

Public key authentication operates as follows:

The administrator first generates a pair of encryption keys, the “public” key and the “private” key. Messages encrypted with the private key can be decrypted only by the public key, and vice-versa. The administrator keeps the private key on his/her local machine, and loads the public key on to the switch. When the administrator attempts to log into the switch, the protocol sends a brief message, encrypted with the public key. If the switch can decrypt the message (and can send back some proof that it has done so) then the response proves that switch must possess the private key, and user is authenticated without giving a username/password.

The public key method is implemented in the Dell Networking N-Series switch as opposed to an external server. If the user does not present a certificate, it is not considered an error, and authentication will continue with challenge-response authentication.

Challenge-response authentication works as follows:

The switch sends an arbitrary “challenge” text and prompts for a response. SSH-2 allows multiple challenges and responses; SSH-1 is restricted to one challenge/response only. Examples of challenge-response authentication include BSD Authentication.

Finally, if all other authentication methods fail, SSH prompts the user for a password.

Access Lines (AAA)

Table 10-3 shows the method lists assigned to the various access lines by default.

Table 10-3. Default AAA Methods

AAA Service (type)	Console	Telnet	SSH
Authentication (login)	defaultList	networkList	networkList
Authentication (enable)	enableList	enableNetList	enableNetList

Table 10-3. Default AAA Methods (Continued)

AAA Service (type)	Console	Telnet	SSH
Authorization (exec)	dfltExecAuthList	dfltExecAuthList	dfltExecAuthList
Authorization (commands)	dfltCmdAuthList	dfltCmdAuthList	dfltCmdAuthList
Accounting (exec)	none	none	none
Accounting (commands)	none	none	none

Access Lines (Non-AAA)

Table 10-4 shows the default configuration of the access lines that do not use method lists.

Table 10-4. Default Configuration for Non-AAA Access Lines

Access Line	Authentication	Authorization
HTTP	local	n/a
HTTPS	local	n/a
802.1X	none	none

Authentication

Authentication is the process of validating a user's identity. During the authentication process, only identity validation is done. There is no determination made of which switch services the user is allowed to access. This is true even when RADIUS is used for authentication; RADIUS cannot perform separate transactions for authentication and authorization. However, the RADIUS server can provide attributes during the authentication process that are used in the authorization process.

Authentication Types

There are three types of authentication:

- **Login—** Login authentication grants access to the switch if the user credentials are validated. Access is granted only at privilege level one.
- **Enable—**Enable authentication grants access to a higher privilege level if the user credentials are validated for the higher privilege level. When RADIUS is used for enable authentication, the username for this request is always $enab15$. The username used to log into the switch is not used for RADIUS enable authentication.
- **802.1X—**802.1X authentication is used to grant an 802.1X supplicant access to the network. For more information about 802.1X, see "Port and System Security " on page 623.

Table 10-5 shows the valid methods for each type of authentication:

Table 10-5. Valid Methods for Authentication Types

Method	Login	Enable	Dot1x
enable	yes	yes	no
ias	no	no	yes
line	yes	yes	no
local	yes	no	no
none	yes	yes	yes
radius	yes	yes	yes
tacacs	yes	yes	no

Authentication Manager

Overview

The Authentication Manager supports the hierarchical configuration of host authentication methods on an interface. Use of the Authentication Manager is optional, but it is recommended when using multiple types of authentication on an interface, e.g., Captive Portal in conjunction with MAB or IEEE 802.1X. Dell switches support the following host authentication methods:

- IEEE 802.1x
- MAC Authentication Bypass (MAB)
- Captive portal

Using the Authentication Manager, the administrator can configure an authentication method list on a per-port basis. Authentication can be enabled or disabled. If authentication is disabled, then no authentication method is applied and the port is provided with open access. The default behavior is that authentication is disabled for all ports.

The configured authentication methods are attempted in the configured order. If an authentication method times out (an error), then the next configured method is attempted. If an authentication method fails, i.e., an incorrect password was entered, then the next method is not attempted and authentication begins again from the first method. If all the methods return an error, then the Authentication Manager starts a timer for reauthentication. The value of the timer is equal to the re-authentication restart timer. Failure in this context means that host authentication was attempted and the host was unable to successfully authenticate. At the expiry of the timer, the Authentication Manager starts the authentication process again from the first method in the list.

The Authentication Manager supports configuring a priority for each authentication method on a port. The authentication priority allows a higher priority method (not currently running) to interrupt an authentication in progress with a lower-priority method. If a client is already authenticated, an interrupt from a higher-priority method can cause a client previously authenticated using a lower priority method to reauthenticate.

By default, Dell switches are configured with a method list that contains the methods (in order) Dot1x, MAB, and Captive Portal (web-auth) as the default methods for all the ports. Dell switches restrict the configuration such that no method is allowed to follow the Captive Portal method, if configured.

The authentication manager controls only the order in which the authentication methods are executed. The switch administrator is responsible for implementing the required configuration for the respective methods to authenticate successfully.

Authentication Restart

Authentication restarts from the first configured method on any of the following events:

- Link flap
- Authentication fails for all configured methods
- Authentication priority (802.1X packet received when a lower priority method is active)

802.1X Interaction

By default, 802.1X drops all traffic prior to successful 802.1X (or MAB) authentication. If Captive Portal is configured as a method, authentication allows certain traffic types, such as DHCP or DNS, access to the network during the Captive Portal method invocation.

Authentication Priority

The default authentication priority of a method is equivalent to its position in the order of the authentication list. If authentication method priorities are not configured, then the relative priorities (first is highest) are in the same order as that of the per-port based authentication list.

Authentication priority allows a higher-priority method (not currently running) to interrupt an authentication in progress with a lower-priority method. Alternatively, if the client is already authenticated, an interrupt from a higher-priority method can cause a client, which was previously authenticated using a lower-priority method, to reauthenticate.

For example, if a client is already authenticated using a method other than 802.1X (MAB or Captive Portal) and 802.1X has higher priority than the authenticated method, and if an 802.1X frame is received, then the existing authenticated client is removed and the authentication process begins again from the first method in the order. If 802.1X has a lower priority than the authenticated method, then the client is not removed and the 802.1X frames are ignored.

If administrator changes the priority of the methods, then all the users who are authenticated using a lower-priority method are forced to reauthenticate. If an authentication session is in progress and the administrator changes the order of the authentication methods, then the configuration will take effect for the next session onwards.

Configuration Example—802.1X and MAB

In this scenario, the authentication manager selects the first authentication method, 802.1X. If authentication using 802.1X is successful, then the client is allowed network access. If authentication using 802.1X errors out, then authentication manager selects the next authentication method: MAB. If authentication using MAB returns an error, then the port is unauthorized. The authentication manager will start a timer to re-authenticate the client. At the expiry of the timer, the authentication manager restarts authentication by selecting the 802.1X method.

1 Enter global configuration mode and define the RADIUS server.

```
console#configure
console(config)#radius-server host 10.10.10.10
console(Config-radius)#name BigRadius
console(Config-radius)#primary
console(Config-radius)#usage 802.1x
console(Config-radius)#exit
```

2 Define the RADIUS server key.

```
console(config)#radius-server key thatsyoursecret-keepit-
keepit
```

3 Enable authentication and globally enable 802.1x client authentication via RADIUS:

```
console(config)#authentication enable
console(config)#aaa authentication dot1x default radius
console(config)#dot1x system-auth-control
```

4 On the interface, enable MAC based authentication mode, enable MAB, and set the order of authentication to 802.1X followed by MAC authentication. Also enable periodic re-authentication.

```
console(config)#interface te1/0/4
console(config-if-Te1/0/4)#dot1x port-control mac-based
console(config-if-Te1/0/4)#dot1x mac-auth-bypass
console(config-if-Te1/0/4)#authentication order dot1x mab
console(config-if-Te1/0/4)#dot1x reauthentication
console(config-if-Te1/0/4)#exit
```

Configuration Example—MAB Client

This example shows how to configure a MAB client on interface Gi1/0/2 using the IAS database for authentication.

1 Enter global configuration mode and create VLAN 3.

```
console#configure
console(config)#configure
console(config)#vlan 3
console(config-vlan3)#exit
```

2 Enable the authentication manager and globally enable 802.1x.

```
console(config)#authentication enable
console(config)#dot1x system-auth-control
```

3 Set IEEE 802.1x to use the local IAS user database.

```
console(config)#aaa authentication dot1x default ias
```

4 Configure the IAS database with the client MAC address as the user name and password. The password MUST be entered in upper case or the authentication will fail with an MD5 Validation Failure, as the MD5 password hashes would not match.

```
console(config)#aaa ias-user username F8B1562BA1D9
console(config-ias-user)#password F8B1562BA1D9
console(config-ias-user)#exit
```

5 Configure interface gi1/0/2 to use VLAN 3 in access mode.

```
console(config)#interface Gi1/0/2
console(config-ifGi1/0/2)#switchport access vlan 3
```

6 On the interface, configure the port to use MAC based authentication and enable MAB. The authentication manager is configured to only use MAB and the priority is set to MAB.

```
console(config-ifGi1/0/2)#dot1x port-control mac-based
console(config-ifGi1/0/2)#dot1x mac-auth-bypass
console(config-ifGi1/0/2)#authentication order mab
console(config-ifGi1/0/2)#authentication priority mab
console(config-ifGi1/0/2)#exit
```

If it is possible that an 802.1x aware client may be connected, it is advisable to configure a re-authentication timer on the port using the **dot1x timeout re-authperiod** command.

The following command shows the 802.1x configuration on the interface:

```
console(config-if-Gi1/0/1)#show dot1x interface gi1/0/2

Administrative Mode............... Enabled
Dynamic VLAN Creation Mode........ Disabled
VLAN Assignment Mode.............. Disabled
Monitor Mode...................... Disabled

Port      Admin Mode         Oper Mode    Reauth   Reauth
                                          Control  Period
--------- ------------------ ------------ -------- ----------
Gi1/0/2   mac-based          Authorized   FALSE    3600

Quiet Period................................... 60
Transmit Period................................ 30
Maximum Requests............................... 2
Max Users...................................... 64
Guest-vlan Timeout............................. 90
Server Timeout (secs).......................... 30
MAB mode (configured).......................... Enabled
MAB mode (operational)......................... Enabled

Logical Supplicant    AuthPAE          Backend    VLAN Username
Filter
Port    MAC-Address    State            State      Id
Id
------- -------------- ------------------ ---------- ---- ---------
---- ------
64      F8B1.562B.A1D9 Authenticated      Idle       3
F8B1562BA1D9

console(config-if-Gi1/0/1)#show dot1x clients all

Clients Authenticated using Monitor Mode....... 0
Clients Authenticated using Dot1x.............. 1
Interface...................................... Gi1/0/2
User Name...................................... F8B1562BA1D9
Supp MAC Address............................... F8B1.562B.A1D9
Session Time................................... 1240
Filter Id......................................
VLAN Assigned.................................. 3
```

Using RADIUS

The RADIUS client on the switch supports multiple RADIUS servers. When multiple authentication servers are configured, they can help provide redundancy. One server can be designated as the primary and the other(s) will function as backup server(s). The switch attempts to use the primary server first. if the primary server does not respond, the switch attempts to use the backup servers. A priority value can be configured to determine the order in which the backup servers are contacted.

How Does RADIUS Control Management Access?

Many networks use a RADIUS server to maintain a centralized user database that contains per-user authentication information. RADIUS servers provide a centralized authentication method for:

- Network Access (IEEE 802.1X)
- User Manager (Management access)
- Captive Portal

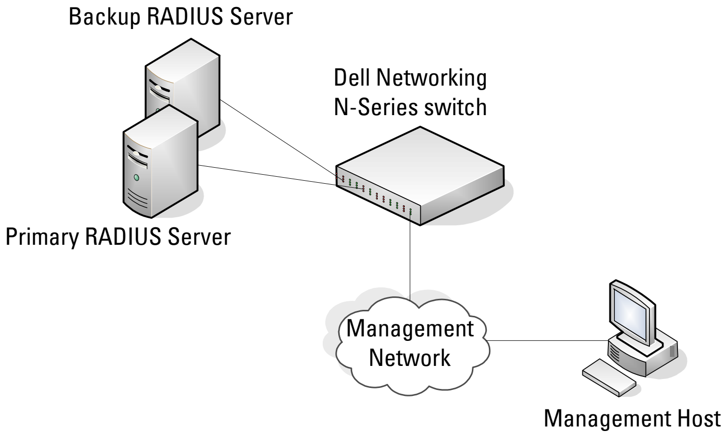
Like TACACS+, RADIUS access control utilizes a database of user information on a remote server. Making use of a single database of accessible information—as in an Authentication Server—can greatly simplify the authentication and management of users in a large network. One such type of Authentication Server supports the Remote Authentication Dial In User Service (RADIUS) protocol as defined by RFC 2865.

For authenticating users, the RADIUS standard has become the protocol of choice by administrators of large networks. To accomplish the authentication in a secure manner, the RADIUS client and RADIUS server must both be configured with the same shared password or “secret”. This “secret” is used to generate one-way encrypted authenticators that are present in all RADIUS packets. The “secret” is never transmitted over the network.

RADIUS conforms to a secure communications client/server model using UDP as a transport protocol. It is extremely flexible, supporting a variety of methods to authenticate and statistically track users. RADIUS is also extensible, allowing for new methods of authentication to be added without disrupting existing functionality.

As a user attempts to connect to the switch management interface, the switch first detects the contact and prompts the user for a name and password. The switch encrypts the supplied information, and a RADIUS client transports the request to a pre-configured RADIUS server.

Figure 10-1. RADIUS Topology



The server can authenticate the user itself or make use of a back-end device to ascertain authenticity. In either case a response may or may not be forthcoming to the client. If the server accepts the user, it returns a positive result with attributes containing configuration information. If the server rejects the user, it returns a negative result. If the server rejects the client or the shared secrets differ, the server returns no result. If the server requires additional verification from the user, it returns a challenge, and the request process begins again.

If using a RADIUS server to authenticate users, the RADIUS administrator must configure user attributes in the user database on the RADIUS server. The user attributes include the user name, password, and privilege level.

NOTE: To set the user privilege level at login, it is required that the Service-Type attribute be used for RADIUS instead of the Cisco AV pair priv-lvl attribute. The Cisco AV priv-lvl is supported only for TACACS authorization.

Which RADIUS Attributes Does the Switch Support?

Table 10-6 lists the RADIUS attributes that the switch supports and indicates whether the 802.1X feature, user management feature, or Captive Portal feature supports the attribute. The RADIUS administrator must configure these attributes on the RADIUS server(s) when utilizing the switch RADIUS service.

Table 10-6. Supported RADIUS Attributes

Type	RADIUS Attribute Name	802.1X	User Manager	Captive Portal
1	User-Name	Yes	Yes	No
2	User-Password	Yes	Yes	No
4	Nas-IP-Address	Yes	Yes	No
5	Nas-Port	Yes	No	No
6	Service-Type	Yes	Yes	No
8	Framed IP Address	Yes	No	No
11	Filter-Id	Yes	No	No
12	Framed-MTU	Yes	No	No
15	Login-Service	No	Yes	No
18	Reply-Message	Yes	Yes	No
24	State	Yes	Yes	No
25	Class	Yes	Yes	No
26	Vendor-Specific	Yes	Yes	Yes
27	Session-Timeout	Yes	No	Yes
28	Idle-Timeout	No	No	Yes
29	Termination-Action	Yes	No	No
30	Called-Station-Id	Yes	No	No
31	Calling-Station-Id	Yes	No	Yes
32	Nas-Identifier	Yes	Yes	No

Table 10-6. Supported RADIUS Attributes (Continued)

Type	RADIUS Attribute Name	802.1X	User Manager	Captive Portal
40	Acct-Status-Type	Set by RADIUS client for Accounting	Yes	No
42	Acct-Input-Octets	Yes	No	No
43	Acct-Output-Octets	Yes	No	No
44	Acct-Session-Id	Set by RADIUS client for Accounting	Yes	No
46	Acct-Session-Time	Yes	Yes	No
49	Acct-Terminate-Cause	Yes	No	No
52	Acct-Input-Gigawords	Yes	No	No
53	Acct-Output-Gigawords	Yes	No	No
61	Nas-Port-Type	Yes	No	Yes
64	Tunnel-Type	Yes	No	No
65	Tunnel-Medium-Type	Yes	No	No
79	EAP-Message	Yes	No	No
80	Message-Authenticator	Set by RADIUS client for Accounting	Yes	No
81	Tunnel-Privategroup-Id	Yes	No	No

How Are RADIUS Attributes Processed on the Switch?

The following attributes are processed in the RADIUS Access-Accept message received from a RADIUS server:

- NAS-PORT

  ifIndex of the port to be authenticated.

- REPLY-MESSAGE

  Trigger to respond to the Access-Accept message with an EAP notification.

- STATE

  RADIUS server state. Transmitted in Access-Request and Accounting-Request messages.

- SESSION-TIMEOUT

  Session time-out value for the session (in seconds). Used by both 802.1x and Captive Portal.

- TERMINATION-ACTION

  Indication as to the action taken when the service is completed.

- EAP-MESSAGE

  Contains an EAP message to be sent to the user. This is typically used for MAB clients.

- VENDOR-SPECIFIC

  The following Cisco AV Pairs are supported:

  - shell:priv-lvl
  - shell:roles

- FILTER-ID

  Name of the diffserv policy for this user.

- TUNNEL-TYPE

  Used to indicate that a VLAN is to be assigned to the user when set to tunnel type VLAN (13).

- TUNNEL-MEDIUM-TYPE

  Used to indicate the tunnel medium type. Must be set to medium type 802 (6) to enable VLAN assignment.

- TUNNEL-PRIVATE-GROUP-ID

  Used to indicate the VLAN to be assigned to the user. May be a string which matches a pre-configured VLAN name or a VLAN ID. If a VLAN ID is given, the string must contain only decimal digits.

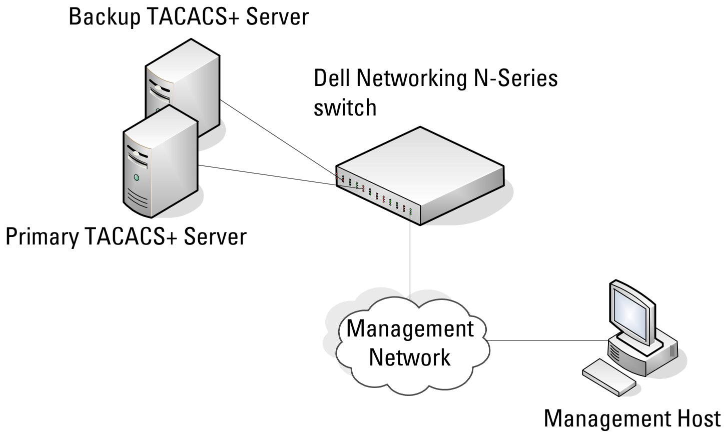
Using TACACS+ Servers to Control Management Access

TACACS+ (Terminal Access Controller Access Control System) provides access control for networked devices via one or more centralized servers. TACACS+ simplifies authentication by making use of a single database that can be shared by many clients on a large network. TACACS+ uses TCP to ensure reliable delivery and a shared key configured on the client and daemon server to encrypt all messages.

If TACACS+ is configured as the authentication method for user login and a user attempts to access the user interface on the switch, the switch prompts for the user login credentials and requests services from the TACACS+ client. The client then uses the configured list of servers for authentication, and provides results back to the switch.

Figure 10-2 shows an example of access management using TACACS+.

Figure 10-2. Basic TACACS+ Topology



The TACACS+ server list can be configured with one or more hosts defined via their network IP addresses. Each can be assigned a priority to determine the order in which the TACACS+ client will contact the servers. TACACS+ contacts the server when a connection attempt fails or times out for a higher priority server.

Each server host can be configured with a specific connection type, port, timeout, and shared key, or the server hosts can be globally configured with the key and timeout.

The TACACS+ server can do the authentication itself, or redirect the request to another back-end device. All sensitive information is encrypted and the shared secret is never passed over the network; it is used only to encrypt the data.

Which TACACS+ Attributes Does the Switch Support?

Table 10-7 lists the TACACS+ attributes that the switch supports and indicates whether the authorization or accounting service supports sending or receiving the attribute. The authentication service does not use attributes. The following attributes can be configured on the TACACS+ server(s) when utilizing the switch TACACS+ service.

Table 10-7. Supported TACACS+ Attributes

Attribute Name	Exec Authorization	Command Authorization	Accounting
cmd	both (optional)	sent	sent
cmd-arg		sent	
elapsed-time			sent
priv-lvl	received		
protocol			sent
roles	both (optional)		
service=shell	both	sent	sent
start-time			sent
stop-time			sent

Authentication Examples

It is important to understand that during authentication, all that happens is that the user is validated. If any attributes are returned from the server, they are not processed during authentication. In the examples below, it is assumed that the default configuration of authorization—that is, no authorization—is used.

Local Authentication Example

Use the following configuration to require local authentication when logging in over a Telnet connection:

1 Create a login authentication list called “loc” that contains the method local:

```
console#config
console(config)#aaa authentication login "loc" local
```

2 Enter the configuration mode for the Telnet line:

```
console(config)#line telnet
```

3 Assign the loc login authentication list to be used for users accessing the switch via Telnet:

```
console(config-telnet)#login authentication loc
console(config-telnet)#exit
```

4 Allow Telnet and SSH users access to Privileged Exec mode. It is required that an enable password be configured to allow local access users to elevate to privileged exec level:

```
console(config)#enable password PaSSW0rd
```

5 Create a user with the name “guest” and password “password”. A simple password can be configured here, since strength-checking has not yet been enabled:

```
console(config)#username guest password password
```

6 Set the minimum number of numeric characters required when password strength checking is enabled. This parameter is enabled only if the **passwords strength minimum character-classes** parameter is set to something greater than its default value of 0:

```
console(config)#passwords strength minimum numeric-characters
2
```

7 Set the minimum number of character classes that must be present in the password. The possible character classes are: upper-case, lower-case, numeric and special:

```
console(config)#passwords strength minimum character-classes 4
```

8 Enable password strength checking:

```
console(config)#passwords strength-check
```

9 Create a user with the name “admin” and password “paSS1&word2”. This user is enabled for privilege level 15. Note that, because password strength checking was enabled, the password was required to have at least two numeric characters, one uppercase character, one lowercase character, and one special character:

```
console(config)#username admin password paSS1&word2 privilege
15
```

10 Configure the switch to lock out a local user after three failed login attempts:

```
console(config)#passwords lock-out 3
```

This configuration allows either user to log into the switch. Both users will have privilege level 1. If no enable password was configured, neither user would be able to successfully execute the **enable** command, which grants access to Privileged EXEC mode, because there is no enable password set by default (the default method list for Telnet enable authentication is only the “enable” method).

NOTE: It is recommend that the password strength checking and password lockout features be enabled when configuring local users.

RADIUS Authentication Example

Use the following configuration to require RADIUS authentication to login over a Telnet connection:

1 Create a login authentication list called “rad” that contains the method radius. If this method returns an error, the user will fail to login:

```
console#config
console(config)#aaa authentication login “rad” radius
```

2 Create an enable authentication list called “raden” that contains the method radius. If this method fails, then the user will fail to execute the enable command:

```
console(config)#aaa authentication enable “raden” radius
```

3 The following command is the first step in defining a RADIUS server at IP address 1.2.3.4. The result of this command is to place the user in radius-server mode to allow further configuration of the server:

```
console(config)#radius-server host 1.2.3.4
```

4 Define the shared secret. This must be the same as the shared secret defined on the RADIUS server:

```
console(config-radius)#key “secret”
console(config-radius)#exit
```

5 Enter the configuration mode for the Telnet line:

```
console(config)#line telnet
```

6 Assign the rad login authentication method list to be used for users accessing the switch via Telnet:

```
console(config-telnet)#login authentication rad
```

7 Assign the raden enable authentication method list to be used for users executing the enable command when accessing the switch via Telnet:

```
console(config-telnet)#enable authentication raden
console(config)#exit
```

ACL Using Authentication Manager to Configure MAB with RADIUS Server

The following is a relatively complex example of using an ACL to control access to Gi1/0/1, using the Authentication Manager to configure MAB in conjunction with a RADIUS server.

1 Create VLAN 60 which will be used for management access via Gi1/0/1:

```
console#config
console(config)#vlan 60
console(config-vlan60)#exit
```

2 Enable 802.1x authentication:

```
console(config)#authentication enable
```

3 Create an access list limiting IP communication exclusively to host 172.25.129.299. All other IP addresses are excluded. This address is in the Bogons address space:

```
console(config)#ip access-list RADIUSCAP
console(config-ip-acl)#permit ip any 172.25.129.229 0.0.0.0
console(config-ip-acl)#permit ip 172.25.129.229 0.0.0.0 any
console(config-ip-acl)#deny ip any any
console(config-ip-acl)#permit every
console(config-ip-acl)#exit
```

4 Set a default gateway for the switch:

```
console(config)#ip default-gateway 172.25.128.254
```

5 Set a default route with administrative distance 253:

```
console(config)#ip route 0.0.0.0 0.0.0.0 172.25.128.254 253
```

6 Assign an IP address to the management VLAN:

```
console(config)#interface vlan 60
console(config-vlan60)#ip address 172.25.128.214 255.255.0.0
console(config-vlan60)#exit
```

7 Enable 802.1x client authentication via RADIUS and allow VLAN assignment to 802.1x clients:

```
console(config)#dot1x system-auth-control
console(config)#aaa authentication dot1x default radius
console(config)#aaa authorization network default radius
```

8 Allow 802.1x client VLANs to be dynamically created via RADIUS:

```
console(config)#dot1x dynamic-vlan enable
```

9 Configure the primary RADIUS sever:

```
console(config)#radius-server host auth 172.25.129.229
```

```
console(config-auth-radius)#primary
console(config-auth-radius)#name "Default-RADIUS-Server"
console(config-auth-radius)#usage 802.1x
console(config-auth-radius)#key "dellSecret"
console(config)#exit
```

10 Configure the management interface and bypass 802.1x authentication for the connected management host:

```
console(config)#interface Gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 60
console(config-if-Gi1/0/1)#dot1x port-control force-authorized
console(config-if-Gi1/0/1)#ip access-group RADIUSCAP in 1
console(config)#exit
```

11 Configure a dedicated printer port. This ports is enabled for MAB only. The VLAN is assigned by the RADIUS server:

```
console(config)#interface Gi1/0/21
console(config-if-Gi1/0/21)#switchport mode general
console(config-if-Gi1/0/21)#dot1x port-control mac-based
console(config-if-Gi1/0/21)#dot1x mac-auth-bypass
console(config-if-Gi1/0/21)#authentication order mab
console(config-if-Gi1/0/21)#authentication priority mab
console(config-if-Gi1/0/21)#exit
```

12 Configure a port for 802.1x access using MAB. This port will periodically re-authenticate connected clients using the configured timer values. The selected timer values are intended to reduce the time required to authenticate:

```
console(config)#interface Gi1/0/22
console(config-if-Gi1/0/22)#switchport mode general
console(config-if-Gi1/0/22)#dot1x port-control mac-based
console(config-if-Gi1/0/22)#dot1x reauthentication
console(config-if-Gi1/0/22)#dot1x timeout quiet-period 10
console(config-if-Gi1/0/22)#dot1x timeout re-authperiod 300
console(config-if-Gi1/0/22)#dot1x timeout tx-period 7
console(config-if-Gi1/0/22)#dot1x timeout guest-vlan-period 5
console(config-if-Gi1/0/22)#dot1x timeout server-timeout 6
console(config-if-Gi1/0/22)#dot1x mac-auth-bypass
console(config-if-Gi1/0/22)#exit
```

Combined RADIUS, CoA, MAB and 802.1x Example

The following example configures RADIUS in conjunction with IEEE 802.1X to provide network access to switch clients.

1 Enable 802.1x:

```
console#config
console(config)#dot1x system-auth-control
```

2 Configure 802.1x clients to use RADIUS services:

```
console(config)#aaa authentication dot1x default radius
```

3 Enable CoA for RADIUS:

```
console(config)#aaa server radius dynamic-author
```

4 Configure the remote RADIUS server for COA requests at 10.130.191.89 with "shared secret" as the key:

```
console(config-radius-da)#client 10.130.191.89 server-key
"shared secret"
```

5 Specify that any CoA with a matching key identifies a client:

```
console(config-radius-da)#auth-type any
console(config-radius-da)#exit
```

6 Configure a group of RADIUS clients (switches) to act as a single large RADIUS client:

```
console(config)#radius-server attribute 4 10.130.65.4
```

7 Specify that the RADIUS server for host authentication/network access is located at 10.130.191.89:

```
console(config)#radius-server host auth 10.130.191.89
```

8 Name the RADIUS server:

```
console(config-auth-radius)#name "Default-RADIUS-Server"
```

9 Configure the RADIUS shared secret as "shared secret":

```
console(config-auth-radius)#key "shared secret"
console(config-auth-radius)#exit
```

10 Configure Gi1/0/7 to use MAC based authentication. This allows multiple hosts sharing the same network port to be individually allowed or denied access to network resources. CoA requests to terminate a host session can be issued by the RADIUS server. This means that if the RADIUS server terminates the host session and subsequently refuses to authorize the host (based upon the MAC address), the host is denied access to the network:

```
console(config)#interface Gi1/0/7
console(config-if-Gi1/0/7)#dot1x port-control mac-based
console(config-if-Gi1/0/7)#exit
```

11 Configure Gi1/0/6 to allow connected hosts access to network resources, regardless of RADIUS configuration. RADIUS CoA disconnect requests are ignored for clients on this port:

```
console(config)#interface Gi1/0/6
console(config-if-Gi1/0/6)#dot1x port-control force-authorized
console(config-if-Gi1/0/6)#exit
```

12 Configure Gi1/0/5 to use standard 802.1x authentication:

```
console(config)#interface Gi1/0/5
console(config-if-Gi1/0/5)#dot1x port-control auto
console(config-if-Gi1/0/5)#exit
```

TACACS+ Authentication Example

Use the following configuration to require TACACS+ authentication when logging in over a Telnet connection:

1 Create a login authentication list called "tacplus" that contains the method tacacs. If this method returns an error, the user will fail to login:

```
console#config
console(config)#aaa authentication login "tacplus" tacacs
```

2 Create an enable authentication list called "tacp" that contains the method tacacs. If this method fails, then the user will fail to execute the enable command:

```
console(config)#aaa authentication enable "tacp" tacacs
```

3 The following command is the first step in defining a TACACS+ server at IP address 1.2.3.4. The result of this command is to place the user in tacacs-server mode to allow further configuration of the server:

```
console(config)#tacacs-server host 1.2.3.4
```

4 Define the shared secret. This must be the same as the shared secret defined on the TACACS+ server:

```
console(tacacs)#key "secret"
console(tacacs)#exit
```

5 Enter the configuration mode for the Telnet line.

```
console(config)#line telnet
```

6 Assign the tacplus login authentication method list to be used for users accessing the switch via Telnet:

```
console(config-telnet)#login authentication tacplus
```

7 Assign the tacp enable authentication method list to be used for users executing the enable command when accessing the switch via Telnet:

```
console(config-telnet)#enable authentication tacp
console(config-telnet)#exit
```

NOTE: A user logging in with this configuration would be placed in User EXEC mode with privilege level 1. To access Privileged EXEC mode with privilege level 15, use the enable command.

NOTE: Dell Networking TACACS supports setting the maximum user privilege level in the authorization response. Configure the TACACS server to send priv-lvl=X, where X is either 1 (Non-privileged mode), or 15 (Privileged mode).

Public Key SSH Authentication Example

The following is an example of a public key configuration for SSH login. Using a tool such as putty and a private/public key infrastructure, one can enable secure login to the Dell Networking N-Series switch without a password. Instead, a public key is used with a private key kept locally on the administrator's computer. The public key can be placed on multiple devices, allowing the administrator secure access without needing to remember multiple passwords. It is strongly recommended that the private key be protected with a password.

This configuration requires entering a public key, which can be generated by a tool such as PuTTYgen. Be sure to generate the correct type of key. In this case, we use an RSA key with the SSH-2 version of the protocol.

Switch Configuration

1 Create a switch administrator:

```
console#config
console(config)#username "admin" password
f4d77eb781360c5711ecf3700a7af623 privilege 15 encrypted
```

2 Set the login and enable methods for line to NOAUTH.

```
console(config)#aaa authentication login "NOAUTH" line
console(config)#aaa authentication enable "NOAUTH" line
```

3 Generate an internal RSA key. This step is not required if an internal RSA key has been generated before on this switch:

```
console(config)#crypto key generate rsa
```

4 Set SSH to use a public key for the specified administrator login. The user login is specified by the **username** command, not the **ias-user** command:

```
console(config)#crypto key pubkey-chain ssh user-key "admin"
rsa
```

5 Enter the public key obtained from a key authority or from a tool such as PuTTyGen. This command is entered as a single line, not as multiple lines as it appears in the following text.

```
console(config)#key-string row
AAAAB3NzaC1yc2EAAAABJQAAAIBor6DPjYDpSy8Qcji68xrS/4Lf8c9Jq4xXKI
Z5Pvv20AkRFE0ifVI9EH4jyZagR3wzH5Xl9dyjA6bTuqMgN15C1xJC1l59FU88
JaY7ywGdRppmoaJrNRPM7RZtQPaDVIunzm3eMr9PywwQ0umsHWGNexUrDYHFWR
IAmJp689AAxw==
console(config)#exit
```

6 Set the line method to SSH:

```
console(config)#line ssh
```

7 Configure the authentication method to the defaultList. The defaultList contains a single method — none — which is equivalent to no authentication. Since the authentication is provided by the public key, a second layer of authentication is not required:

```
console(config-ssh)#login authentication defaultList
console(config-ssh)#exit
```

8 The following three lines enable the SSH server, configure it to use public key authentication, and specify use of the SSH-2 protocol.

```
console(config)#ip ssh server
console(config)#ip ssh pubkey-auth
console(config)#ip ssh protocol 2
```

The following command shows the configured authentication methods:

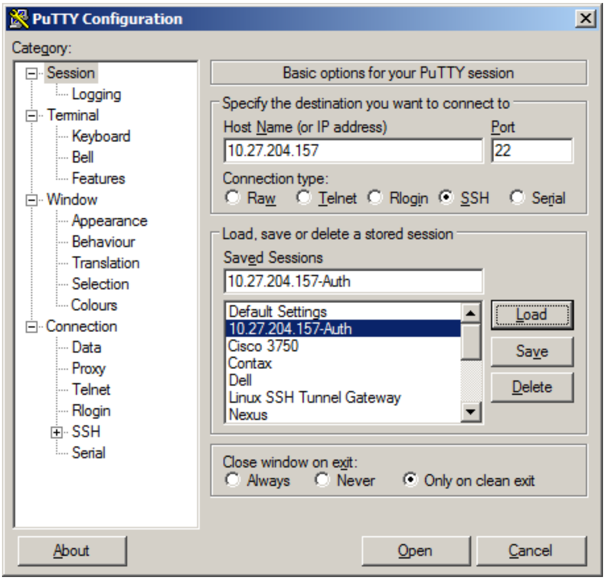
```
console (config)#show authentication methods

Login Authentication Method Lists
---------------------------------
defaultList         :  none
networkList         :  local
NOAUTH              :  line

Enable Authentication Method Lists
----------------------------------
enableList          :  enable   none
enableNetList       :  enable
NOAUTH              :  line

Line     Login Method List    Enable Method List
-------  -----------------    ------------------
Console  defaultList          enableList
Telnet   networkList          enableList
```

```
SSH       defaultList          enableList

HTTPS        :local
HTTP         :local
DOT1X        :
```

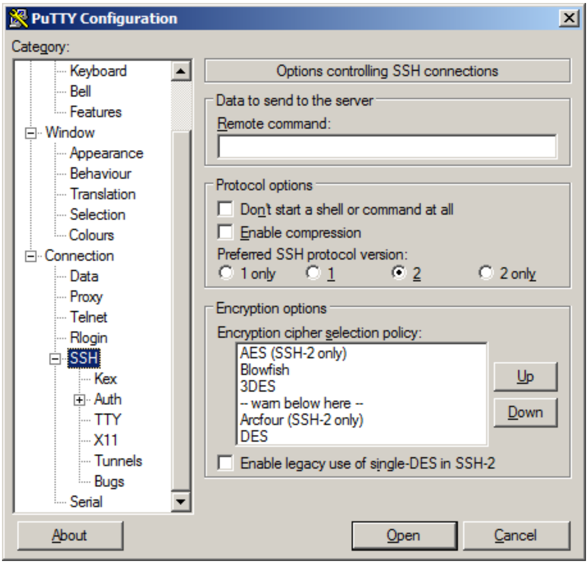
PUTTY Configuration

Main Screen

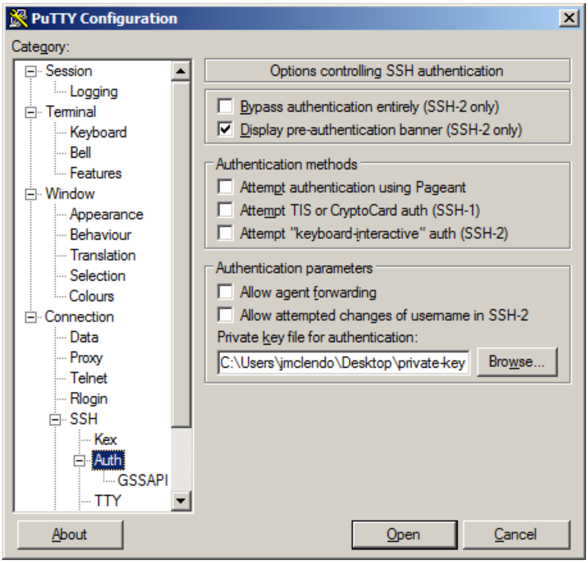
On the following screen, the IP address of the switch is configured and SSH is selected as the secure login protocol.



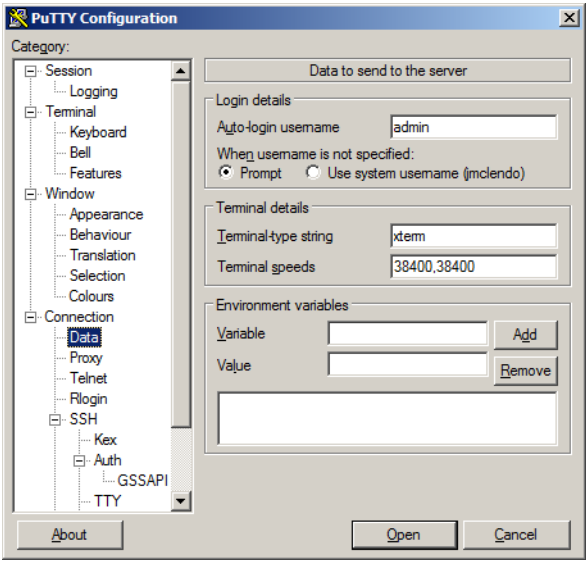
On the next screen, PUTTY is configured to use SSH-2 only. This is an optional step that accelerates the login process.



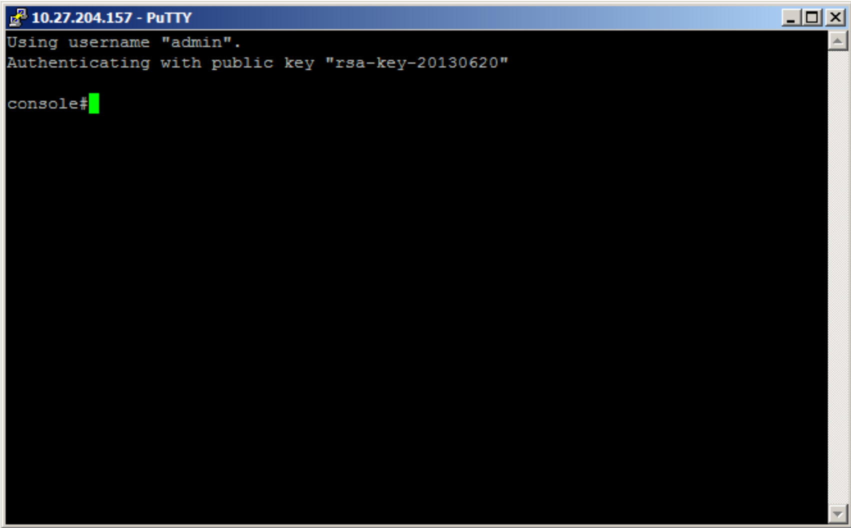
The following screen is the key to the configuration. It is set to display the authentication banner, disable authentication with Pageant, disable keyboard-interactive authentication (unless desired), disable attempted changes of user name, and select the private key file used to authenticate with the switch.



The following screen configures the user name to be sent to the switch. A user name is always required. Alternatively, leave Auto-login name blank and the system will prompt for a user name.

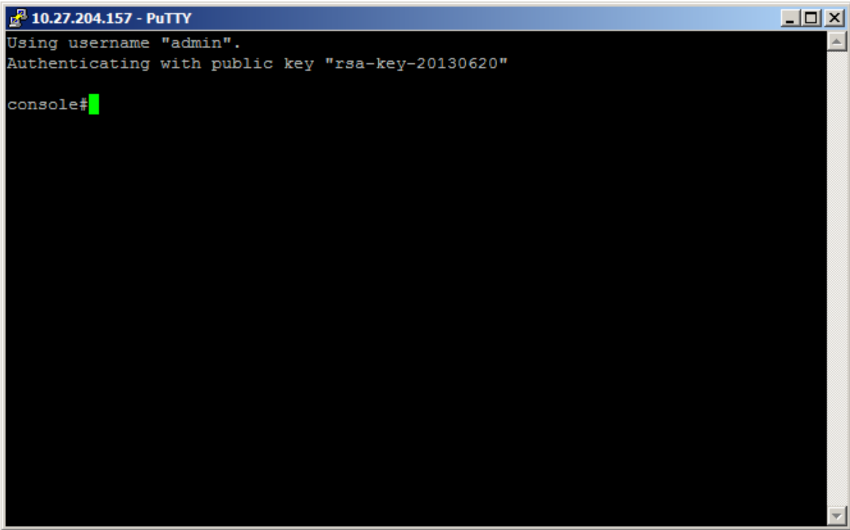


After configuring Putty, be sure to save the configuration. The following screen shows the result of the login process. The user name is entered automatically and the switch confirms that public key authentication occurs.



Authenticating Without a Public Key

When authenticating without the public key, the switch prompts for the user name and password. This is a SSH function, not a switch function. If the user knows the administrator login and password, then they are able to authenticate in this manner.



Associating a User With an SSH Key

The following example shows how to associate a user with an externally generated SSH key. SSH, RSA, or DSA keys can be generated by using the **ssh-keygen** command on a Unix system or with other publicly available utilities.

1 Create the local user:

```
console#config
console(config)#username mylogin password XXXXXXXX privilege
15
```

2 Enter the externally generated key:

```
console(config)#crypto key pubkey-chain ssh
```

3 Associate the key with the newly added user login:

```
console(config-pubkey-chain)#user-key mylogin dsa
```

4 Add the externally generated key. All of the key information is entered between double quotes.

```
console(config-pubkey-key)#key-string "ssh-dss
AAAAB3NzaC1kc3MAAACBAJRwUAD3AuRACp1MObBeh1AgyZb18wf9Btdip+t+1C
bAqiqNEh4lBiew184DSKk0T6SnSSXuCN+bJnQPxJeiQt+OFnmjiYhnHcvI04Q5
KnQhloZcEFgSsmQ7zJnReWtLvUQI0QvBIStanzedmQVGHvDrQ5X2R729ToSH0i
bBrnYtAAAAFQDNord7S9EJvUkKKxVBpWE6/skCmQAAAIBMjMO+BPP5KXzNWfZh
qAhxBSoBvif/z6pzi9xWLlYy99A03zmRYCpcGIoLWiRHsR7NVpxFqwbqvez8KS
0CDJ5aoKKLrpBlpg5ETkYEew/uTZ14lQQRBrzPwGBfxvTXKCWiI2j5KFa/WKLS
nmWJX0/98qpxW/lMXoXsA9iK4pnMKwAAAIB4Jrt6jmoLybpzgOPOI0DsJ7jQwW
acinD0jliz8k+qzCpanhd2wH+DEdj/xO2sFRfnYlME3hmXoB+7NByVUtheVjuQ
2CWhcGFIKm9tbuPC6DtXh1xxT0NJ7rspvLgb0s6y/0tk+94ZP5RCoAtLZ7wirS
hy3/KJ4RE0y2SFZjIVjQ=="

console(config-pubkey-key)#exit
console(config-pubkey-chain)#exit
console(config)#exit
```

5 Use the following command to show the user and SSH association:

```
console#show crypto key pubkey-chain ssh username mylogin
Username : mylogin
ssh-dss
AAAAB3NzaC1kc3MAAACBAJRwUAD3AuRACp1MObBeh1AgyZb18wf9Btdip+t+1C
bAqiqNEh4lBiew184DSKk0T6SnSSXuCN+bJnQPxJeiQt+OFnmjiYhnHcvI04Q5
KnQhloZcEFgSsmQ7zJnReWtLvUQI0QvBIStanzedmQVGHvDrQ5X2R729ToSH0i
bBrnYtAAAAFQDNord7S9EJvUkKKxVBpWE6/skCmQAAAIBMjMO+BPP5KXzNWfZh
qAhxBSoBvif/z6pzi9xWLlYy99A03zmRYCpcGIoLWiRHsR7NVpxFqwbqvez8KS
0CDJ5aoKKLrpBlpg5ETkYEew/uTZ14lQQRBrzPwGBfxvTXKCWiI2j5KFa/WKLS
nmWJX0/98qpxW/lMXoXsA9iK4pnMKwAAAIB4Jrt6jmoLybpzgOPOI0DsJ7jQwW
acinD0jliz8k+qzCpanhd2wH+DEdj/xO2sFRfnYlME3hmXoB+7NByVUtheVjuQ
2CWhcGFIKm9tbuPC6DtXh1xxT0NJ7rspvLgb0s6y/0tk+94ZP5RCoAtLZ7wirS
hy3/KJ4RE0y2SFZjIVjQ==
Fingerprint : d9:d1:21:ad:26:41:ba:43:b1:dc:5c:6c:b9:57:07:6c
```

Authorization

Authorization is used to determine which services the user is allowed to access. For example, the authorization process may assign a user's privilege level, which determines the set of commands the user can execute. There are three kinds of authorization: commands, exec, and network.

- **Commands**: Command authorization determines which CLI commands the user is authorized to execute.
- **Exec**: Exec authorization determines what the user is authorized to do on the switch; that is, the user's privilege level and an administrative profile.
- **Network**: Network authorization enables a RADIUS server to assign a particular 802.1X supplicant to a VLAN. For more information about 802.1X, see "Port and System Security " on page 623.

Table 10-8 shows the valid methods for each type of authorization:

Table 10-8. Authorization Methods

Method	Commands	Exec	Network
local	no	yes	no
none	yes	yes	no
radius	no	yes	yes
tacacs	yes	yes	no

Exec Authorization Capabilities

Dell Networking N-Series switches support two types of service configuration with exec authorization: privilege level and administrative profiles.

Privilege Level

By setting the privilege level during exec authorization, a user can be placed directly into Privileged EXEC mode when they log into the command line interface.

Administrative Profiles

The Administrative Profiles feature allows the network administrator to define a list of rules that control the CLI commands available to a user. These rules are collected in a “profile.” The rules in a profile can define the set of commands, or a command mode, to which a user is permitted or denied access.

Within a profile, rule numbers determine the order in which the rules are applied. When a user enters a CLI command, rules within the first profile assigned to the user are applied in descending order until there is a rule that matches the input. If no rule permitting the command is found, then the other profiles assigned to the user (if any) are searched for rules permitting the command. Rules may use regular expressions for command matching. All profiles have an implicit “deny all” rule, such that any command that does not match any rule in the profile is considered to have been denied by that profile.

A user can be assigned to more than one profile. If there are conflicting rules in profiles, the “permit” rule always takes precedence over the “deny” rule. That is, if any profile assigned to a user permits a command, then the user is permitted access to that command. A user may be assigned up to 16 profiles.

A number of profiles are provided by default. These profiles cannot be altered by the switch administrator. See "Administrative Profiles " on page 266 for the list of default profiles.

If the successful authorization method does not provide an administrative profile for a user, then the user is permitted access based upon the user's privilege level. This means that, if a user successfully passes enable authentication or if exec authorization assigns a privilege level, the user is permitted access to all commands. This is also true if none of the administrative profiles provided are configured on the switch. If some, but not all, of the profiles provided in the authentication are configured on the switch, then the user is assigned the profiles that exist, and a message is logged that indicates which profiles could not be assigned.

The administrative profiles shown in Table 10-9 are system-defined and may not be deleted or altered. To see the rules in a profile, use the **show admin-profiles name** *profile name* command.

Table 10-9. Default Administrative Profiles

Name	Description
network-admin	Allows access to all commands.
network-security	Allows access to network security features such as 802.1X, Voice VLAN, Dynamic ARP Inspection and IP Source Guard.
router-admin	Allows access to Layer 3 features such as IPv4 Routing, IPv6 Routing, OSPF, RIP, etc.
multicast-admin	Allows access to multicast features at all layers, this includes L2, IPv4 and IPv6 multicast, IGMP, IGMP Snooping, etc.
dhcp-admin	Allows access to DHCP related features such as DHCP Server and DHCP Snooping.
CP-admin	Allows access to the Captive Portal feature.
network-operator	Allows access to all User EXEC mode commands and show commands.

Authorization Examples

Authorization allows the administrator to control which services a user is allowed to access. Some of the things that can be controlled with authorization include the user's initial privilege level and which commands the user is allowed to execute. When authorization fails, the user is denied access to the switch, even though the user has passed authentication.

The following examples assume that the configuration used in the previous examples has already been applied.

Local Authorization Example—Direct Login to Privileged EXEC Mode

Apply the following configuration to use the local user database for authorization, such that a user can enter privileged EXEC mode directly:

```
aaa authorization exec “locex” local
line telnet
authorization exec locex
exit
```

With the users that were previously configured, the guest user will still log into user EXEC mode, since the guest user only has privilege level 1 (the default). The admin user will be able to login directly to privileged EXEC mode since his privilege level was configured as 15.

RADIUS Authorization Example—Direct Login to Privileged EXEC Mode

Apply the following configuration to use RADIUS for authorization, such that a user can enter privileged exec mode directly:

```
aaa authorization exec "rad" radius
line telnet
authorization exec rad
exit
```

Configure the RADIUS server so that the RADIUS attribute Service Type (6) is sent with value Administrative. Any value other than Administrative is interpreted as privilege level 1.

The following describes each line in the above configuration:

- The **aaa authorization exec "rad" radius** command creates an exec authorization method list called "rad" that contains the method radius.
- The **authorization exec rad** command assigns the rad exec authorization method list to be used for users accessing the switch via Telnet.

NOTES:

- If the privilege level is zero (that is, blocked), then authorization will fail and the user will be denied access to the switch.
- If the privilege level is higher than one, the user will be placed directly in Privileged EXEC mode. Note that all commands in Privileged EXEC mode require privilege level 15, so assigning a user a lower privilege level will be of no value.
- A privilege level greater than 15 is invalid and treated as if privilege level zero had been supplied.

RADIUS Authorization Example—Administrative Profiles

The switch should use the same configuration as in the previous authorization example.

The RADIUS server should be configured such that it will send the Cisco AV Pair attribute with the “roles” value. For example:

```
shell:roles=router-admin
```

The above example attribute gives the user access to the commands permitted by the router-admin profile.

RADIUS Change of Authorization

Dell Networking N-Series switches support the Change of Authorization Disconnect-Request per RFC 3575. The Dell Networking N-Series switch listens for the Disconnect-Request on UDP port 3799. The Disconnect-Request identifies the user session to be terminated using the following attributes:

- State (IETF attribute #24)
- Acct-Session-Id (IETF attribute #44)
- Calling-Station-Id (IETF attribute #31, which contains the host MAC address)

The following messages from RFC 3575 are supported:

40 – Disconnect-Request

41 – Disconnect-ACK

42 – Disconnect-NAK

A CoA Disconnect-Request terminates the session without disabling the switch port. Instead, CoA Disconnect-Request termination causes reinitialization of the authenticator state machine for the specified host. MAC-based authentication can be enabled for 802.1X sessions in conjunction with CoA. In this case, if the RADIUS server successfully terminates a MAB session and subsequently does not re-authorize the host MAC address to access network resources, the host is effectively denied network access.

If the session cannot be located, the device returns a Disconnect-NAK message with the “Session Context Not Found” error-code attribute. If the session is located, the device terminates the session. After the session has been completely removed, the device returns a Disconnect-ACK message. The attributes returned within a CoA ACK can vary based on the CoA Request.

The administrator can configure whether all or any of the session attributes are used to identify a client session. If all is configured, all session identification attributes included in the CoA Disconnect-Request must match a session or the device returns a Disconnect-NAK or CoA-NAK with the "Invalid Attribute Value" error-code attribute.

Dell Networking N-Series switches support the following attributes in responses:

- State (IETF attribute #24)
- Calling-Station-ID (IETF attribute #31)
- Acct-Session-ID (IETF attribute #44)
- Message-Authenticator (IETF attribute #80)
- Error-Cause (IETF attribute #101)

A CoA NAK message is not sent for all CoA requests with a key mismatch. The message is sent only for the first three requests for a client. After that, all the packets from that client are dropped. When there is a key mismatch, the response authenticator sent with the CoA NAK message is calculated from a dummy key value.

The Dell Networking N-Series switch will start listening to the client again based on the re-authentication timer.

RADIUS COA Example

The following example configures the Dell Networking N-Series switch to listen for and respond to RADIUS COA messages:

1. Configure the switch to use the new model CLI command set. Dell Networking N-Series switches do not support old model commands:

```
console#config
console(config)#aaa new-model
```

2. Configure the switch to listen to RADIUS CoA requests.

```
console(config)#aaa server radius dynamic-author
```

3. Configure a local RADIUS client connection to RADIUS server 10.11.12.13 using the shared secret "secret sauce". The default port number is used.

```
console(config-radius-da)#client 10.11.12.13 server-key
"secret sauce"
```

4 Disconnect-request client identification must match on all keys.

```
console(config-radius-da)#auth-type all
console(config-radius-da)#exit
```

RADIUS COA Example with Telnet and SSH

The following example configures telnet and SSH clients in conjunction with RADIUS CoA.

1 Configure a login list named "login-list" that uses RADIUS as the only method:

```
console#config
console(config)#aaa authentication login "login-list" radius
```

2 Enable RADIUS COA:

```
console(config)#aaa server radius dynamic-author
```

3 Enable the switch RADIUS client connecting to the RADIUS server at 10.130.191.89:

```
console(config-radius-da)#client 10.130.191.89 server-key
"shared secret"
```

4 Allow matching of the client session on any of the key values:

```
console(config-radius-da)#auth-type any
console(config-radius-da)#exit
```

5 Configure the RADIUS server attribute 4 (NAS-IP-Address). This attribute is sent in the RADIUS message to the RADIUS server but does not change the source IP address sent in the RADIUS messages. It allows a group of NASs to simulate a large RADIUS NAS:

```
console(config)#radius-server attribute 4 10.130.65.4
```

6 Configure the remote RADIUS server address with name Default-RADIUS-Server and key "shared secret":

```
console(config)#radius-server host auth 10.130.191.89
console(Config-auth-radius)#name "Default-RADIUS-Server"
console(Config-auth-radius)#key "shared secret"
console(Config-auth-radius)#exit
```

7 Configure telnet sessions to the switch to use RADIUS authentication (the only login-list method):

```
console(config)#line telnet
console(config-telnet)#login authentication login-list
console(config-telnet)#exit
```

8 Configure SSH sessions to the switch to use RADIUS authentication:

```
console(config)#line ssh
console(config-ssh)#login authentication login-list
console(config-ssh)#exit
```

9 Enable the SSH server (the telnet server is enabled by default):

```
console(config)#ip ssh server
```

TACACS Authorization

TACACS+ Authorization Example—Direct Login to Privileged EXEC Mode

Apply the following configuration to use TACACS+ for authorization, such that a user can enter privileged EXEC mode directly:

1. Create an exec authorization method list called "tacex" which contains the method tacacs.

```
console#config
console(config)#aaa authorization exec "tacex" tacacs
```

2. Assign the tacex exec authorization method list to be used for users accessing the switch via Telnet.

```
console(config)#line telnet
console(config-telnet)#authorization exec tacex
console(config-telnet)#exit
```

3. Configure the TACACS+ server so that the shell service is enabled and the priv-lvl attribute is sent when user authorization is performed. For example:

   shell:priv-lvl=15

NOTES:

- If the privilege level is zero (that is, blocked), then authorization will fail and the user will be denied access to the switch.
- If the privilege level is higher than one, the user will be placed directly in Privileged EXEC mode. Note that all commands in Privileged EXEC mode require privilege level 15, so assigning a user a lower privilege level will be of no value.
- A privilege level greater than 15 is invalid and treated as if privilege level zero had been supplied.
- The shell service must be enabled on the TACACS+ server. If this service is not enabled, authorization will fail and the user will be denied access to the switch.

TACACS+ Authorization Example—Administrative Profiles

The switch should use the same configuration as for the previous authorization example.

The TACACS+ server should be configured such that it will send the "roles" attribute. For example:

```
shell:roles=router-admin
```

The above example attribute will give the user access to the commands permitted by the router-admin profile.

NOTE: If the priv-lvl attribute is also supplied, the user can also be placed directly into privileged EXEC mode.

TACACS+ Authorization Example—Custom Administrative Profile

This example creates a custom profile that allows the user to control user access to the switch by configuring a administrative profile that only allows access to AAA related commands. Use the following commands to create the administrative profile:

1 Create an administrative profile called “aaa” and place the user in admin-profile-config mode.

```
console#config
console(config)#admin-profile aaa
```

2 Enter **rule** *number* **permit command** *regex* commands to allows any command that matches the regular expression.

   The command rules use regular expressions as implemented by Henry Spencer's regex library (the POSIX 1003.2 compliant version). In the regular expressions used in this example, the caret (^) matches the null string at the beginning of a line, the period (.) matches any single character, and the asterisk (*) repeats the previous match zero or more times.

```
console(config)#rule 99 permit command "^show aaa .*"
console(admin-profile)#rule 98 permit command "^show
authentication .*"
console(admin-profile)#rule 97 permit command "^show
authorization .*"
console(admin-profile)#rule 96 permit command "^show
accounting .*"
console(admin-profile)#rule 95 permit command "^show tacacs
.*"
console(admin-profile)#rule 94 permit command "^aaa .*"
console(admin-profile)#rule 93 permit command "^line .*"
console(admin-profile)#rule 92 permit command "^login .*"
console(admin-profile)#rule 91 permit command "^authorization
.*"
console(admin-profile)#rule 90 permit command "^accounting .*"
console(admin-profile)#rule 89 permit command "^configure .*"
```

```
console(admin-profile)#rule 88 permit command "^password .*"
console(admin-profile)#rule 87 permit command "^username .*"
console(admin-profile)#rule 86 permit command "^show user.*"
console(admin-profile)#rule 85 permit command "^radius-server
.*"
console(admin-profile)#rule 84 permit command "^tacacs-server
.*"
```

3 Enter **rule** *number* **permit mode** *mode-name* commands to allows all commands in the named mode.

```
console(admin-profile)#rule 83 permit mode radius-auth-config
console(admin-profile)#rule 82 permit mode radius-acct-config
console(admin-profile)#rule 81 permit mode tacacs-config
console(admin-profile)#exit
```

4 Assign this profile to a user by configuring the TACACS+ server so that it sends the following "roles" attribute for the user:

```
shell:roles=aaa
```

If it is desired to also permit the user access to network-operator commands (basically, all the command in User EXEC mode), then the "roles" attribute would be configured as follows:

```
shell:roles=aaa,network-operator
```

TACACS+ Authorization Example—Per-command Authorization

An alternative method for command authorization is to use the TACACS+ feature of per-command authorization. With this feature, every time the user enters a command, a request is sent to the TACACS+ server to ask if the user is permitted to execute that command. Exec authorization does not need to be configured to use per-command authorization.

Apply the following configuration to use TACACS+ to authorize commands:

1 Creates a command authorization method list called "taccmd" that includes the method tacacs.

```
console#config
console(config)#aaa authorization commands "taccmd" tacacs
```

- Assigns the taccmd command authorization method list to be used for users accessing the switch via Telnet.

```
console(config)#line telnet
console(config-telnet)#authorization commands taccmd
console(config-telnet)#exit
```

The TACACS+ server must be configured with the commands that the user is allowed to execute. If the server is configured for command authorization as "None", then no commands will be authorized. If both administrative profiles and per-command authorization are configured for a user, any command must be permitted by both the administrative profiles and by per-command authorization.

TACACS Authorization—Privilege Level

Dell Networking TACACS supports setting the maximum user privilege level in the TACACS authorization response. Configure the TACACS server to send priv-lvl=X, where X is either 1 (Non-privileged mode), or 15 (Privileged mode).

Accounting

Accounting is used to record security events, such as a user logging in or executing a command. Accounting records may be sent upon completion of an event (stop-only) or at both the beginning and end of an event (start-stop). There are three types of accounting: commands, Dot1x, and exec.

- **Commands**—Sends accounting records for command execution.
- **Dot1x**—Sends accounting records for network access.
- **Exec**—Sends accounting records for management access (logins).

For more information about the data sent in accounting records, see "Which RADIUS Attributes Does the Switch Support? " on page 243 and "Using TACACS+ Servers to Control Management Access " on page 246.

Table 10-10 shows the valid methods for each type of accounting:

Table 10-10. Accounting Methods

Method	Commands	Dot1x	Exec
radius	no	yes	yes
tacacs	yes	no	yes

RADIUS Accounting

Dell Networking N-Series switches support RADIUS accounting. The supported accounting types are start-only or start-stop.

The following attributes are sent in the Accounting Stop record that is sent to the RADIUS server when the switch is configured for 802.1X accounting:

- User-Name (1)
- Called-Station-Id (30)
- Calling-Station-Id (31)
- NAS-Port-Type (61)
- State (24)
- Acct-Terminate-Cause(49)
- NAS-IP-Address(4)
- Class (25)

- Acct-Session Time(46)
- Acct-Input-Octets (42)
- Acct-Output-Octets (43)
- Acct-Input-Gigawords(52)
- Acct-Output-Gigawords (53)

Certain of the attributes above are sent only if received from the RADIUS server during the Access Request process, e.g., Class or State.

The following attributes are sent in the Accounting Start record sent to the RADIUS server when the switch is configured for 802.1x accounting:

- User-Name (1)
- Called-Station-Id (30)
- Calling-Station-Id (31)
- NAS-Port-Type (61)
- State (24)
- Class (25)
- NAS-IP-Address(4)
- Tunnel-Private-Group-Id(81) - VLAN ID

IEEE 802.1X

What is IEEE 802.1X?

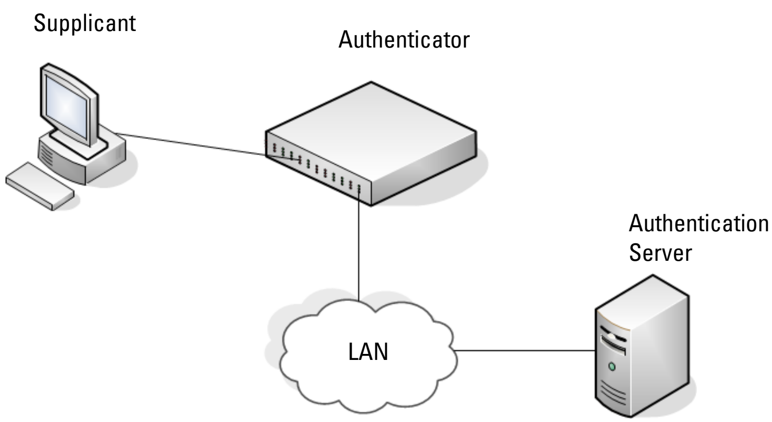
The IEEE 802.1X standard provides a means of preventing unauthorized access by supplicants (clients) to the services the switch offers, such as access to the LAN.

The 802.1X network has three components:

- **Supplicant** — The client connected to the authenticated port that requests access to the network.
- **Authenticator** — The network device that prevents network access prior to authentication.
- **Authentication Server** — The network server (such as a RADIUS server) that performs the authentication on behalf of the authenticator, and indicates whether the user is authorized to access system services. Dell Networking supports interoperability with a variety of external authentication servers. Refer to "Authentication, Authorization, and Accounting " on page 229 for more information.

Figure 10-3 shows the 802.1X network components.

Figure 10-3. IEEE 802.1X Network



As shown in Figure 10-3, the Dell Networking N1500, N2000, N3000, or N4000 Series switch, is the authenticator and ensures that the supplicant (a PC) that is attached to an 802.1X-controlled port is authenticated by an authentication server (a RADIUS server). The result of the authentication process determines whether the supplicant is authorized to access network services on that controlled port. Dell Networking N-Series switches support authentication using remote RADIUS or TACACS servers and also support authentication using a local authentication service.

Supported security methods for supplicant communication with remote authentication servers include MD5, PEAP, EAP-TTL, EAP-TTLS, and EAP-TLS. Only EAP-MD5 is supported when using the local authentication server (IAS) for communication with the supplicant.

For a list of RADIUS attributes that the switch supports, see "Using RADIUS " on page 241.

What are the 802.1X Port Authentication Modes?

The 802.1X port authentication mode determines whether to allow or prevent network traffic on the port. A port can configured to be in one of the following 802.1X authentication modes:

- Auto (default)
- MAC-based
- Force-authorized
- Force-unauthorized.

These modes control the behavior of the port. The port state is either Authorized or Unauthorized.

If the port is in the force-authorized mode, the port state is Authorized and the port sends and receives normal traffic without client port-based authentication. When a port is in a forced-unauthorized mode, the port state is Unauthorized and the port ignores supplicant authentication attempts and does not provide authentication services. By default, when 802.1X is globally enabled on the switch, all ports are in auto authentication mode, which means the port will be unauthorized until a successful authentication exchange has taken place. Auto mode is suitable for authentication of a single supplicant attached to a port.

In addition to force-authorized, force-unauthorized, and auto modes, the 802.1X mode of a port can be MAC based, as the following section describes.

NOTE: Only MAC-Based and Auto modes actually use 802.1X to authenticate. Authorized and Unauthorized modes are manual overrides.

What is MAC-Based 802.1X Authentication?

MAC-based authentication allows multiple supplicants connected to the same port to authenticate individually. For example, a 5-port hub might be connected to a single port on the switch. Each host connected to the hub must authenticate separately in order to gain access to the network. Hosts that do not authenticate (or are not configured with MAB or a guest or unauthenticated VLAN) are denied access to the network.

The hosts are distinguished by their MAC addresses. Internally, the switch adds an ACL to the port to allow packets from the host MAC address to pass into the switch.

NOTE: By default, all ports are in switchport access mode. A port that uses MAC-based authentication must be configured to be in General mode.

When multiple hosts (for example, a PC, a printer, and a phone in the same office) are connected to the switch on the same port, each of the connected hosts authenticates separately with the RADIUS server.

If a port uses MAC-based 802.1X authentication, the option to use MAC Authentication Bypass (MAB) is available. MAB is a supplemental authentication mechanism that allows 802.1X unaware clients – such as printers, fax machines, and some IP phones — to authenticate to the network using the client MAC address as an identifier.

The known and allowable MAC address and corresponding access rights of the client must be pre-populated in the authentication server.

When a port configured for MAB receives traffic from an unauthenticated client, the switch (Authenticator):

- Sends a EAP Request packet to the unauthenticated client
- Waits a pre-determined period of time for a response
- Retries – resends the EAP Request packet up to three times

- Considers the client to be 802.1X unaware client (if it does not receive an EAP response packet from that client)

The authenticator sends a request to the authentication server with the MAC address of the client in a hexadecimal format as the username and the MD5 hash of the MAC address as the password. The authentication server checks its database for the authorized MAC addresses and returns an Access-Accept or an Access-Reject response, depending on whether the MAC address is found in the database. MAB also allows 802.1X-unaware clients to be placed in a RADIUS-assigned VLAN or to apply a specific Filter ID to the client traffic.

NOTE: MAB initiates only after the dot1x guest VLAN period times out. If the client responds to any of the EAPOL identity requests, MAB does not initiate for that client.

What is the Role of 802.1X in VLAN Assignment?

Dell Networking N-Series switches allow a port to be placed into a particular VLAN based on the result of the authentication. The authentication server can provide information to the switch about which VLAN to assign the supplicant or the administrator can configure the level of access provided when authentication fails or is never attempted.

When a host connects to a switch that uses an authentication server to authenticate, the host authentication will have one of three outcomes:

- The host is authenticated.
- The host attempts to authenticate but fails because it lacks certain security credentials.
- The host does not try to authenticate at all (802.1X unaware).

Three separate VLANs can be created on the switch to handle a host depending on whether the host authenticates, fails the authentication, or does not attempt authentication. The RADIUS server informs the switch of the selected VLAN as part of the authentication.

Authenticated VLANs

Hosts that authenticate normally use a VLAN that includes access to network resources. This VLAN may be assigned by the RADIUS server. Hosts that fail authentication might be denied access to the network or placed into a guest

VLAN. Host that do not attempt authentication may be placed into an unauthenticated VLAN. The network administrator can configure the type of access provided to the authenticated, guest, and unauthenticated VLANs.

Much of the configuration to assign authenticated hosts to a particular VLAN takes place on the 802.1X authenticator server (for example, a RADIUS server). If an external RADIUS server is used to manage VLANs, configure the server to use Tunnel attributes in Access-Accept messages in order to inform the switch about the selected VLAN. These attributes are defined in RFC 2868 and their use for dynamic VLAN is specified in RFC 3580.

The VLAN attributes defined in RFC3580 and required for VLAN assignment via RADIUS are as follows:

- Tunnel-Type (64) = VLAN (13)
- Tunnel-Medium-Type (65) = 802 (6)
- Tunnel-Private-Group-ID (81) = VLANID

The tag value for the Tunnel-Private-Group-ID is parsed as the length of the VLAN ID. The VLAN ID may consist of a VLAN name (not to exceed 32 characters) or a numeric value in ASCII (no alphabetic characters are allowed) in the range 1–4093.

Dynamic VLAN Creation

If RADIUS-assigned VLANs are enabled though the Authorization Network RADIUS configuration option, the RADIUS server is expected to include the VLAN ID in the 802.1X tunnel attributes of its response message to the switch. If dynamic VLAN creation is enabled on the switch and the RADIUS-assigned VLAN does not exist, then the assigned VLAN is dynamically created and the port is made a member of the VLAN. If the VLAN is already created on the switch, the port is simply made a member of the VLAN. This implies that the client can connect from any port and be assigned to the appropriate VLAN based on the RADIUS server configuration. This gives flexibility for clients to move around the network without much additional configuration required on the switches in the network. Dynamic VLAN assignment requires that the port be configured in general mode if the port authentication mode is MAC-based and be configured in general or access mode if the port authentication mode is auto.

Unauthenticated VLAN

The network administrator may choose to configure an unauthorized VLAN. Hosts that attempt authentication and fail three times are placed in the unauthenticated VLAN. Once in the unauthenticated VLAN, authentication is not reattempted until:

- the re-authentication timer expires
- the supplicant disconnects from the port
- the port is shut down and re-enabled

The number of re-authentication failures required to place a supplicant in the unauthenticated VLAN is not configurable.

The network administrator can configure the unauthenticated VLAN to provide the desired level of network access, i.e., a black hole or a guest VLAN type of access.

Guest VLAN

The Guest VLAN feature provides a mechanism to allow users access to a guest VLAN. For example, the administrator might provide a guest VLAN to visitors and contractors to permit network access that allows visitors to connect to external network resources, such as the Internet, with no ability to browse information on the internal LAN.

On a port configured in auto authentication mode (**dot1x port-control auto**), connected to a client that does not support 802.1X, the client does not respond to the 802.1X requests from the switch. The port remains in the unauthorized state and the client is not granted access to the network. If a guest VLAN is configured for that port, the port is placed in the configured guest VLAN and the port is moved to the authorized state, allowing access to the client over the guest VLAN.

NOTE: MAB and the guest VLAN feature are mutually exclusive on a port. If MAB is enabled on a port concurrently with guest VLAN, the port will not move to the authorized state.

When the guest VLAN is disabled, users authorized by the guest VLAN are removed from the VLAN and denied network access.

What is Monitor Mode?

The monitor mode is a special mode that can be enabled in conjunction with 802.1X authentication. Monitor mode provides a way for network administrators to identify possible issues with the 802.1X configuration on the switch without affecting the network access to the users of the switch. It allows network access even in case where there is a failure to authenticate but logs the results of the authentication process for diagnostic purposes.

The monitor mode can be configured globally on a switch. If the switch fails to authenticate a user for any reason (for example, RADIUS access reject from RADIUS server, RADIUS timeout, or the client itself is dot1x-unaware), the client is authenticated and is undisturbed by the failure condition(s). The reasons for failure are logged for tracking purposes.

Table 10-11 provides a summary of the 802.1X Monitor Mode behavior.

Table 10-11. IEEE 802.1X Monitor Mode Behavior

Case	Sub-case	Regular Dot1x	Dot1x Monitor Mode
RADIUS/IAS Success	Success	Port State: Permit VLAN: Assigned Filter: Assigned	Port State: Permit VLAN: Assigned Filter: Assigned
	Incorrect NAS Port	Port State: Deny	Port State: Permit VLAN: Default PVID of the port
	Invalid VLAN Assignment	Port State: Deny	Port State: Permit VLAN: Default PVID of the port
	Invalid Filter-id	Port State: Deny	Port State: Permit VLAN: Default PVID of the port
	Bad RADIUS packet	Port State: Deny	Port State: Permit VLAN: Default PVID of the port
RADIUS/IAS Failure	Default behavior	Port State: Deny	Port State: Permit VLAN: Default PVID of the port
	Unauth VLAN enabled	Port State: Permit VLAN: Unauth	Port State: Permit VLAN: Unauth

Table 10-11. IEEE 802.1X Monitor Mode Behavior (Continued)

Case	Sub-case	Regular Dot1x	Dot1x Monitor Mode
RADIUS Timeout	Default behavior	Port State: Deny	Port State: Permit VLAN: Default PVID of the port
	Unauth VLAN enabled	Port State: Deny	Port State: Permit VLAN: Unauth
EAPOL Timeout	Default behavior	Port State: Deny	Port State: Permit
3 × EAPOL Timeout (Guest VLAN timer expiry or MAB timer expiry)	Guest VLAN enabled	Port State: Permit VLAN: Guest	Port State: Permit VLAN: Guest
	MAB Success Case	Port State: Permit VLAN: Assigned Filter: Assigned	Port State: Permit VLAN: Assigned Filter: Assigned
	MAB Fail Case	Port State: Deny	Port State: Permit VLAN: Default PVID of the port
Supplicant Timeout		Port State: Deny	Port State: Deny
Port/Client Authenticated on Guest VLAN	Delete Guest VLANID through Dot1Q	Port State: Deny	Port State: Permit VLAN: Default PVID of the port

How Does the Authentication Server Assign DiffServ Policy?

The Dell Networking N-Series switches allow the external 802.1X Authenticator or RADIUS server to assign DiffServ policies to users that authenticate to the switch. When a host (supplicant) attempts to connect to the network through a port, the switch contacts the 802.1X authenticator or RADIUS server, which then provides information to the switch about which DiffServ policy to assign the host (supplicant). The application of the policy is applied to the host after the authentication process has completed. The Diffserv policy is always applied for the "in" direction of the interface and applies to the interface as a whole.

For additional guidelines about using an authentication server to assign DiffServ policies, see "Configuring Authentication Server DiffServ Policy Assignments " on page 310.

What is the Internal Authentication Server?

The Internal Authentication Server (IAS) is a dedicated database for localized authentication of users for network access through 802.1X. In this database, the switch maintains a list of username and password combinations to use for 802.1X authentication. Entries can be created in the database manually, or the IAS information can be uploaded to the switch.

If the authentication method for 802.1X is IAS, the switch uses the locally stored list of username and passwords to provide port-based authentication to users instead of using an external authentication server. Authentication using the IAS supports the EAP-MD5 method only.

NOTE: The IAS database does not handle VLAN assignments or DiffServ policy assignments.

Default 802.1X Values

Table 10-12 lists the default values for the 802.1X features.

Table 10-12. Default Port-Based Security Values

Feature	Description
Global 802.1X status	Disabled
802.1X authentication method	None
Per-port 802.1X status	Disabled
Port authentication mode	Auto mode
Port authentication state	Unauthorized
Periodic reauthentication	Disabled
Seconds between reauthentication attempts	3600
Authentication server timeout	30 seconds
Resending EAP identity Request	30 seconds

Table 10-12. Default Port-Based Security Values

Feature	Description
Quiet period	60 seconds
Supplicant timeout	30 seconds
Max EAP request	2 times
Maximum number of supplicants per port for MAC-based authentication mode	48
Guest VLAN	Disabled
Unauthenticated VLAN	Disabled
Dynamic VLAN creation	Disabled
RADIUS-assigned VLANs	Disabled
IAS users	none configured
Port security	Unlocked
Port security traps	Disabled
Maximum learned MAC addresses	100 (when locked)
Monitor mode	Disabled

Configuring IEEE 802.1X (Web)

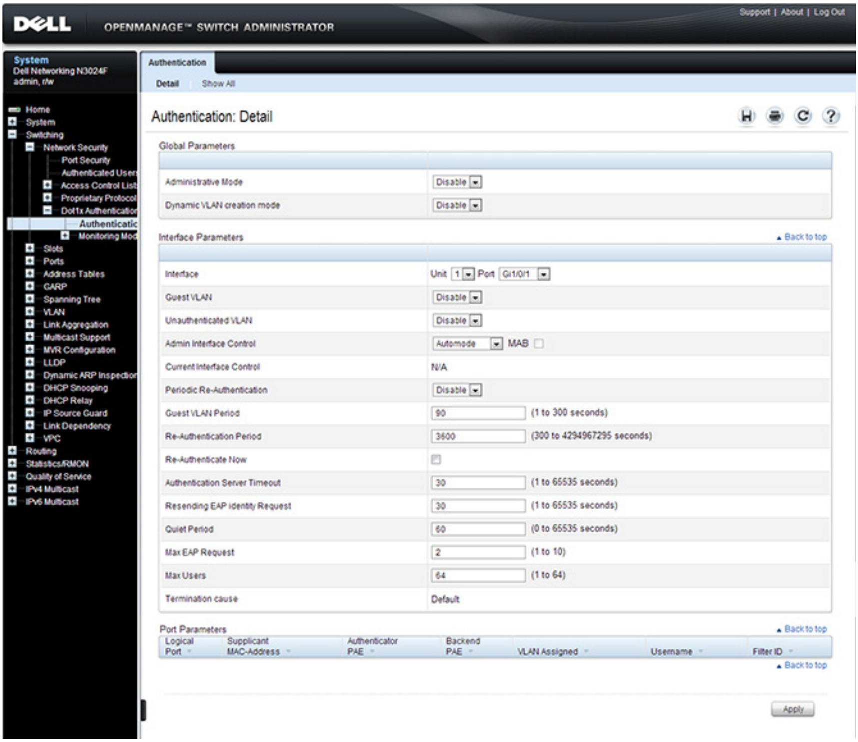
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring the IEEE 802.1X features and Port Security on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

Dot1x Authentication

Use the **Dot1x Authentication** page to configure the 802.1X administrative mode on the switch and to configure general 802.1X parameters for a port.

To display the **Dot1x Authentication** page, click **Switching → Network Security → Dot1x Authentication → Authentication** in the navigation panel.

Figure 10-4. Dot1x Authentication



To configure 802.1X authentication on multiple ports:

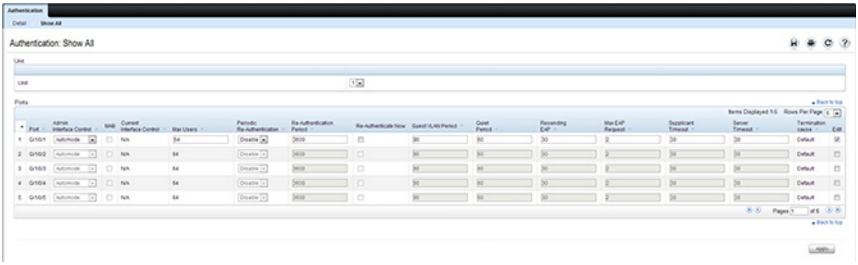
1 Open the **Dot1x Authentication** page.

2 Click **Show All** to display the **Dot1x Authentication Table** page.

3 In the Ports list, select the check box in the **Edit** column for the port to configure.

4 Select the desired settings to change for all ports that are selected for editing.

Figure 10-5. Configure Dot1x Settings



5 Click **Apply**.

To reauthenticate a port:

1 Open the **Dot1x Authentication** page.

2 Click **Show All**.

The **Dot1x Authentication Table** displays.

3 Check **Edit** to select the Unit/Port to re-authenticate.

4 Check **Reauthenticate Now**.

5 Click **Apply**.

The authentication process is restarted on the specified port.

To reauthenticate multiple ports:

1 Open the **Dot1x Authentication** page.

2 Click **Show All**.

The **Dot1x Authentication Table** displays.

3 Check **Edit** to select the Units/Ports to re-authenticate.

4 To re-authenticate on a periodic basis, set **Periodic Re-Authentication** to Enable, and specify a **Re-Authentication Period** for all desired ports.

5 To re-authenticate immediately, check **Reauthenticate Now** for all ports to be re-authenticated.

6 Click **Apply**.

The authentication process is restarted on the specified ports (either immediately or periodically).

To change the administrative port control:

1 Open the **Dot1x Authentication** page.

2 Click **Show All**.

   The **Dot1x Authentication Table** displays.

3 Scroll to the right side of the table and select the **Edit** check box for each port to configure. Change **Admin Port Control** to **Authorized**, **Unauthorized**, or **Automode** as needed for chosen ports. Only **MAC-Based** and **Automode** actually use 802.1X to authenticate. **Authorized** and **Unauthorized** are manual overrides.

4 Click **Apply**.

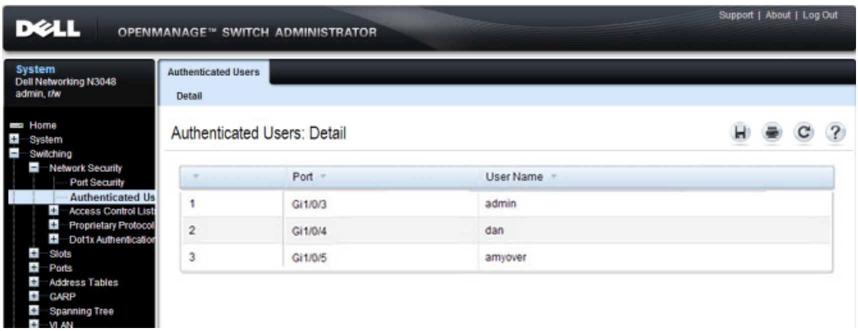
   Admin Port Control is updated for the specified ports, and the device is updated.

Authenticated Users

The **Authenticated Users** page is used to display lists of ports that have authenticated users.

To display the **Authenticated Users** page, click **Switching** → **Network Security** → **Authenticated Users** in the navigation panel.

Figure 10-6. Network Security Authenticated Users



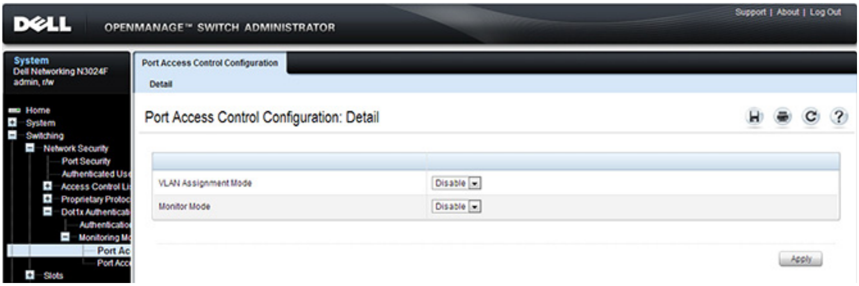
Port Access Control Configuration

Use the **Port Access Control Configuration** page to globally enable or disable RADIUS-assigned VLANs and to enable Monitor Mode to help troubleshoot 802.1X configuration issues.

NOTE: The VLAN Assignment Mode field is the same as the Admin Mode field on the **System** → **Management Security** → **Authorization Network RADIUS** page.

To display the **Port Access Control Configuration** page, click **Switching** → **Network Security** → **Dot1x Authentication** → **Monitor Mode** → **Port Access Control Configuration** in the navigation panel.

Figure 10-7. Port Access Control Configuration

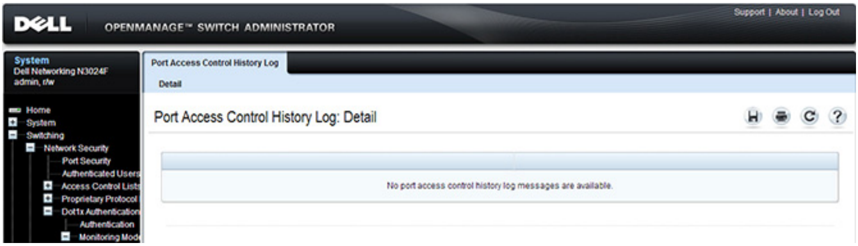


Port Access Control History Log Summary

Use the **Port Access Control History Log Summary** page to view log messages about 802.1X client authentication attempts. The information on this page can help you troubleshoot 802.1X configuration issues.

To display the **Port Access Control History Log Summary** page, click **Port Access Control Configuration** page, click **Switching** → **Network Security** → **Dot1x Authentication** → **Monitor Mode** → **Port Access Control History Log Summary** in the navigation panel.

Figure 10-8. Port Access Control History Log Summary

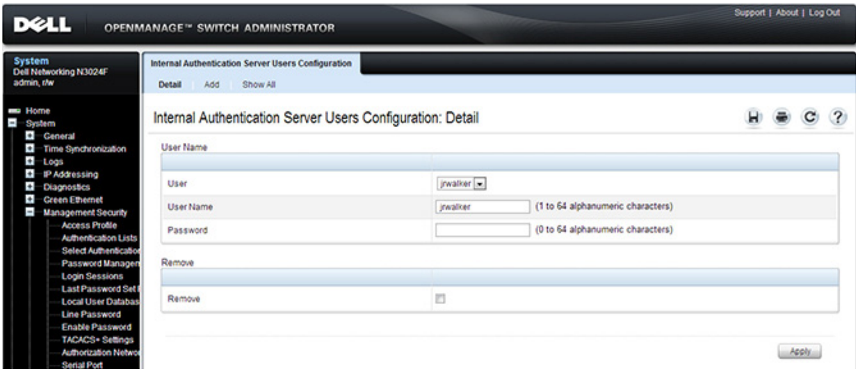


Internal Authentication Server Users Configuration

Use the **Internal Authentication Server Users Configuration** page to add users to the local IAS database and to view the database entries.

To display the **Internal Authentication Server Users Configuration** page, click **System** → **Management Security** → **Internal Authentication Server Users Configuration** in the navigation panel.

Figure 10-9. Internal Authentication Server Users Configuration



NOTE: If no users exist in the IAS database, the IAS Users Configuration Page does not display the fields shown in the image.

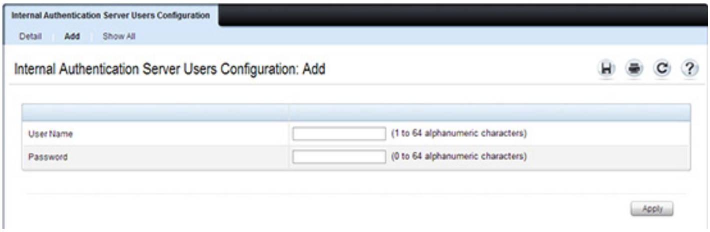
To add IAS users:

1 Open the **Internal Authentication Server Users Configuration** page.

2 Click **Add** to display the **Internal Authentication Server Users Add** page.

3 Specify a username and password in the appropriate fields.

Figure 10-10. Adding an IAS User



4 Click **Apply**.

To view the Internal Authentication Server Users Table page, click **Show All**.

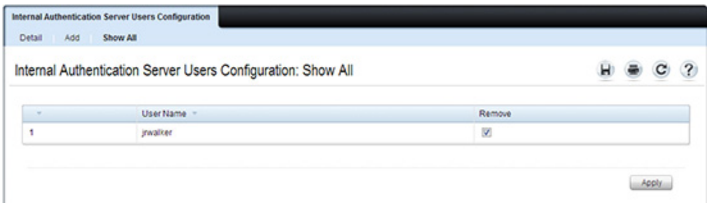
To delete an IAS user:

1 Open the **Internal Authentication Server Users Configuration** page.

2 From the User menu, select the user to remove, select the user to remove.

3 Select the **Remove** check box.

Figure 10-11. Removing an IAS User



4 Click **Apply**.

Configuring IEEE 802.1X (CLI)

This section provides information about commands you use to configure 802.1X and Port Security settings. For additional information about the commands in this section, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Basic 802.1X Authentication Settings

Beginning in Privileged EXEC mode, use the following commands to enable and configure 802.1X authentication on the switch.

Command	Purpose
configure	Enter Global Configuration mode.
aaa accounting dot1x default	Sets 802.1X accounting to the default operational mode
aaa authentication dot1x default *method1*	Specify the authentication method to use to authenticate 802.1X clients that connect to the switch. *method1*—The method keyword can be **radius**, **none, or ias**.
dot1x system-auth-control	Globally enable 802.1X authentication on the switch.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.

Command	Purpose
dot1x port-control {**force-authorized** | **force-unauthorized** | **auto** | **mac-based**}	Specify the 802.1X mode for the port. **NOTE:** For standard 802.1X implementations in which one client is connected to one port, use the **dot1x port-control auto** command to enable 802.1X authentication on the port. • **auto** — Enables 802.1X authentication on the interface and causes the port to transition to the authorized or unauthorized state based on the 802.1X authentication exchange between the switch and the client. • **force-authorized** — Disables 802.1X authentication on the interface and causes the port to transition to the authorized state without any authentication exchange required. The port sends and receives normal traffic without 802.1X-based authentication of the client. • **force-unauthorized** — Denies all access through this interface by forcing the port to transition to the unauthorized state, ignoring all attempts by the client to authenticate. The switch cannot provide authentication services to the client through the interface. • **mac-based** — Enables 802.1X authentication on the interface and allows multiple hosts to authenticate on a single port. The hosts are distinguished by their MAC addresses.
dot1x mac-auth-bypass	If the 802.1X mode on the interface is **mac-based**, this command can optionally be used to enable MAB on an interface.
CTRL + Z	Exit to Privileged EXEC mode.
show dot1x	View the current 802.1X status.
show dot1x clients {**all** | *interface*}	View information about 802.1X clients that have successfully authenticated and are connected to the switch. The *interface* variable includes the interface type and number.
show dot1x users [**username** *username*]	View the 802.1X authenticated users for the switch.

NOTE: To enable 802.1X Monitor Mode to help troubleshoot authentication issues, use the **dot1x system-auth-control monitor** command in Global Configuration mode. To view 802.1X authentication events and information, use the **show dot1x authentication-history** {*interface* | **all**} [**failed-auth-only**] [**detail**] command in Privileged EXEC mode. To clear the history, use the **clear dot1x authentication-history** command.

Configuring Additional 802.1X Interface Settings

Beginning in Privileged EXEC mode, use the following commands to configure 802.1X interface settings such as the reauthentication period and switch-to-client retransmission time.

Command	Purpose
configure	Enter Global Configuration mode.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
dot1x reauthentication	Enable periodic re-authentication of the client.
dot1x timeout re-authperiod *seconds*	Set the number of seconds between re-authentication attempts.
dot1x timeout server-timeout *seconds*	Set the time that the switch waits for a response from the authentication server.
dot1x timeout tx-period *seconds*	Set the number of seconds that the switch waits for a response to an Extensible Authentication Protocol (EAP)-request/identity frame from the client before resending the request.
dot1x timeout quiet-period *seconds*	Set the number of seconds that the switch remains in the quiet state following a failed authentication exchange (for example, the client provided an invalid password).

Command	Purpose
dot1x timeout supp-timeout *seconds*	Set the time that the switch waits for a response before retransmitting an Extensible Authentication Protocol (EAP)-request frame to the client.
dot1x max-req *count*	Set the maximum number of times that the switch sends an Extensible Authentication Protocol (EAP)-request frame (assuming that no response is received) to the client before restarting the authentication process.
dot1x max-users *users*	Set the maximum number of clients supported on the port when MAC-based 802.1X authentication is enabled on the port.
CTRL + Z	Exit to Privileged EXEC mode.
dot1x re-authenticate [*interface*]	Manually initiate the re-authentication of all 802.1X-enabled ports or on the specified 802.1X-enabled port. The *interface* variable includes the interface type and number.
dot1x initialize [*interface*]	Start the initialization sequence on all ports or on the specified port. **NOTE:** This command is valid only if the port-control mode for the specified port is auto or MAC-based.
show dot1x [interface *interface]*	View 802.1X settings for the switch or for the specified interface.
show dot1x interface *interface* **statistics**	View 802.1X statistics for the specified interface.

Configuring 802.1X Settings for RADIUS-Assigned VLANs

Beginning in Privileged EXEC mode, use the following commands to configure 802.1X settings that affect the RADIUS-assigned VLAN.

Command	Purpose
configure	Enter Global Configuration mode.
aaa authorization network default radius	Allow the RADIUS server to assign VLAN IDs to clients.

Command	Purpose
dot1x dynamic-vlan enable	If the RADIUS assigned VLAN does not exist on the switch, allow the switch to dynamically create the assigned VLAN.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
dot1x guest-vlan *vlan-id*	Specify the guest VLAN.
dot1x unauth-vlan *vlan-id*	Specify the unauthenticated VLAN. The VLAN must already have been created.
CTRL + Z	Exit to Privileged EXEC mode.
show dot1x advanced *interface*	View the current 802.1X configuration.

NOTE: When dynamically creating VLANs, the uplink port should be in trunk mode so that it will automatically participate in all dynamically-created VLANs. Otherwise, the supplicant may be placed in a VLAN that does not go beyond the switch because no other ports are participating.

Configuring Internal Authentication Server Users

Beginning in Privileged EXEC mode, use the following commands to add users to the IAS database and to use the database for 802.1X authentication.

Command	Purpose
configure	Enter Global Configuration mode.
aaa ias-user username *user*	Add a user to the IAS user database. This command also changes the mode to the AAA User Config mode.
password *password* [**encrypted**]	Configure the password associated with the user.
CTRL + Z	Exit to Privileged EXEC mode.
show aaa ias-users	View all configured IAS users.
clear aaa ias-users	Delete all IAS users from the database.

IEEE 802.1X Configuration Examples

This section contains the following examples:

- Configuring 802.1X Authentication
- Controlling Authentication-Based VLAN Assignment
- Allowing Dynamic Creation of RADIUS-Assigned VLANs
- Configuring Authentication Server DiffServ Policy Assignments

Configuring 802.1X Authentication

The network in this example requires clients to use 802.1X authentication to access the network through the switch ports. The administrator must configure the following settings on systems other than the switch before configuring the switch:

1. Add the users to the client database on the Authentication Server, such as a RADIUS server with Cisco® Secure Access Control Server (ACS) software.
2. Configure the settings on the client, such a PC running Microsoft® Windows, to require 802.1X authentication.

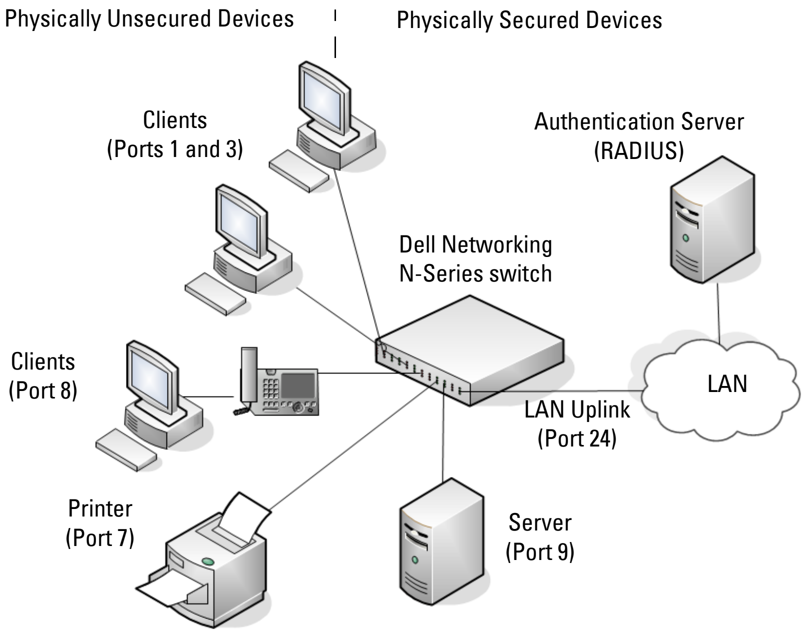
The switch uses an authentication server with an IP address of 10.10.10.10 to authenticate clients. Port 7 is connected to a printer in the unsecured area. The printer is an 802.1X unaware client, so Port 7 is configured to use MAC-based authentication with MAB.

NOTE: The printer requires an entry in the client database that uses the printer MAC address as the username.

An IP phone is directly connected to Port 8, and a PC is connected to the IP phone. Both devices are authenticated through MAC-based authentication, which allows multiple hosts to authenticate on a single port. The hosts are distinguished by their MAC addresses, and hosts authenticate separately with the RADIUS server.

Port 9 is connected to a server in a part of the network that has secure physical access (i.e. the doors to the wiring closet and data center are locked), so this port is set to the Authorized state, meaning that the device connected to this port does not need to authenticate using 802.1X. Port 24 is the uplink to a router and is also in the Authorized state.

Figure 10-12. 802.1X Example



The following example shows how to configure the example shown in Figure 10-12.

1 Configure the RADIUS server IP address and shared secret (*secret*).

```
console#configure
console(config)#radius-server host 10.10.10.10
console(Config-radius)#exit
console(config)#radius-server key secret
console(config)#exit
```

2 Enable 802.1X port-based access control on the switch.

```
console(config)#dot1x system-auth-control
```

3 Configure ports 9 and 24 to be in the Authorized state, which allows the devices to connect to these ports to access the switch services without authentication.

```
console(config)#interface range Gi1/0/9,Gi1/0/24
console(config-if)#dot1x port-control force-authorized
console(config-if)#exit
```

4 Configure Port 7 to require MAC-based authentication with MAB.

```
console(config)#interface gi1/0/7
console(config-if-Gi1/0/7)#dot1x port-control mac-based
console(config-if-Gi1/0/7)#dot1x mac-auth-bypass
```

5 Set the port to an 802.1Q VLAN. The port must be in general mode in order to enable MAC-based 802.1X authentication.

```
console(config-if-Gi1/0/7)#switchport mode general
console(config-if-Gi1/0/7)#exit
```

6 Enable MAC-based authentication on port 8 and limit the number of devices that can authenticate on that port to 2.

```
console(config)#interface gi1/0/8
console(config-if-Gi1/0/8)#dot1x port-control mac-based
console(config-if-Gi1/0/8)#dot1x max-users 2
```

7 Set Port 8 to switchport mode general. The port must be in general mode in order to enable MAC-based 802.1X authentication.

```
console(config-if-Gi1/0/8)#switchport mode general
console(config-if-Gi1/0/8)#exit
console(config)#exit
```

8 View the client connection status.

When the clients on Ports 1, 3, and 7(supplicants), attempt to communicate via the switch, the switch challenges the supplicants for 802.1X credentials. The switch encrypts the provided information and transmits it to the RADIUS server. If the RADIUS server grants access, the system sets the 802.1X port state of the interface to authorized and the supplicants are able to access network resources.

```
console#show dot1x clients all

Interface...................................... Gi1/0/1
User Name...................................... aoversmit
Supp MAC Address............................... 0012.1753.031A
Session Time................................... 756
Filter Id......................................
VLAN Assigned.................................. 1 (Default)

Interface...................................... Gi1/0/3
User Name...................................... dflint
Supp MAC Address............................... 0004.5A55.EFAD
Session Time................................... 826
Filter Id......................................
```

```
VLAN Assigned................................. 1 (Default)

Interface..................................... Gi1/0/7
User Name..................................... 0006.6B33.06BA
Supp MAC Address.............................. 0006.6B33.06BA
Session Time.................................. 826
Filter Id.....................................
VLAN Assigned................................. 1 (Default)
```

9 View a summary of the port status.

```
console#show dot1x
Administrative Mode............... Enabled

Port     Admin            Oper         Reauth   Reauth
         Mode             Mode         Control  Period
-------  ---------------- ------------ -------- ----------
Gi1/0/1  auto             Authorized   FALSE    3600
Gi1/0/2  auto             N/A          FALSE    3600
Gi1/0/3  auto             Authorized   FALSE    3600
Gi1/0/4  auto             N/A          FALSE    3600
Gi1/0/5  auto             N/A          FALSE    3600
Gi1/0/6  auto             N/A          FALSE    3600
Gi1/0/7  mac-based        Authorized   FALSE    3600
Gi1/0/8  mac-based        N/A          FALSE    3600
Gi1/0/9  force-authorized Authorized   FALSE    3600
Gi1/0/10 force-authorized Authorized   FALSE    3600
Gi1/0/11 auto             N/A          FALSE    3600
```

10 View 802.1X information about Port 8.

```
console#show dot1x interface Gi1/0/8

Administrative Mode............... Enabled
Dynamic VLAN Creation Mode........ Enabled
Monitor Mode...................... Disabled

Port     Admin              Oper         Reauth    Reauth
         Mode               Mode         Control   Period
-------  ---------------- ------------  --------  ----------
Gi1/0/8  mac-based          Authorized   FALSE     3600

Quiet Period.................................. 60
Transmit Period............................... 30
Maximum Requests.............................. 2
Max Users..................................... 2
VLAN Assigned................................. 1 (Default)
Supplicant Timeout............................ 30
Guest-vlan Timeout............................ 90
Server Timeout (secs)......................... 30
MAB mode (configured)......................... Disabled
MAB mode (operational)........................ Disabled
```

Controlling Authentication-Based VLAN Assignment

The network in this example uses three VLANs to control access to network resources. When a client connects to the network, it is assigned to a particular VLAN based on one of the following events:

- It attempts to contact the 802.1X server and is authenticated.
- It attempts to contact the 802.1X server and fails to authenticate.
- It does not attempt to contact the 802.1X server.

The following table describes the three VLANs:

VLAN ID	VLAN Name	VLAN Purpose
100	Authorized	Data from authorized clients
200	Unauthorized	Data traffic from clients that fail the authentication with the RADIUS server
300	Guest	Data traffic from clients that do not attempt to authenticate with the RADIUS server

NOTE: Dynamic VLAN creation applies only to authorized ports. The VLANs for unauthorized and guest users must be configured on the switch and cannot be dynamically created based on RADIUS-based VLAN assignment.

NOTE: RADIUS VLAN assignment is supported on general mode ports only if MAC-based authentication is enabled. VLAN assignment is supported for all port modes other than trunk mode if auto assignment is enabled (**dot1x port-control auto**).

The commands in this example show how to configure the switch to control VLAN assignment for the example network. This example also contains commands to configure the uplink, or trunk, port (a port connected to a router or the internal network), and to configure the downlink, or access, ports (ports connected to one or more hosts). Ports 1–23 are downstream ports. Port 24 is an uplink port. An external RADIUS server handles the VLAN assignment.

NOTE: The configuration to control the VLAN assignment for authorized users is done on the external RADIUS server.

To configure the switch:

1 Create the VLANs and configure the VLAN names.

```
console(config)#vlan 100
console(config-vlan100)#name Authorized
console(config-vlan100)#exit

console(config)#vlan 200
console(config-vlan200)#name Unauthorized
console(config-vlan200)#exit

console(config)#vlan 300
console(config-vlan300)#name Guest
console(config-vlan300)#exit
```

2 Configure information about the external RADIUS server the switch uses to authenticate clients. The RADIUS server IP address is 10.10.10.10, and the shared secret is qwerty123.

```
console(config)#radius-server key qwerty123
console(config)#radius-server host 10.10.10.10
console(Config-auth-radius)#exit
```

3 Enable 802.1X on the switch.

```
console(config)#dot1x system-auth-control
```

4 Create a default authentication login list and use the RADIUS server for port-based authentication for connected clients.

```
console(config)#aaa authentication dot1x default radius
```

5 Allow the switch to accept VLAN assignments by the RADIUS server.

```
console(config)#aaa authorization network default radius
```

6 Enter interface configuration mode for the downlink ports.

```
console(config)#interface range Gi1/0/1-23
```

7 Set the downlink ports to the access mode because each downlink port connects to a single host that belongs to a single VLAN. Set the port control mode to auto (default) to allow VLAN assignment from the RADIUS server.

```
console(config-if)#switchport mode access
console(config-if)#dot1x port-control auto
```

8 Enable periodic reauthentication of the client on the ports and set the number of seconds to wait between reauthentication attempts to 300 seconds. Reauthentication is enabled to increase security. If the client

information is removed from the RADIUS server after it has been authenticated, the client will be denied access when it attempts to reauthenticate.

```
console(config-if)#dot1x reauthentication
console(config-if)#dot1x timeout re-authperiod 300
```

9 Set the unauthenticated VLAN on the ports to VLAN 200 so that any client that connects to one of the ports and fails the 802.1X authentication is placed in VLAN 200.

```
console(config-if)#dot1x unauth-vlan 200
```

10 Set the guest VLAN on the ports to VLAN 300. This command automatically enables the Guest VLAN Mode on the downlink ports. Any client that connects to the port and does not attempt to authenticate is placed on the guest VLAN.

```
console(config-if)#dot1x guest-vlan 300
console(config-if)#exit
```

11 Enter Interface Configuration mode for port 24, the uplink (trunk) port.

```
console(config)#interface Gi1/0/24
```

12 Disable 802.1X authentication on the interface. This causes the port to transition to the authorized state without any authentication exchange required. This port does not connect to any end-users, so there is no need for 802.1X-based authentication.

```
console(config-if-Gi1/0/24)#dot1x port-control force-
authorized
```

13 Set the uplink port to trunk mode so that it accepts tagged traffic and transmits it to the connected device (another switch or router).

```
console(config-if-Gi1/0/24)#switchport mode trunk
```

Allowing Dynamic Creation of RADIUS-Assigned VLANs

The network in this example uses a RADIUS server to provide VLAN assignments to host that connect to the switch. In this example, the VLANs are not configured on the switch. Instead, the switch is configured to allow the dynamic creation of VLANs when a RADIUS-assigned VLAN does not already exist on the switch.

In this example, Ports 1–23 are configured as downlink, or access, ports, and Port 24 is the trunk port. As a trunk port, Port 24 is automatically added as a member to all VLANs that are statically or dynamically configured on the switch. However, the network administrator in this example has determined that traffic in VLANs 1000–2000 should not be forwarded on the trunk port, even if the RADIUS server assigns a connected host to a VLAN in this range, and the switch dynamically creates the VLAN.

NOTE: The configuration to control the VLAN assignment for hosts is done on the external RADIUS server.

To configure the switch:

1 Configure information about the external RADIUS server the switch uses to authenticate clients. The RADIUS server IP address is 10.10.10.10, and the shared secret is qwerty123.

```
console(config)#radius-server key qwerty123
console(config)#radius-server host 10.10.10.10
console(Config-auth-radius)#exit
```

2 Enable 802.1X on the switch.

```
console(config)#dot1x system-auth-control
```

3 Create a default authentication login list and use the RADIUS server for port-based authentication for connected clients.

```
console(config)#aaa authentication dot1x default radius
```

4 Allow the switch to accept VLAN assignments by the RADIUS server.

```
console(config)#aaa authorization network default radius
```

5 Allow the switch to dynamically create VLANs when a RADIUS-assigned VLAN does not exist on the switch.

```
console(config)#dot1x dynamic-vlan enable
```

6 Enter interface configuration mode for the downlink ports.

```
console(config)#interface range Gi1/0/1-23
```

7 Set the downlink ports to the access mode because each downlink port connects to a single host that belongs to a single VLAN. Set the port-control mode to auto (the default) to allow assignment of the dynamically created VLANs to the host connected port.

```
console(config-if)#switchport mode access
console(config-if)#dot1x port-control auto
console(config-if)#exit
```

8 Enter Interface Configuration mode for port 24, the uplink (trunk) port.

```
console(config)#interface Gi1/0/24
```

9 Disable 802.1X authentication on the interface. This causes the port to transition to the authorized state without any authentication exchange required. This port does not connect to any end-users, so there is no need for 802.1X-based authentication.

```
console(config-if-Gi1/0/24)#dot1x port-control force-
authorized
```

10 Set the uplink port to trunk mode so that it accepts tagged traffic and transmits it to the connected device (another switch or router). The trunk port will automatically become a member of any dynamically created VLANs unless configured to exclude them.

```
console(config-if-Gi1/0/24)#switchport mode trunk
```

11 Forbid the trunk from forwarding traffic that has VLAN tags for any VLAN from 1000–2000, inclusive.

```
console(config-if-Gi1/0/24)#switchport trunk allowed vlan
remove 1000-2000
console(config-if-Gi1/0/24)#exit
```

Configuring Authentication Server DiffServ Policy Assignments

To enable DiffServ policy assignment by an external server, the following conditions must be true:

- The port that the host is connected to must be enabled for MAC-based port access control by using the following command in Interface Config mode:

  ```
  dot1x port-control mac-based
  ```

- The RADIUS or 802.1X server must specify the name of the policy to assign.

  For example, if the DiffServ policy to assign is named internet_access, include the following attribute in the RADIUS server configuration:

  Filter-id (11) = “internet_access”

- The DiffServ policy specified in the attribute must already be configured on the switch, and the policy names must be identical.

  For information about configuring a DiffServ policy, see "DiffServ Configuration Examples " on page 1467. The example "Providing Subnets

Equal Access to External Network " on page 1467, describes how to configure a policy named internet_access.

If you use an authentication server to assign DiffServ policies to an authenticated user, note the following guidelines:

- If the policy specified within the server Filter-id attribute does not exist on the switch, authentication will fail.
- Do not delete policies used as the Filter-id by the RADIUS server while 802.1X is enabled.
- Do not use the DiffServ **service-policy** command to apply the filter to an interface if you configure the RADIUS server or 802.1X authenticator to assign the DiffServ filter.

In the following example, Company XYZ uses IEEE 802.1X to authenticate all users. Contractors and temporary employees at Company XYZ are not permitted to have access to SSH ports, and data rates for Web traffic is limited. When a contractor is authenticated by the RADIUS server, the server assigns a DiffServ policy to control the traffic restrictions.

The network administrator configures two DiffServ classes: *cl-ssh* and *cl-http*. The class *cl-ssh* matches all incoming SSH packets. The class *cl-http* matches all incoming HTTP packets. Then, the administrator configures a traffic policy called *con-pol* and adds the *cl-ssh* and *cl-http*. The policy is configured so that SSH packets are to be dropped, and HTTP data rates are limited to 1 MB with a burst size of 64 Kbps. HTTP traffic that exceeds the limit is dropped. The host ports, ports 1–23, are configured to use MAC-based dot1x authentication to allow the DiffServ policy to be applied. Finally, the administrator configures the RADIUS server with the attribute Filter-id (11) = "con-pol" (steps not shown).

To configure the switch:

1 Configure the DiffServ traffic class that matches SSH traffic.

```
console#configure
console(config)#class-map match-all cl-ssh
console(config-classmap)#match dstl4port 22
console(config-classmap)#exit
```

2 Configure the DiffServ traffic class that matches HTTP traffic.

```
console(config)#class-map match-all cl-http
console(config-classmap)#match dstl4port 80
console(config-classmap)#exit
```

3 Configure the DiffServ policy.

```
console(config)#policy-map con-pol in
console(config-policy-map)#class cl-ssh
console(config-policy-classmap)#drop
console(config-policy-classmap)#exit
console(config-policy-map)#class cl-http
console(config-policy-classmap)#police-simple 1000000 64
conform-action transmit violate-action drop
console(config-policy-classmap)#exit
console(config-policy-map)#exit
```

4 Enable DiffServ on the switch. (Optional as diffserv is enabled by default.)

```
console(config)#diffserv
```

5 Configure information about the external RADIUS server the switch uses to authenticate clients. The RADIUS server IP address is 10.10.10.10, and the shared secret is qwerty123.

```
console(config)#radius-server key qwerty123
console(config)#radius-server host 10.10.10.10
console(Config-auth-radius)#exit
```

6 Enable 802.1X on the switch.

```
console(config)#dot1x system-auth-control
```

7 Create a default authentication login list and use the RADIUS server for port-based authentication for connected clients.

```
console(config)#aaa authentication dot1x default radius
```

8 Enter Interface Configuration mode for ports 1–23 and enable MAC-based authentication.

```
console(config)#interface range Gi1/0/1-23
console(config-if)#dot1x port-control mac-based
```

9 Set the ports to an 802.1Q VLAN. The ports must be in general mode to enable MAC-based 802.1X authentication. Enable the policy on the ports.

```
console(config-if)#switchport mode general
console(config-if)#service-policy in con-pol
console(config-if)#exit
console(config)#exit
```

Captive Portal

This section describes how to configure the Captive Portal feature.

The topics covered in this section include:

- Captive Portal Overview
- Default Captive Portal Behavior and Settings
- Configuring Captive Portal (Web)
- Configuring Captive Portal (CLI)
- IEEE 802.1X Configuration Examples

Captive Portal Overview

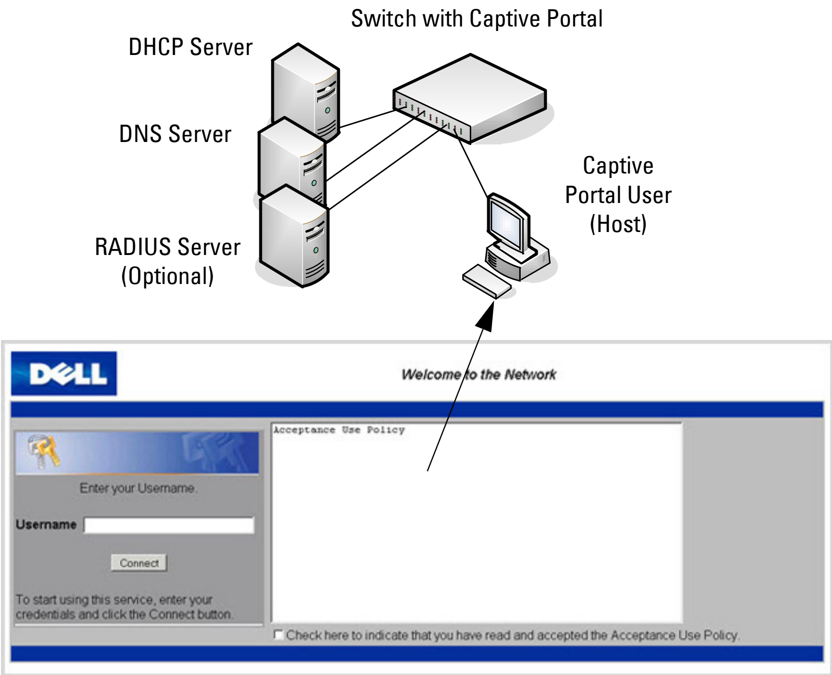
A Captive Portal (CP) helps manage or restrict network access. CPs are often used in locations that provide wired Internet access to customers, such as business centers and hotels. For example, a hotel might provide an Ethernet port in each room so that guests can connect to the Internet during their stay. The hotel might charge for Internet use, or the hotel might allow guests to connect only after they indicate that they have read and agree to the acceptable use policy.

What Does Captive Portal Do?

The CP feature allows you to require a user to enter login information on a custom web page before gaining access to the network. When the user connects to the port and opens a browser, the user is presented with a welcome screen. To gain network access, the user must enter a username (for guest access) or a username and password (for authenticated access) and accept the terms of use. The network administrator can also configure the CP feature to redirect the user to another web page after successful authentication, for example a company home page.

CP is supported in IPv4 networks only.

Figure 10-13. Connecting to the Captive Portal



Default Captive Portal Welcome Screen (Displays in Captive Portal User's Browser)

The CP feature blocks hosts connected to the switch from most network access until user verification has been established. Access to 802.1X, DHCP, ARP, NetBIOS, and DNS services is allowed. The network administrator can configure CP verification to allow access for both guest and authenticated users. Authenticated users must be validated against a database of authorized CP users before access is granted. The database can be stored locally on the switch or on a RADIUS server.

Is Captive Portal Dependent on Any Other Feature?

If security procedures require RADIUS authentication, the administrator must configure the RADIUS server information on the switch (see "Using RADIUS " on page 241). The RADIUS administrator must also configure the RADIUS attributes for CP users on the RADIUS server. For information about the RADIUS attributes to configure, see Table 10-15.

For a list of RADIUS attributes that the switch supports, see "Which RADIUS Attributes Does the Switch Support? " on page 243.

To support redirection of user entered URLs from a web browser, a DNS server must be configured in the network. If routing is enabled on the switch, IP helper should be enabled to allow hosts to obtain an IP address via DHCP.

A DHCP server must be available if it is expected that hosts will obtain IP addresses dynamically. In addition, if routing is enabled, DHCP relay must be configured.

The only type of interface where CP can be enabled is a physical port. CP is not supported on multi-access VLANs or on LAGs.

A physical port's VLAN membership does not affect CP. A physical port enabled for CP can be a member of any VLAN or multiple VLANs, which can be switching or routing VLANs.

A port enabled for CP may be directly connected to a single client (e.g., an access switch), or the port may serve many clients (e.g., a port on an aggregation switch).

Port security and CP cannot both be enabled on the same interface.

If a physical port configured with CP is made a member of a LAG, CP is disabled on the port.

Dell Networking does not support configuring spanning tree on a CP port. BPDUs received on a port enabled for CP will not receive their normal prioritization.

CP can coexist on an interface with DHCP snooping and Dynamic ARP Inspection (DAI).

The administrator can configure the switch to send SNMP trap messages to any enabled SNMP Trap Receivers for several CP events, such as when a CP user has an authentication failure or when a CP user successfully connects to the network. If traps are enabled, the switch also writes a message to the trap log when the event occurs. To enable the CP traps, see "Configuring SNMP Notifications (Traps and Informs) " on page 461.

What Factors Should Be Considered When Designing and Configuring a Captive Portal?

Before enabling the CP feature, decide what type (or types) of authentication will be supported. Since Dell Networking N-Series switches support up to 10 different CP instances, it is possible to configure one CP that requires a username and password and another that only requires the username. For each CP, the administrator can customize the welcome screen, including the colors and logo.

If network policy requires authentication, consider the number of users that must exist in the user database. The local user database supports up to 128 users. If there is a need to support more than 128 authenticated users, use a remote RADIUS server for authentication.

The administrator can specify whether the CP uses HTTP or HTTPS as the protocol during the user verification process. HTTP does not use encryption during verification, and HTTPS uses the Secure Sockets Layer (SSL), which requires a certificate to provide encryption. The certificate is presented to the user at connection time.

If the authenticating user requires DNS or DHCP services, these will need to be configured in the network and the switch will need to relay DHCP packets.

The initial Web page that a user sees when he or she connects to the CP can be customized. The logo, color schemes, welcome messages, and all text on the page can be customized, including the field and button labels. The welcome page the user sees after a successful verification or authentication can also be customized.

Figure 10-14. Customized Captive Portal Welcome Screen

How Does Captive Portal Work?

When a port is enabled for CP, all the traffic coming onto the port from the unverified clients is dropped except for the ARP, DHCP, NetBIOS, and DNS packets. These packets are forwarded by the switch so that the unverified clients can get an IP address and are able to resolve host or domain names. If an unverified client opens a web browser and tries to connect to the network, CP redirects all the HTTP/HTTPS traffic from the unverified client to the authenticating server on the switch. If the network administrator has configured an additional web server port, packets with this destination TCP port number are also forwarded to the authenticating server. A CP web page is sent back to the unverified client. If the verification mode for the CP associated with the port is Guest, the client can be verified without providing authentication information. If the verification mode is Local or RADIUS, the client must provide credentials that are compared against the information in the Local or RADIUS client database. After the user successfully provides the required information, the CP feature grants access to the network.

What Captive Portal Pages Can Be Customized?

The following three CP pages can be customized:

- Authentication Page —This page displays when a client attempts to connect to the network. The images, text, and colors that display on this page can be customized.

- Logout Page — If the user logout mode is enabled, this page displays in a pop-up window after the user successfully authenticates. This window contains the logout button.
- Logout Success Page — If the user logout mode is enabled, this page displays after a user clicks the logout button and successfully deauthenticates.

Understanding User Logout Mode

The User Logout Mode feature allows a user who successfully authenticates to the network through the CP to explicitly deauthenticate from the network. When User Logout Mode is disabled or the user does not specifically request logout, the connection status will remain authenticated until the CP deauthenticates the user based on the configured session timeout value. In order for the user logout feature to function properly, the client browser must have JavaScript enabled an must allow popup windows.

Localizing Captive Portal Pages

The CP localization feature allows you to create up to three language-specific web pages for each CP as long as all pages use the same verification type; either guest or authorized user web pages. This allows you to create pages in a variety of languages to accommodate a diverse group of users.

To customize the pages that the user sees, click the language tab. By default, the English tab is available. The settings for the **Authentication Page** display.

Captive Portal IP Address Selection

CP automatically associates with one of the IP addresses assigned to the switch. The automatic IP address selection algorithm is outlined below:

1. On switching-only devices or when routing is disabled, CP uses the out-of-band interface IP address, if available.
2. If routing is enabled, CP uses a loopback interface if one is defined, and a routing interface as the second choice.
3. If routing is enabled and no active routing interface is available, the CP goes down.
4. If the CP IP address changes due to administrator action or due to an interface going down, then the CP is automatically disabled and re-enabled. All active sessions are dropped.

Captive Portal and DNS

CP allows unauthenticated users access to DNS services on TCP and UDP destination port 53. CP inspects all DNS traffic to ensure that it conforms with the DNS protocol (RFC 1035/1996). CP checks the format of DNS messages and discards packets that do not conform to the minimum standards. Specifically, CP performs the following checks on a DNS packet:

- The packet must have a full-size header and at least one question field
- The packet must have a valid DNS response code
- The first question field must not exceed 63 octets in length, nor must the length field be greater than 63
- The first question class field must be valid.

Captive Portal Troubleshooting

The following table explains the status values for CP authentication sessions and the resulting actions taken, if any. CP global status, interface status, and session status are available in the user interfaces.

Table 10-13. Captive Portal Status Values

Status Value	Description	Browser Action
Default	Initial request from the client.	Used to detect initial request.
Serve	Default serve.	Used when serving the initial connection page.
Validate	Actual validation request.	Indicates that the user has submitted credentials and requests authentication.
WIP	Indicates that validation is in progress.	The validation page begins to poll the server until the status flag changes. The actual poll request is the same http(s) request used to “validate” as described above. While waiting between polls, the browser displays an “authorization in process” message.

Table 10-13. Captive Portal Status Values (Continued)

Status Value	Description	Browser Action
RADIUS_WIP	Indicates that RADIUS validation is in progress.	The browser action is the same as for the WIP status.
Success	Indicates that authentication is a success.	Displays either the customized welcome page or an external URL.
Denied	Indicates that the user has failed to enter credentials that match the expected configuration.	The default serve page is resubmitted and includes the appropriate failure message.
Resource	Indicates that the system has rejected authentication due to system resource limitations or session timeout.	The default serve page is resubmitted and includes the appropriate failure message.
No Accept	Indicates that the user did not accept the acceptance use policy.	The default serve page is resubmitted and includes the appropriate failure message.
Timeout	Indicates that the authentication transaction took too long. This could be due to user input time, or a timeout due to the overall transaction.	The default serve page is resubmitted and includes the appropriate failure message.

Default Captive Portal Behavior and Settings

CP is disabled by default. If you enable CP, no interfaces are associated with the default CP. After you associate an interface with the CP and globally enable the CP feature, a user who connects to the switch through that interface is presented with the CP Welcome screen shown in Figure 10-15.

Figure 10-15. Default Captive Portal Welcome Screen

The user types a name in the Username field, selects the Acceptance Use Policy check box, and clicks **Connect** to gain network access. By default, the user does not need to be defined in a database or enter a password to access the network because the default verification mode is Guest. Note that duplicate Username entries can exist in this mode because the client IP and MAC addresses are obtained for identification.

Table 10-14 shows the default values for the CP feature.

Table 10-14. Default Captive Portal Values

Feature	Value
Global Captive Portal Operational Status	Disabled
Additional HTTP or HTTPS Ports	Disabled CP can be configured to use an additional HTTP and/or HTTPS port (in support of Proxy networks).
Authentication Timeout	300 seconds

Table 10-14. Default Captive Portal Values

Feature	Value
Configured Captive Portals	1
Captive Portal Name	Default
Protocol Mode	HTTP
Verification Mode	Guest
URL Redirect Mode	Off
User Group	1-Default
Session Timeout	86400 seconds
Local Users	None configured
Interface associations	None
Interface status	Not blocked
	If the CP is blocked, users cannot gain access to the network through the CP. Use this function to temporarily protect the network during unexpected events, such as denial of service attacks.
Supported Captive Portal users	1024
Supported local users	128
Supported Captive Portals	10

Configuring Captive Portal (Web)

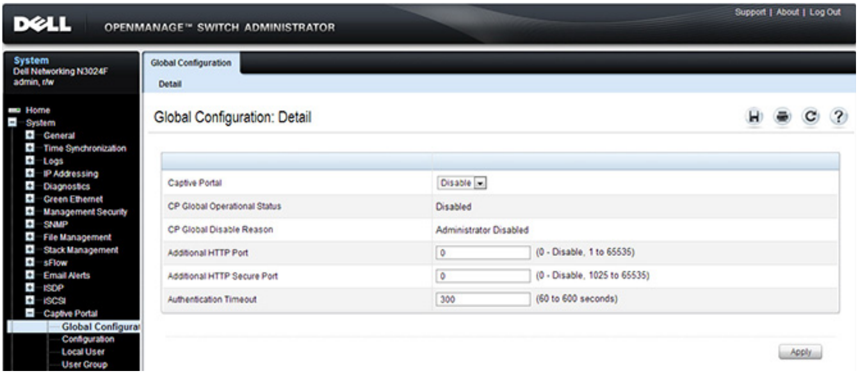
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring CP settings on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click at the top of the page.

Captive Portal Global Configuration

Use the **Captive Portal Global Configuration** page to control the administrative state of the CP feature and configure global settings that affect all CPs configured on the switch.

To display the **Captive Portal Global Configuration** page, click **System → Captive Portal → Global Configuration.**

Figure 10-16. Captive Portal Global Configuration



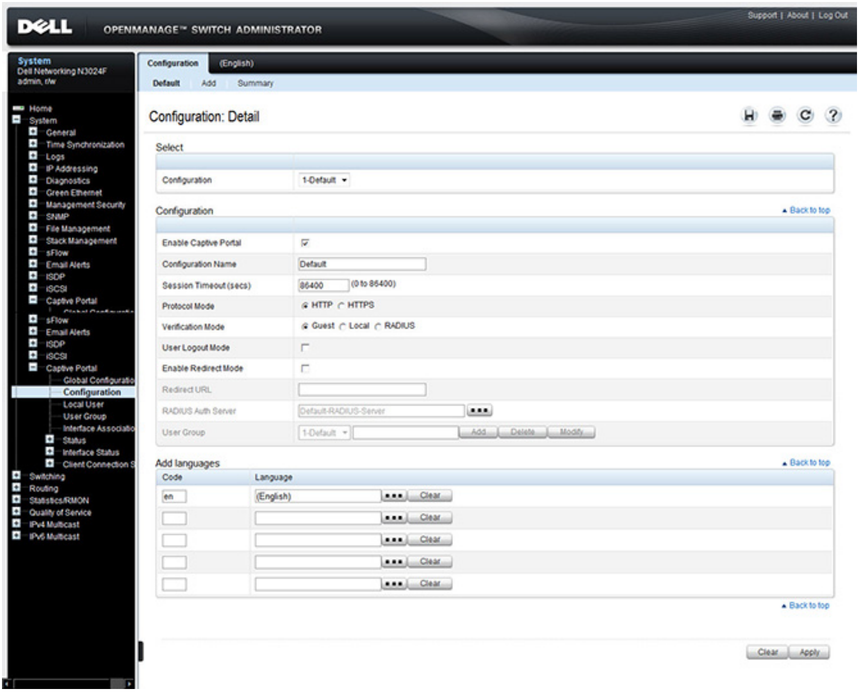
Captive Portal Configuration

Use the **Captive Portal Configuration** page to view summary information about CPs on the system, add a CP, and configure existing CPs.

The switch supports 10 CP configurations. CP configuration 1 is created by default and cannot be deleted. Each CP configuration can have unique guest or group access modes and a customized acceptance use policy that displays when the client connects.

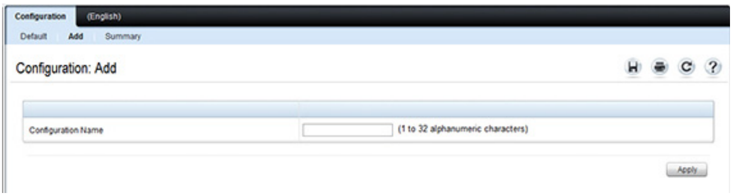
To display the **Captive Portal Configuration** page, click **System → Captive Portal → Configuration.**

Figure 10-17. Captive Portal Configuration



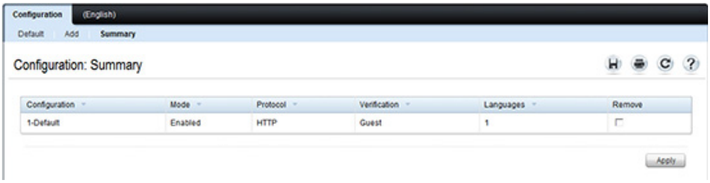
From the **Captive Portal Configuration** page, click **Add** to create a new CP instance.

Figure 10-18. Add Captive Portal Configuration



From the **Captive Portal Configuration** page, click **Summary** to view summary information about the CP instances configured on the switch.

Figure 10-19. Captive Portal Summary



Customizing a Captive Portal

The procedures in this section customize the pages that the user sees when he or she attempts to connect to (and log off of) a network through the CP. These procedures configure the English version of the Default Captive Portal.

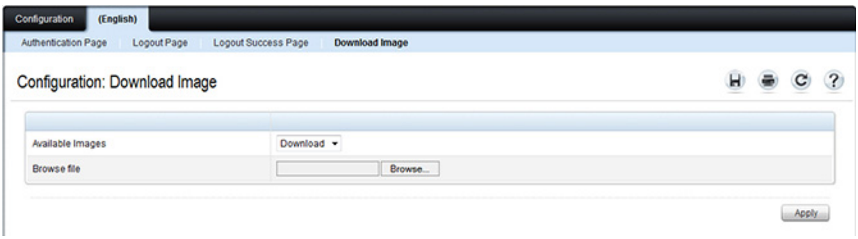
To configure the switch:

1 From the **Captive Portal Configuration** page click the **(English)** tab. The settings for the **Authentication Page** display, and the links to the CP customization appear.

2 Click **Download Image** to download one or more custom images to the switch. A downloaded custom image can be used for the branding logo (default: Dell logo) on the Authentication Page and Logout Success page, for the account image (default: blue banner with keys) on the Authentication Page, and for the background image (default: blank) on the Logout Success Page.

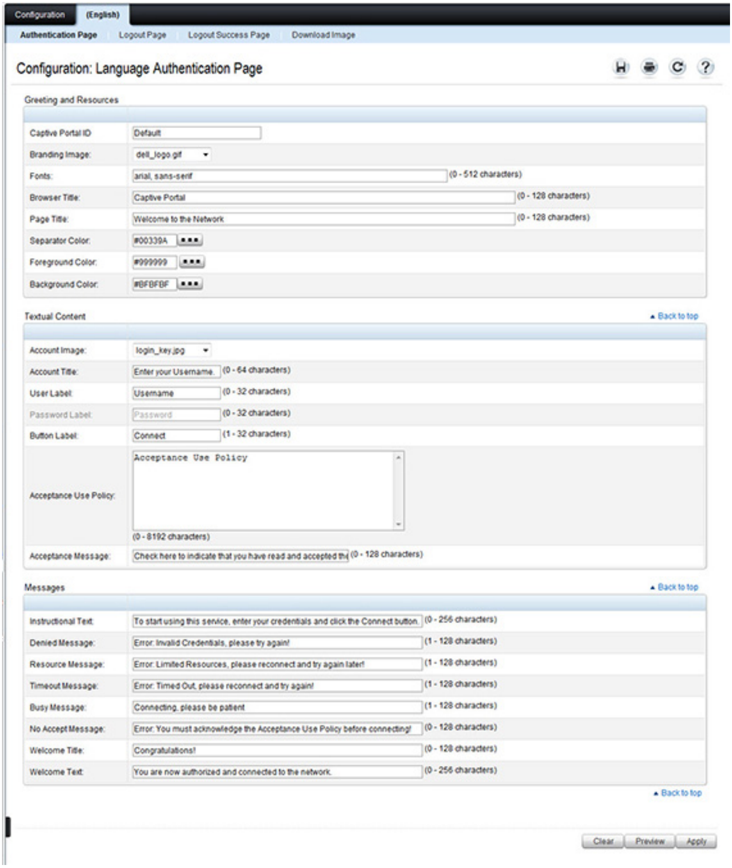
> **NOTE:** The image to download must be accessible from your local system. The image should be 5 KB max, 200x200 pixels, GIF or JPG format.

Figure 10-20. Captive Portal Download Image Page



3. Make sure Download is selected in the **Available Images** menu, and click **Browse**.
4. Browse to the directory where the image to be downloaded is located and select the image.
5. Click **Apply** to download the selected file to the switch.
6. To customize the Authentication Page, which is the page that a user sees upon attempting to connect to the network, click the **Authentication Page** link.

Figure 10-21. Captive Portal Authentication Page



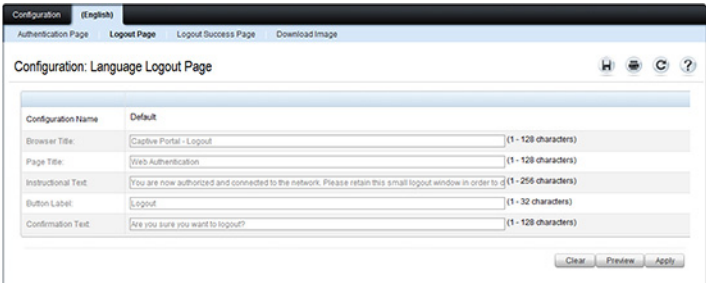
7 Select the branding image to use and customize other page components such as the font for all text the page displays, the page title, and the acceptance use policy.

8 Click **Apply** to save the settings to the running configuration or click **Preview** to view what the user will see. To return to the default views, click **Clear**.

9 Click the **Logout Page** link to configure the page that contains the logout window.

NOTE: The Logout Page settings can be configured only if the User Logout Mode is selected on the **Configuration** page. The User Logout Mode allows an authenticated client to deauthenticate from the network.

Figure 10-22. Captive Portal Logout Page

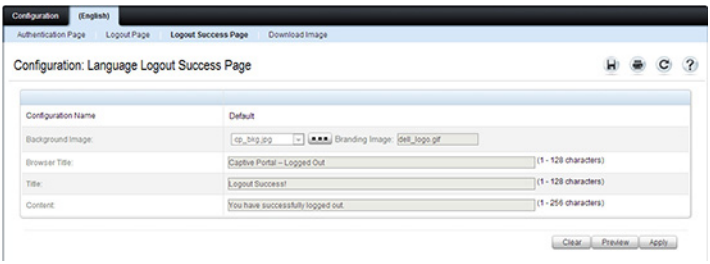


10 Customize the look and feel of the Logout Page, such as the page title and logout instructions.

11 Click **Apply** to save the settings to the running configuration or click **Preview** to view what the user will see. To return to the default views, click **Clear**.

12 Click the **Logout Success Page** link to configure the page that contains the logout window. A user is required to logout only if the User Logout Mode is selected on the **Configuration** page.

Figure 10-23. Captive Portal Logout Success Page



13 Customize the look and feel of the Logout Page, such as the background image and successful logout message.

14 Click **Apply** to save the settings to the running configuration or click **Preview** to view what the user will see. To return to the default views, click **Clear**.

Local User

A portal can be configured to accommodate guest users and authorized users. Guest users do not have assigned user names and passwords. Authorized users provide a valid user name and password that must first be validated against a local database or RADIUS server. Authorized users can gain network access once the switch confirms the user's credentials.

By default, each CP instance contains the default group. The default group can be renamed, or a different group can be created and assigned to each CP instance. A CP instance can be associated to one user group only. A user, however, can be assigned to multiple groups.

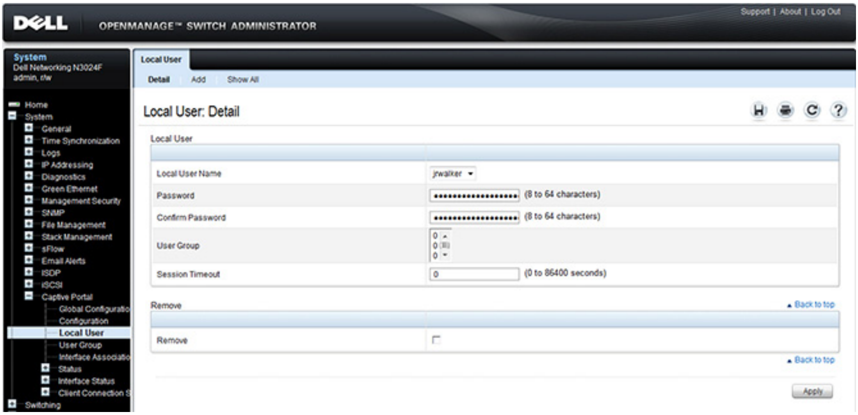
The **Local User** page allows you to add authorized users to the local database, which can contain up to 128 user entries. Users can be added to and deleted from the local database using the **Local User** page.

To display the **Local User** page, click **System → Captive Portal → Local User.**

Figure 10-24 shows the **Local User** page after a user has been added. If no users have been added to the switch, many of the fields do not display on the screen.

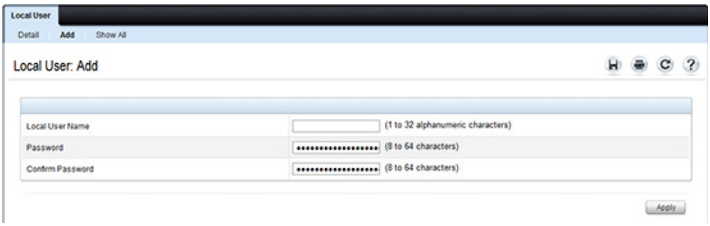
NOTE: Multiple user groups can be selected by holding the CTRL key down while clicking the desired groups.

Figure 10-24. Local User Configuration



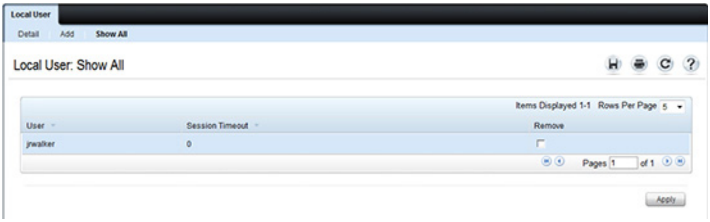
From the **Local User** page, click **Add** to add a new user to the local database.

Figure 10-25. Add Local User



From the **Local User** page, click **Show All** to view summary information about the local users configured in the local database.

Figure 10-26. Captive Portal Local User Summary



To delete a configured user from the database, select the Remove check box associated with the user and click **Apply**.

Configuring Users in a Remote RADIUS Server

A remote RADIUS server client authorization can be used. All users must be added to the RADIUS server. The local database does not share any information with the remote RADIUS database.

Table 10-15 indicates the RADIUS attributes you use to configure authorized CP clients. The table indicates both RADIUS attributes and vendor-specific attributes (VSA). VSAs are denoted in the Attribute column and are comma delimited (vendor ID, attribute ID).

Table 10-15. Captive Portal User RADIUS Attributes

Attribute	Number	Description	Range	Usage	Default
User-Name	1	User name to be authorized	1-32 characters	Required	None
User-Password	2	User password	8-64 characters	Required	None

Table 10-15. Captive Portal User RADIUS Attributes (Continued)

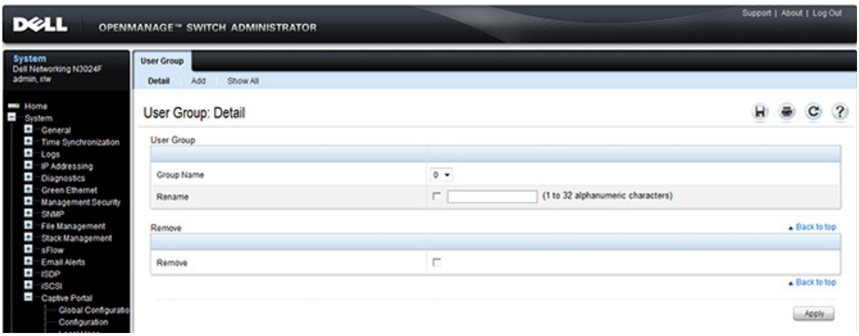
Attribute	Number	Description	Range	Usage	Default
Session-Timeout	27	Logout once session timeout is reached (seconds). If the attribute is 0 or not present then use the value configured for the CP.	Integer (seconds)	Optional	0
Dell-Captive-Portal-Groups	6231, 127	A comma-delimited list of group names that correspond to the configured CP instance configurations.	String	Optional	None. The default group is used if not defined here

User Group

Local Users can be assigned to User Groups. If the Verification Mode is Local or RADIUS, a User Group is assigned to a CP Configuration. All users who belong to the group are permitted to access the network through this portal. The User Group list is the same for all CP configurations on the switch.

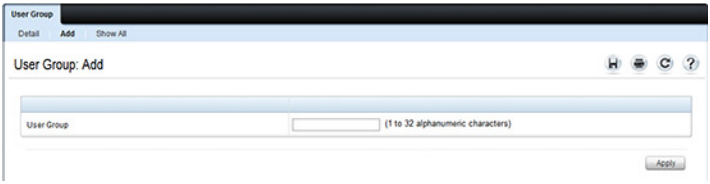
To display the **User Group** page, click **System → Captive Portal → User Group.**

Figure 10-27. User Group



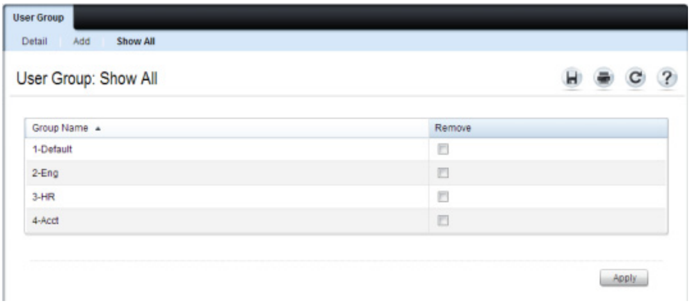
From the **User Group** page, click **Add** to configure a new user group.

Figure 10-28. Add User Group



From the **User Group** page, click **Show All** to view summary information about the user groups configured on the switch.

Figure 10-29. Captive Portal User Group Summary



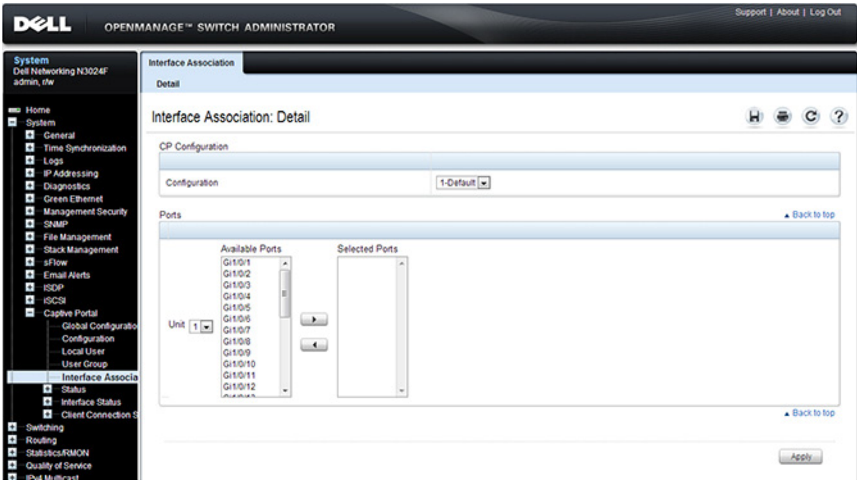
To delete a configured group, select the Remove check box associated with the group and click **Apply**.

Interface Association

Using the **Interface Association** page, a configured CP can be associated with specific interfaces. The CP feature only runs on the interfaces that you specify. A CP can have multiple interfaces associated with it, but an interface can be associated to only one CP at a time.

To display the **Interface Association** page, click **System** → **Captive Portal** → **Interface Association**.

Figure 10-30. Captive Portal Interface Association



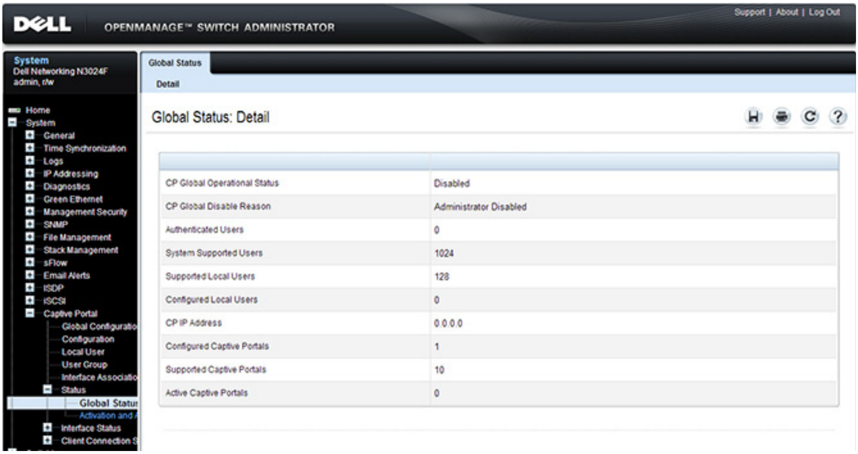
NOTE: When you associate an interface with a CP, the interface is disabled in the Interface List. Each interface can be associated with only one CP at a time.

Captive Portal Global Status

The **Captive Portal Global Status** page contains a variety of information about the CP feature, including information about the CP activity and interfaces.

To display the **Global Status** page, click **System → Captive Portal → Status Global Status.**

Figure 10-31. Captive Portal Global Status



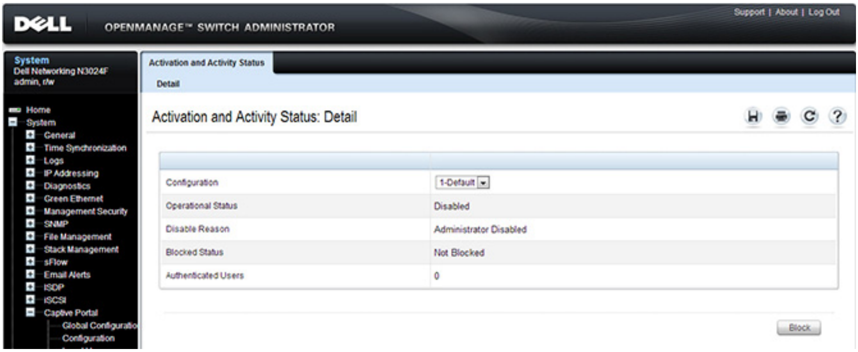
Captive Portal Activation and Activity Status

The **Captive Portal Activation and Activity Status** page provides information about each CP configured on the switch.

The **Captive Portal Activation and Activity Status** page has a drop-down menu that contains all CPs configured on the switch. When you select a CP, the activation and activity status for that portal displays.

To display the **Activation and Activity Status** page, click **System** → **Captive Portal** → **Status Activation and Activity Status.**

Figure 10-32. Captive Portal Activation and Activity Status



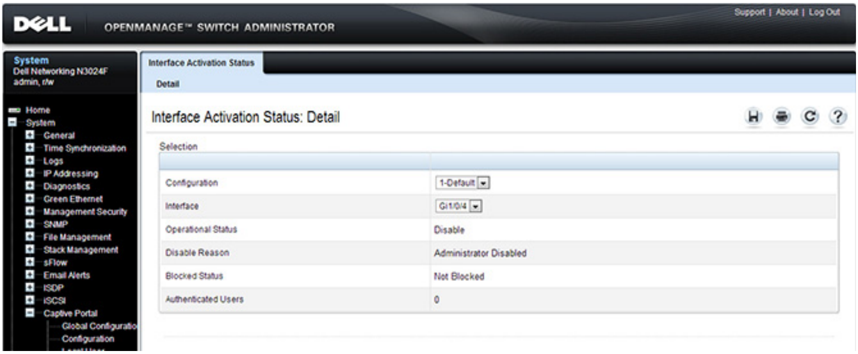
NOTE: Use the Block and Unblock buttons to control the blocked status. If the CP is blocked, users cannot gain access to the network through the CP. Use this function to temporarily protect the network during unexpected events, such as denial of service attacks.

Interface Activation Status

The **Interface Activation Status** page shows information for every interface assigned to a CP instance.

To display the **Interface Activation Status** page, click **System** → **Captive Portal** → **Interface Status Interface Activation Status.**

Figure 10-33. Interface Activation Status

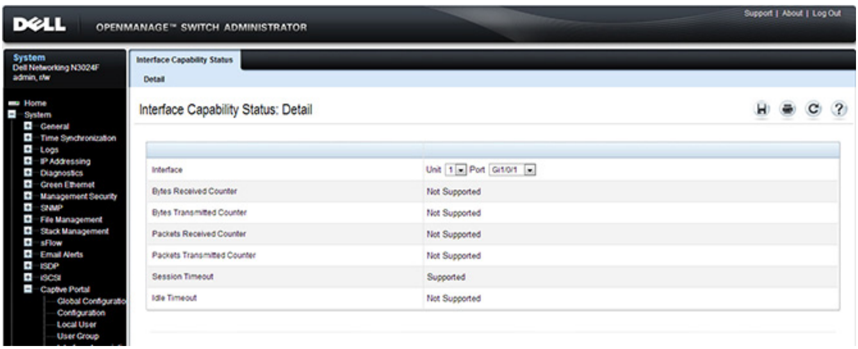


Interface Capability Status

The **Interface Capability Status** page contains information about interfaces that can have CPs associated with them. The page also contains status information for various capabilities. Specifically, this page indicates what services are provided through the CP to clients connected on this interface. The list of services is determined by the interface capabilities.

To display the **Interface Capability Status** page, click **System** → **Captive Portal** → **Interface Status Interface Capability Status.**

Figure 10-34. Interface Capability Status

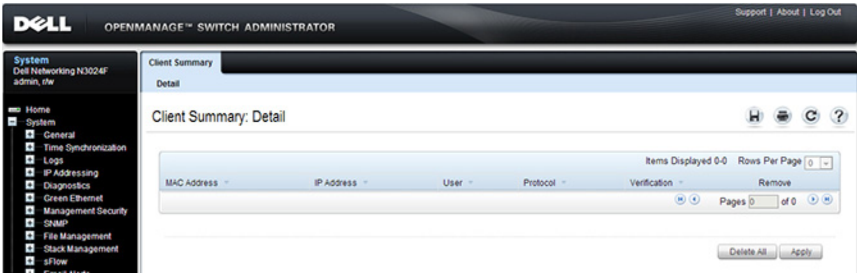


Client Summary

Use the **Client Summary** page to view summary information about all authenticated clients that are connected through the CP. From this page, the CP can be manually forced to disconnect one or more authenticated clients. The list of clients is sorted by client MAC address.

To display the **Client Summary** page, click **System** → **Captive Portal** → **Client Connection Status Client Summary.**

Figure 10-35. Client Summary



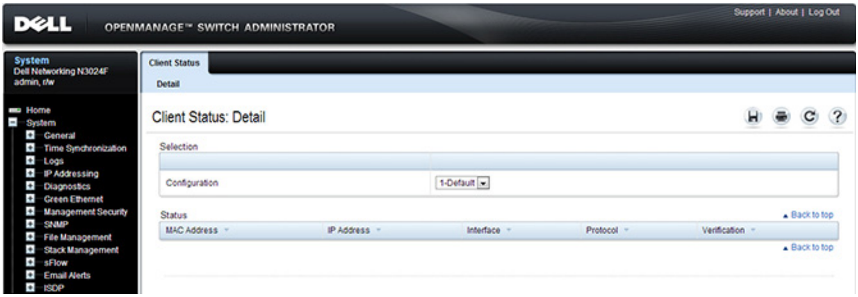
To force the CP to disconnect an authenticated client, select the **Remove** check box next to the client MAC address and click **Apply**. To disconnect all clients from all CPs, click **Delete All.**

Client Detail

The **Client Detail** page shows detailed information about each client connected to the network through a CP.

To display the **Client Detail** page, click **System** → **Captive Portal** → **Client Connection Status Client Detail.**

Figure 10-36. Client Detail

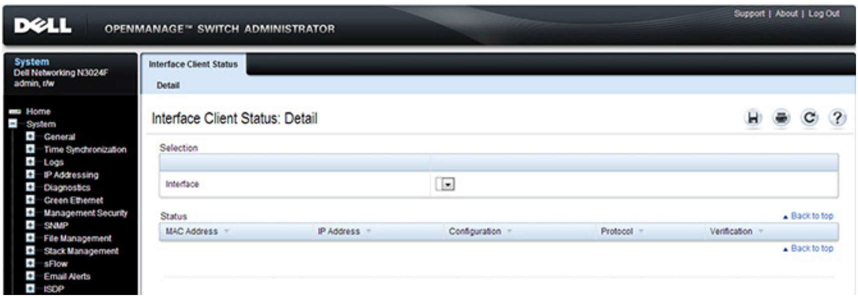


Captive Portal Interface Client Status

Use the **Interface Client Status** page to view clients that are authenticated to a specific interface.

To display the **Interface Client Status** page, click **System** → **Captive Portal Client Connection Status Interface Client Status.**

Figure 10-37. Interface - Client Status

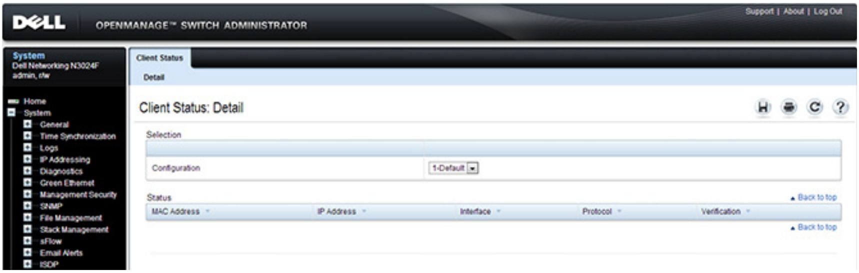


Captive Portal Client Status

Use the **Client Status** page to view clients that are authenticated to a specific CP configuration.

To display the **Client Status** page, click **System** → **Captive Portal Client Connection Status Client Status.**

Figure 10-38. Captive Portal - Client Status



Configuring Captive Portal (CLI)

This section provides information about the commands you use to create and configure CP settings. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Global Captive Portal Settings

Beginning in Privileged EXEC mode, use the following commands to configure global CP settings.

Command	Purpose
configure	Enter global configuration mode.
captive-portal	Enter Captive Portal mode.
http port *port-num*	(Optional) Configure an additional HTTP port for CP to monitor. Use this command on networks that use an HTTP proxy server. *port-num* — The port number to monitor (Range: 1–65535, excluding ports 80, 443, and the configured switch management port).
https port *port-num*	(Optional) Configure an additional HTTPS port for CP to monitor. Use this command on networks that use an HTTPS proxy server. *port-num* — The port number to monitor Range: 1–65535, excluding ports 80, 443, and the configured switch management port).
authentication timeout *timeout*	(Optional) Configure the number of seconds the user has to enter valid credentials into the verification page. If the user exceeds the configured timeout, the verification page needs to be served again in order for the client to gain access to the network. *timeout* — The authentication timeout (Range: 60–600 seconds).
enable	Globally enable the CP feature.

Command	Purpose
CTRL + Z	Exit to Privileged EXEC mode.
show captive-portal [**status**]	View the CP administrative and operational status. Use the **status** keyword to view additional global CP information and summary information about all configured CP instances.

Creating and Configuring a Captive Portal

Beginning in Privileged EXEC mode, use the following commands to create a CP instance and configure its settings.

Command	Purpose
configure	Enter global configuration mode.
captive-portal	Enter Captive Portal mode.
configuration *cp-id*	Enter the CP instance mode *cp-id* — The CP instance (Range: 1–10). The CP configuration identified by CP ID 1 is the default CP configuration.
name *string*	Add a name to the CP instance. *string* — CP configuration name (Range: 1–32 characters).
protocol {**http** | **https**}	Specify whether to use HTTP or HTTPs during the CP user verification process.
verification {**guest** | **local** | **radius**}	Specify how to process user credentials the user enters on the verification page. • guest — Allows access for unauthenticated users (users that do not have assigned user names and passwords). • local — Authenticates users against a local user database. • radius — Authenticates users against a remote RADIUS database.
radius-auth-server *name*	(Optional) Specify the name of the RADIUS server to use for RADIUS verification. Use the commands described in "Using RADIUS " on page 241 to configure RADIUS server settings for the switch. This command is not required if **local** or **guest** verification is configured.

Command	Purpose
user-logout	(Optional) Enable user logout mode to allow an authenticated client to deauthenticate from the network. If this option is clear or the user does not specifically request logout, the client connection status remains authenticated until the CP deauthenticates the user, for example by reaching the idle timeout or session timeout values.
redirect	(Optional) Enable the redirect mode for a CP configuration so that the user is redirected to a specific Web page after the verification or authentication process. When the redirect mode is not enabled, the user sees the CP welcome page after the verification or authentication process.
redirect-url *url*	(Optional) Specify the web page that the users sees after successful verification or authentication through the CP. *url* — The URL for redirection (Range: 1–512 characters).
group *group-number*	(For Local and RADIUS verification) Configure the group number associated with this CP configuration. By default, only the default group exists. To assign a different user group to the CP instance, you must first configure the group. *group-number* — The number of the group to associate with this configuration (Range: 1–10)
session-timeout *timeout*	(Optional) Enter the number of seconds to wait before terminating a session. A user is logged out once the session timeout is reached. The session timeout can be set for each user if the CP requires authentication. *timeout* — Session timeout. 0 indicates timeout not enforced (Range: 0–86400 seconds)
interface *interface*	Associate an interface with this CP. (The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**.
enable	Enable the CP instance.

Command	Purpose
block	(Optional) Block all traffic for a CP configuration. If the CP is blocked, users cannot gain access to the network through the CP. Use this function to temporarily protect the network during unexpected events, such as denial of service attacks.
CTRL + Z	Exit to Privileged EXEC mode.
show captive-portal configuration *cp-id* [**status** | **interface**]	View summary information about a CP instance. • *cp-id* — The CP instance (Range: 1–10). • **status** — View additional information about the CP instance. • interface — View information about the interface(s) associated with the specified CP.
show captive-portal interface configuration *cp-id* **status**	View information about the interfaces associated with the specified CP instance. *cp-id* — The CP instance (Range: 1–10).

NOTE: To return the default CP instance to its default values, use the **clear** command in the Captive Portal Instance mode. You must also use the **no interface** *interface* command to remove any associated interfaces from the instance.

Configuring Captive Portal Groups and Users

Beginning in Privileged EXEC mode, use the following commands to create a CP group. The default group can be used, or a new group can be created.

Command	Purpose
configure	Enter global configuration mode.
captive-portal	Enter Captive Portal mode.

Command	Purpose
user group *group-id* [**name** *name*]	Configure a group. Each CP that requires authentication has a group associated with it. Only the users who are members of that group can be authenticated if they connect to the CP. • *group-id* — Group ID (Range: 1–10). • *name* — Group name (Range: 1–32 characters).
user *user-id* **name** *name*	Create a new user for the local user authentication database. • *user-id*—User ID (Range: 1–128). • *name*—user name (Range: 1–32 characters).
user *user-id* **password** *password*	Configure the password for the specified user. • *user-id*—User ID (Range: 1–128). • *password*—User password (Range: 8–64 characters).
user *user-id* **group** *group-id*	Associate a group with a CP user. A user can be associated with more than one group. • *user-id* — User ID (Range: 1–128). • *group-id* — Group ID (Range: 1–10).
user *user-id* **session-timeout** *timeout*	Enter the number of seconds to wait before terminating a session for the specified user. The user is logged out once the session timeout is reached. • *user-id* — User ID (Range: 1–128). • *timeout* — Session timeout. 0 indicates timeout not enforced (Range: 0–86400 seconds)
user group *group-id* **moveusers** *new-group-id*	(Optional) Move all of the users in a group to a different group. This command removes the users from the group specified by *group-id*. • *group-id* — Group ID (Range: 1–10). • *new-group-id* — Group ID (Range: 1–10).
CTRL + Z	Exit to Privileged EXEC mode.
show captive-portal user [*user-id*]	View summary information about all users configured in the local database. Specify the user ID to view additional information about a user. *user-id* — User ID (Range: 1–128).

Command	Purpose
clear captive portal users	(Optional) Delete all CP user entries from the local database.

Managing Captive Portal Clients

The commands in this section are all executed in Privileged EXEC mode. Use the following commands to view and manage clients that are connected to a CP.

Command	Purpose
show captive-portal configuration [*cp-id*] **client status**	Display information about the clients authenticated to all CP configurations or a to specific configuration. *cp-id* — The CP instance (Range: 1–10).
show captive-portal interface *interface* **client status**	Display information about clients authenticated on all interfaces or no a specific interface. *interface* — Specific Ethernet interface, such as gi1/0/8.
show captive-portal client [*macaddr*] **status**	Display client connection details or a connection summary for connected CP users. *macaddr* — The MAC address of the client.
captive-portal client deauthenticate *macaddr*	Deauthenticate a specific CP client. *macaddr* — The MAC address of the client.

Captive Portal Configuration Example

The manager of a resort and conference center needs to provide wired Internet access to each guest room at the resort and in each conference room. Due to legal reasons, visitors and guests must agree to the resort’s acceptable use policy to gain network access. Additionally, network access from the conference rooms must be authenticated. The person who rents the conference room space receives a list username and password combinations upon arrival. Hotel employees have their own CP.

The network administrator for the resort and conference center decides to configure the three CPs Table 10-16 describes.

Table 10-16. Captive Portal Instances

CP Name	Description
Guest	Free Internet access is provided in each guest room, but guests must enter a name and agree to the acceptable use policy before they can gain access. The manager wants guests to be redirected to the resort’s home web page upon successful verification. No logout is required.
Conference	Because physical access to the conference rooms is less secure than access to each guest room, the manager wants to ensure that people who connect to the network through a port in a conference room are authenticated. The Conference CP uses the local database for authentication.
Employee	To gain network access, resort employees must enter a username and password that is stored on a RADIUS server.

Configuration Overview

The following steps provide an overview of the process you use to configure the CP feature. In addition to the following steps, IP Helper/DHCP relay should be configured (not shown) if routing is enabled so that clients can obtain an IP address from a DHCP server. Ensure that a DNS server is configured in the network to resolve domain names in user-entered URLs to IP addresses. Refer to "Layer-2 and Layer-3 Relay Features " on page 1157 for further information.

To configure the switch:

1. If a RADIUS server is selected for authentication, configure the RADIUS server settings on the switch.
2. If authentication is required, configure the user groups to associate with each CP.
3. Create (add) the CPs.
4. Configure the CP settings for each CP, such as the verification mode.
5. Associate interfaces with the CP instances.
6. Download the branding images, such as the company logo, to the switch.

   The images you download must be accessible from the switch, either on the system you use to manage the switch or on a server that is on the same network as the switch.

   NOTE: You must use the web interface to download images.

7. Customize the authentication, logout, and logout success web pages that a CP user will see.

   Dell recommends the use of the Dell OpenManage Administrator to customize the CP authentication, logout, and logout success pages. A **Preview** button is available to allow a preview of the pages that a CP user will see.

8. If the local database for user authentication is selected, configure the users on the switch.
9. If a RADIUS server for authentication is selected, add the users to the database on the RADIUS server. If the Captive Portal clients use DNS or DHCP services, configure access to the appropriate services using IP Helper (routed networks) or layer-2 relay (switched networks).
10. Associate interfaces with the CP instances.
11. Test and verify end-user access over the Captive Portal ports and, if satisfied all services are available and working correctly, globally enable CP.

Detailed Configuration Procedures

Use the following steps to perform the CP configuration:

1. Configure the RADIUS server information on the switch.

   In this example, the RADIUS server IP address is 192.168.2.188, and the RADIUS server name is luxury-radius.

```
console#configure
console(config)#radius-server host 192.168.12.182
console(Config-auth-radius)#name luxury-radius
console(Config-auth-radius)#exit
```

2. Configure the CP groups.

```
console(config)#captive-portal
console(config-CP)#user group 2 name Conference
console(config-CP)#user group 3 name Employee
console(config-CP)#exit
```

3. Configure the Guest CP.

```
console(config)#captive-portal
console(config-CP)#configuration 2
console(config-CP 2)#name Guest
console(config-CP 2)#redirect
console(config-CP 2)#redirect-url
http://www.luxuryresorturl.com
console(config-CP 2)#interface te1/0/1
console(config-CP 2)#interface te1/0/2
...
console(config-CP 2)#interface te1/0/4
console(config-CP 2)#exit
```

4. Configure the Conference CP.

```
console(config-CP)#configuration 3
console(config-CP 3)#name Conference
console(config-CP 3)#verification local
console(config-CP 3)#group 2
console(config-CP 4)#interface te1/0/8
...
console(config-CP 4)#interface te1/0/15
console(config-CP 3)#exit
```

5. Configure the Employee CP.

```
console(config-CP)#configuration 4
console(config-CP 4)#name Employee
console(config-CP 4)#verification radius
console(config-CP 4)#group 3
```

```
console(config-CP 4)#interface te1/0/18
...
console(config-CP 4)#interface te1/0/40
console(config-CP 4)#exit
```

6. Use the web interface to customize the CP pages that are presented to users when they attempt to connect to the network.

   NOTE: CP page customization is supported only through the web interface. For information about customizing the CP pages, see "Customizing a Captive Portal " on page 325.

7. Add the Conference users to the local database.

```
console(config-CP)#user 1 name EaglesNest1
console(config-CP)#user 1 password
Enter password (8 to 64 characters): *********
Re-enter password: *********
console(config-CP)#user 1 group 2
```

   Continue entering username and password combinations to populate the local database.

8. Add the User-Name, User-Password, Session-Timeout, and Dell-Captive-Portal-Groups attributes for each employee to the database on the RADIUS server.

9. Globally enable CP.

```
console(config-CP)#enable
```

NOTE: Other items that may need to be configured to obtain a working configuration are:
- DNS servers
- DHCP servers
- DHCP Relay or DHCP L2 Relay
- Default gateway (if routing is enabled)

In Case Of Problems in Captive Portal Deployment

When configuring captive portal, many administrators will find that the web browsers or hosts are not able to reach the captive portal web page. This is most often due to network issues as opposed to issues with the captive portal service.

When deploying captive portal, first ensure that web clients on the internal network can reach the external network by disabling captive portal entirely and verifying connectivity. It may be required to configure DHCP relay or DHCP snooping, add one or more default gateways in routed networks, and ensure access to one or more DNS servers. After outside network connectivity is verified by bringing up web pages from the external network on an internal host on the captive portal disabled port, re-enable captive portal.

11

Monitoring and Logging System Information

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter provides information about the features used for monitoring the switch, including logging, cable tests, and email alerting. The topics covered in this chapter include:

- System Monitoring Overview
- Default Log Settings
- Monitoring System Information and Configuring Logging (Web)
- Monitoring System Information and Configuring Logging (CLI)
- Logging Configuration Examples

System Monitoring Overview

What System Information Is Monitored?

The CLI and web-based interfaces provide information about physical aspects of the switch, such as system health and cable diagnostics, as well as information about system events, such as management login history. The switch also reports system resource usage.

The system logging utility can monitor a variety of events, including the following:

- System events — System state changes and errors that range in severity from Emergency to Debug
- Audit events — Attempts to login or logout from the switch and attempts to perform any operations with files on the flash drive
- CLI commands — Commands executed from the CLI
- Web page visits — Pages viewed by using OpenManage Switch Administrator
- SNMP events — SNMP **set** operations

Why Is System Information Needed?

The information the switch provides can help the switch administrator troubleshoot issues that might be affecting system performance. The cable diagnostics test help the administrator troubleshoot problems with the physical connections to the switch. Auditing access to the switch and the activities an administrator performed while managing the switch can help provide security and accountability.

Where Are Log Messages Sent?

The messages the switch generates in response to events, faults, errors, and configuration changes can be recorded in several locations. By default, these messages are stored locally on the switch in the RAM (cache). This collection of log files is called the RAM log or buffered log. When the RAM log file reaches the configured maximum size, the oldest message is deleted from the RAM when a new message is added. If the system restarts, all messages are cleared.

In addition to the RAM log, log files can be sent to the following sources:

- Console — If the administrator is connected to the switch CLI through the console port, messages display to the screen as they are generated. Use the **terminal monitor** command to control logging of messages to the console when connected via Telnet or SSH.
- Log file — Messages sent to the log file are saved in the flash memory and are not cleared when the system restarts.
- Remote server — Messages can be sent to a remote log server for viewing and storage.
- Email — Messages can be sent to one or more email addresses. Information about the network Simple Mail Transport Protocol (SMTP) server must be configured for email to be successfully sent from the switch.

What Are the Severity Levels?

The severity of the messages to be logged for each local or remote log file can be specified. Each severity level is identified by a name and a number. Table 11-1 provides information about the severity levels.

Table 11-1. Log Message Severity

Severity Keyword	Severity Level	Description
emergencies	0	The switch is unusable.
alerts	1	Action must be taken immediately.
critical	2	The switch is experiencing critical conditions.
errors	3	The switch is experiencing error conditions.
warnings	4	The switch is experiencing warning conditions.
notification	5	The switch is experiencing normal but significant conditions.
informational	6	The switch is providing non-critical information.
debugging	7	The switch is providing debug-level information.

When the severity level is specified, messages with that severity level and higher are sent to the log file. For example, if the severity level is specified as critical, messages with a severity level of alert and emergency are also logged. When the severity level is specified in a CLI command, the keyword or the numerical level can be used.

What Are the System Startup and Operation Logs?

Two types of log files exist in flash (persistent) memory:

- The first log type is the system startup log. The system startup log stores the first 32 messages received after system reboot. The log file stops when it is full.
- The second log type is the system operation log. The system operation log stores the last 1000 messages received during system operation. The oldest messages are overwritten when the file is full.

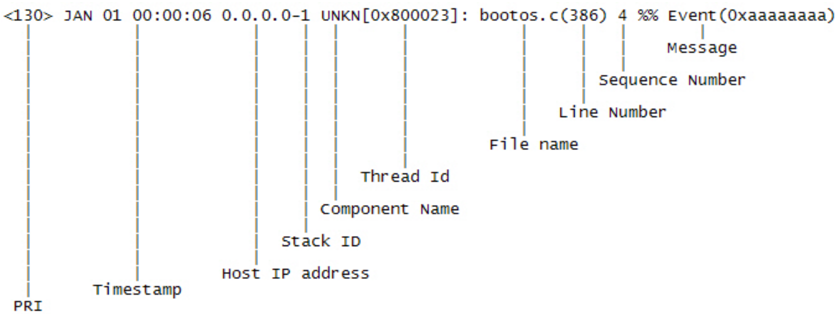
A message is only logged in one file. On system startup, if the Log file is enabled, the startup log stores messages up to its limit. Then the operation log begins to store the messages.

To view the log messages in the system startup and operational log files, the log files must be download to an administrative host. The startup log files are named slogX.txt and the operation log files are named ologX.txt. When enabled, the system stores the startup and operation log files for the last three switch boots. The current log files have a zero (0) in the file name (replacing the X in the name as shown above), the prior log files contain a one (1) in the name, and the oldest log files contain a two (2) in the name. For more information about downloading files, see "Images and File Management " on page 469.

What Is the Log Message Format?

The first part of the log message up to the first left bracket is fixed by the syslog standard (RFC 3164). The second part up to the two percent signs is standardized for all Dell Networking logs. The variable text of the log message follows. The log message is limited to 96 bytes.

Each log message uses the following format:



- PRI—This consists of the facility code (see RFC 3164) multiplied by 8 and added to the severity. The log messages use the local7 facility code (23). This implies that a message of severity 0 will have a priority of 184 and a message of severity 7 will have a priority of 191.
- Timestamp—This is the system up time. For systems that use SNTP, this is UTC. When time zones are enabled, local time will be used.
- Host IP address—This is the IP address of the local system.

- Stack ID —This is the assigned stack ID. For the Dell Networking N1500, N2000, N3000, and N4000 Series switches, the stack ID number is always 1. The number 1 is used for systems without stacking ability. The stack master is used to collect messages for the entire stack.
- Component name—The component name for the logging component. Component "UNKN" is substituted for components that do not identify themselves to the logging component.
- Thread ID—The thread ID of the logging component.
- File name —The name of the file containing the invoking macro.
- Line number —The line number which contains the invoking macro.
- Sequence number —The message sequence number for this stack component. Sequence numbers may be skipped because of filtering but are always monotonically increasing on a per-stack member basis.
- Message — Contains the text of the log message.

What Factors Should Be Considered When Configuring Logging?

Dell recommends that network administrators deploy a syslog server in their network and configure all switches to log messages to the syslog server. Switch administrators should also consider enabling persistent logging on the switch.

When managing logs on a stack of switches, the RAM log and persistent log files exist only on the top of stack platform. Other platforms in the stack forward their messages to the top of stack log.

Logging of debug level messages is intended for use by support personnel. The output is voluminous and cryptic and, because of the large number of messages generated, can adversely affect switch operations. Set the logging level to debug only under the direction of support personnel.

Default Log Settings

System logging is enabled, and messages are sent to the console (severity level: warning and above), and RAM log (severity level: informational and above). Switch auditing, CLI command logging, Web logging, and SNMP logging are disabled. By default, no messages are sent to the log file that is stored in flash, and no remote log servers are defined.

Email alerting is disabled, and no recipient email address is configured. Additionally, no mail server is defined. If a mail server is added, by default, no authentication or security protocols are configured, and the switch uses TCP port 25 for SMTP.

After email alerting is enabled and the mail server and recipient email address are configured, log messages with a severity level of emergency and alert are sent immediately with each log message in a separate mail. The email subject is “Urgent Log Messages.” Log messages with a severity level of critical, error, and warning are sent periodically in a single email. The email subject is “Non Urgent Log Messages.” Messages with a severity level of notice and below are not sent in an email.

Monitoring System Information and Configuring Logging (Web)

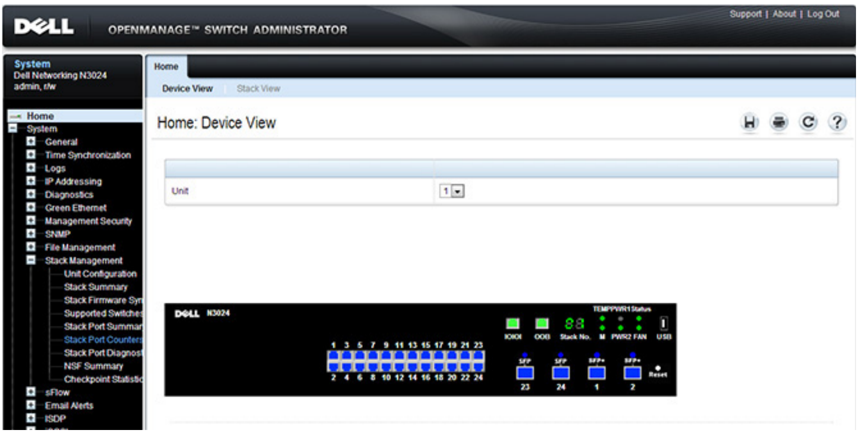
This section provides information about the OpenManage Switch Administrator pages to use to monitor system information and configure logging on the Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click at the top of the page.

Device Information

The **Device Information** page displays after you successfully log on to the switch by using the Dell OpenManage Switch Administrator. This page is a virtual representation of the switch front panel. Use the **Device Information** page to view information about the port status, or system status, and the switch stack. Click on a port to access the **Port Configuration** page for the selected port.

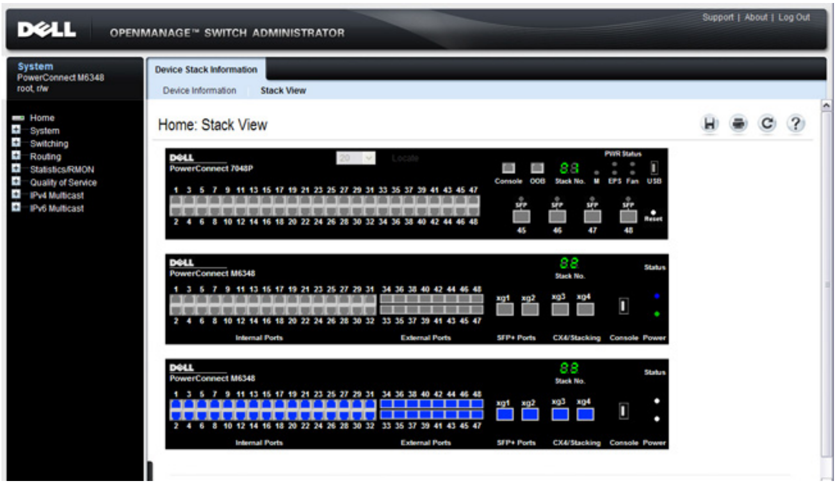
To display the **Device Information** page, click **Home** in the navigation panel.

Figure 11-1. Device Information



Click the **Stack View** link to view front panel representations for all units in the stack.

Figure 11-2. Stack View



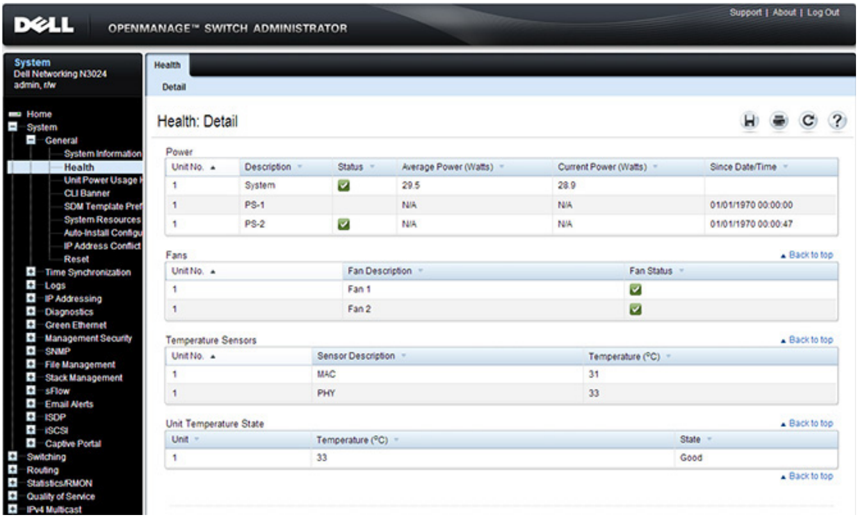
For more information about the device view features, see "Understanding the Device View " on page 150.

System Health

Use the **Health** page to view status information about the switch power and ventilation sources.

To display the **Health** page, click **System** → **General** → **Health** in the navigation panel.

Figure 11-3. Health

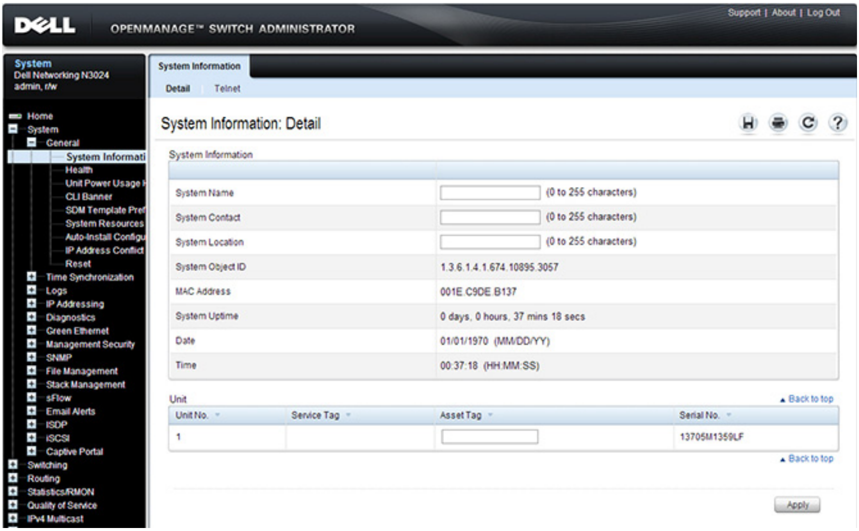


System Resources

Use the **System Resources** page to view information about memory usage and task utilization.

To display the **System Resources** page, click **System → General → System Resources** in the navigation panel.

Figure 11-4. System Resources

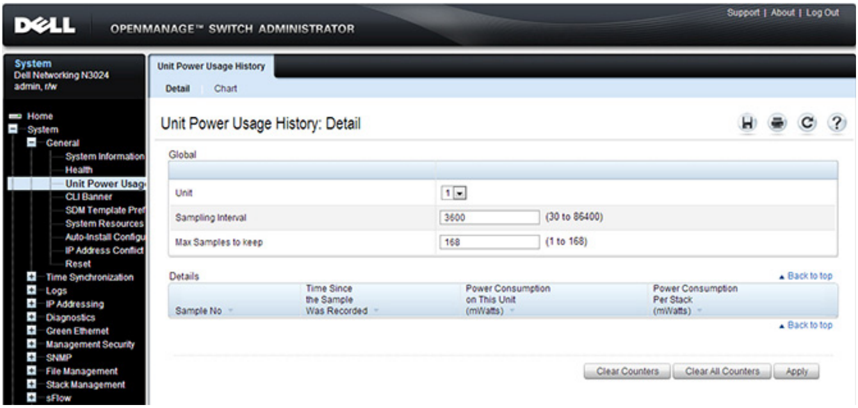


Unit Power Usage History

Use the **Unit Power Usage History** page to view information about switch power consumption.

To display the **Unit Power Usage History** page, click **System** → **General** → **Unit Power Usage History** in the navigation panel.

Figure 11-5. Unit Power Usage History



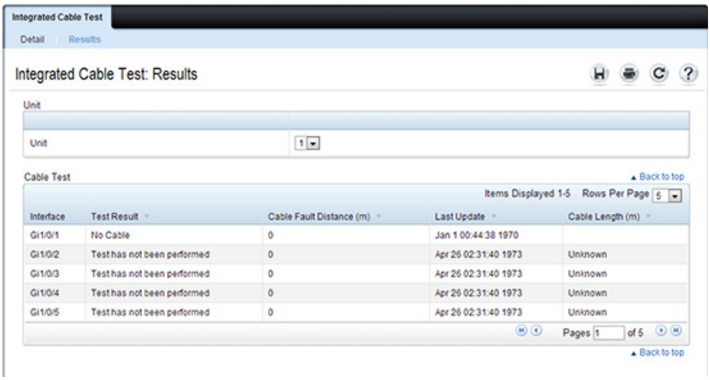
Integrated Cable Test for Copper Cables

Use the **Integrated Cable Test for Copper Cables** page to perform tests on copper cables. Cable testing provides information about where errors occurred in the cable, the last time a cable test was performed, and the type of cable error which occurred. The tests use Time Domain Reflectometry (TDR) technology to test the quality and characteristics of a copper cable attached to a port. Cables up to 120 meters long can be tested. Cables are tested when the ports are in the down state, with the exception of the Approximated Cable Length test. SFP, SFP+, and QSFP cables with passive copper assemblies are not capable of performing TDR tests.

NOTE: Cable diagnostics may give misleading results if any green Ethernet modes are enabled on the port. Disable EEE and energy-detect mode prior to running any cable diagnostics.

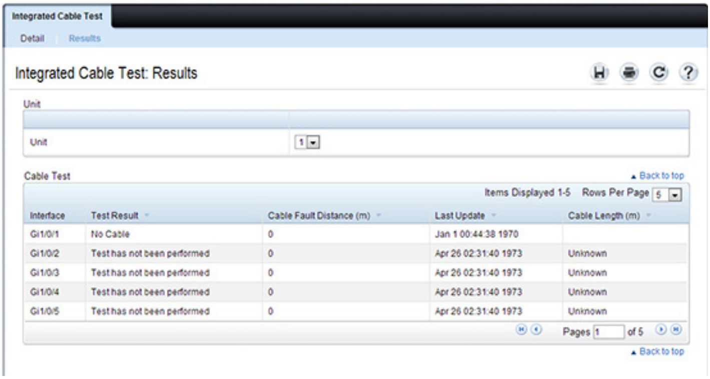
To display the **Integrated Cable Test for Copper Cables** page, click **System** → **Diagnostics** → **Integrated Cable Test** in the navigation panel.

Figure 11-6. Integrated Cable Test for Copper Cables



To view a summary of all integrated cable tests performed, click the **Show All** link.

Figure 11-7. Integrated Cable Test Summary



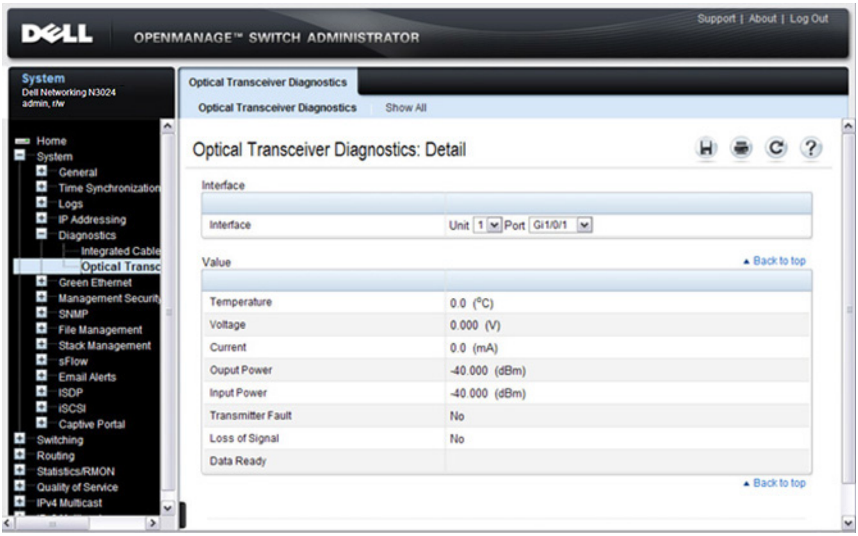
Optical Transceiver Diagnostics

Use the **Transceiver Diagnostics** page to perform tests on Fiber Optic cables.

To display the **Transceiver Diagnostics** page, click **System** → **Diagnostics** → **Transceiver Diagnostics** in the navigation panel.

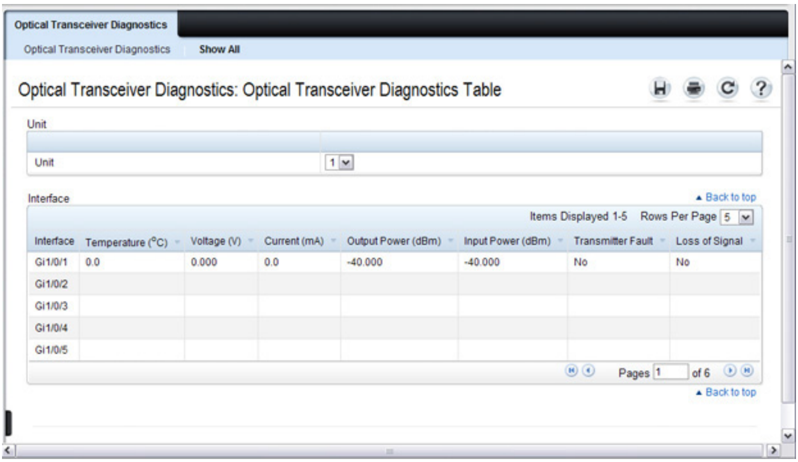
NOTE: Optical transceiver diagnostics can be performed only when the link is present.

Figure 11-8. Transceiver Diagnostics



To view a summary of all optical transceiver diagnostics tests performed, click the **Show All** link.

Figure 11-9. Transceiver Diagnostics Summary



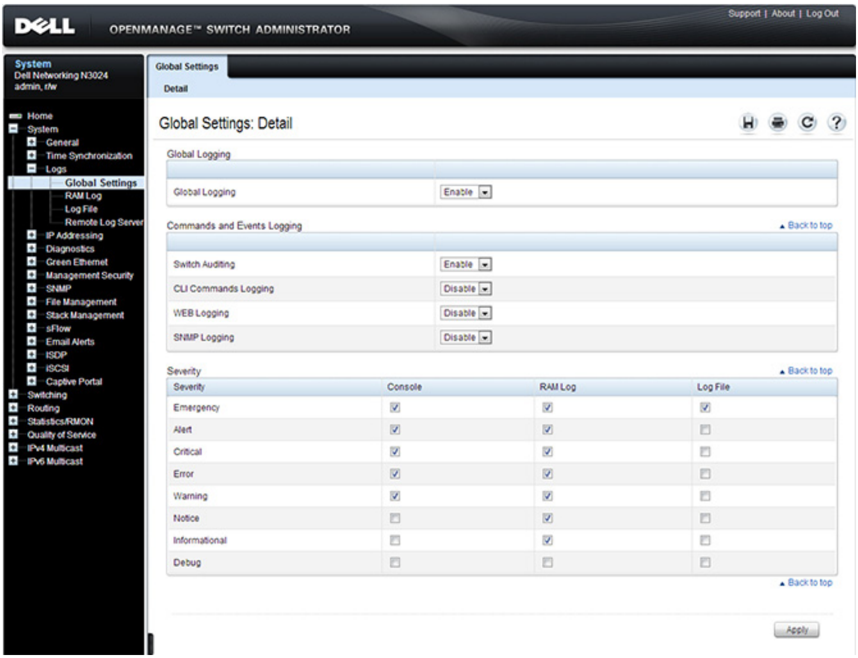
Log Global Settings

Use the **Global Settings** page to enable logging globally, to enable other types of logging. The severity of messages that are logged to the console, RAM log, and flash-based log file can also be specified.

The **Severity** table lists log messages from the highest severity (Emergency) to the lowest (Debug). When a severity level is selected, all higher levels are automatically selected. To prevent log messages from being sent to the console, RAM log, or flash log file, clear all check boxes in the **Severity** column.

To display the **Global Settings** page, click **System** → **Logs** → **Global Settings** in the navigation panel.

Figure 11-10. Global Settings

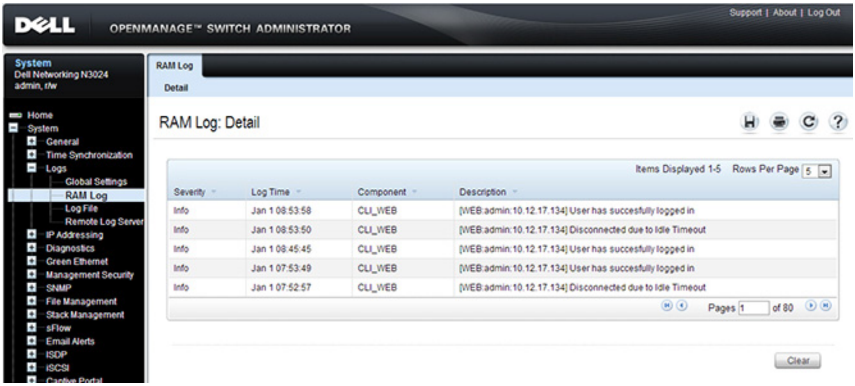


RAM Log

Use the **RAM Log** page to view information about specific RAM (cache) log entries, including the time the log was entered, the log severity, and a description of the log.

To display the **RAM Log**, click **System** → **Logs** → **RAM Log** in the navigation panel.

Figure 11-11. RAM Log Table

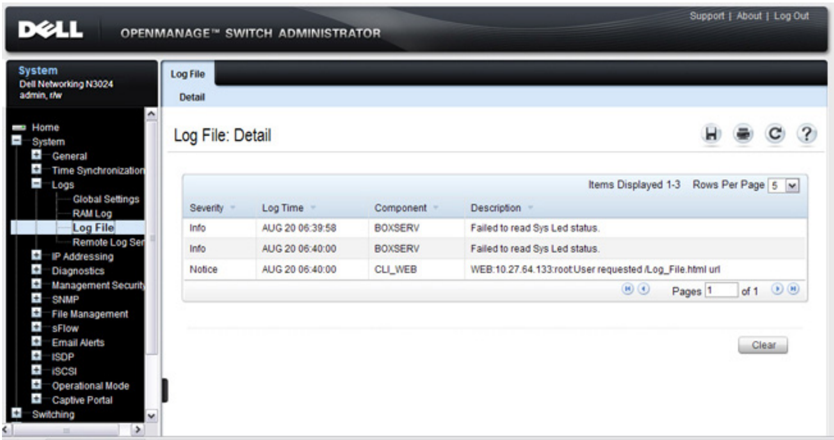


Log File

The **Log File** contains information about specific log entries, including the time the log was entered, the log severity, and a description of the log.

To display the **Log File**, click **System** → **Logs** → **Log File** in the navigation panel.

Figure 11-12. Log File

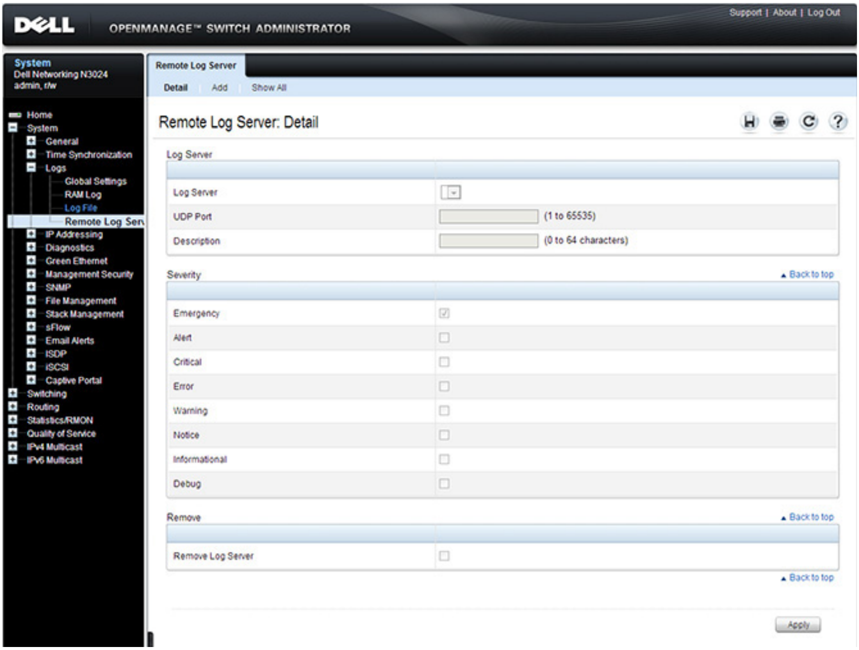


Syslog Server

Use the **Remote Log Server** page to view and configure the available syslog servers, to define new syslog servers, and to set the severity of the log events sent to the syslog server.

To display the **Remote Log Server** page, click **System** → **Logs** → **Remote Log Server**.

Figure 11-13. Remote Log Server

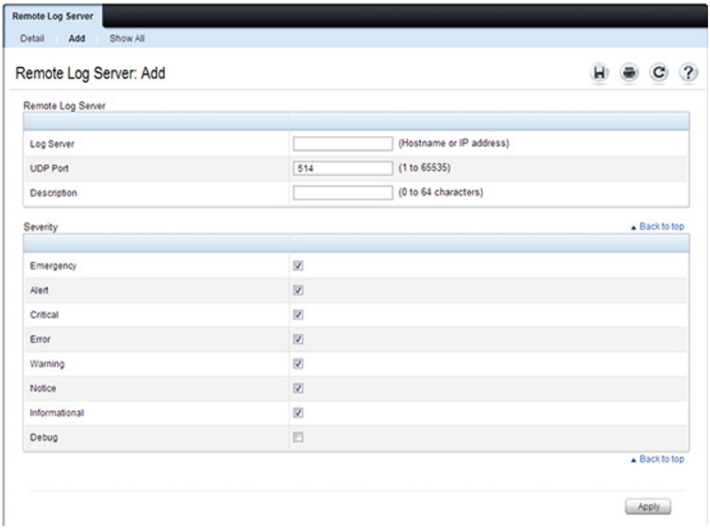


Adding a New Remote Log Server

To add a syslog server:

1. Open the **Remote Log Server** page.
2. Click **Add** to display the **Add Remote Log Server** page.
3. Specify the IP address or hostname of the remote server.
4. Define the **UDP Port** and **Description** fields.

Figure 11-14.   Add Remote Log Server



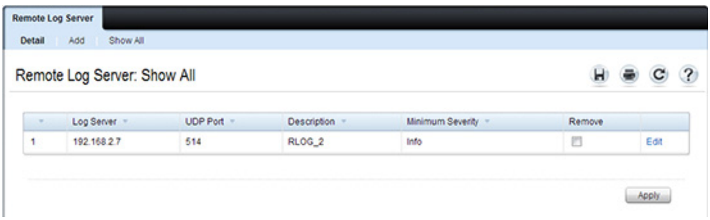
5 Select the severity of the messages to send to the remote server.

NOTE: When a severity level is selected, all higher (numerically lower) severity levels are automatically selected.

6 Click **Apply**.

Click the **Show All** link to view or remove remote log servers configured on the system.

Figure 11-15.   Show All Log Servers

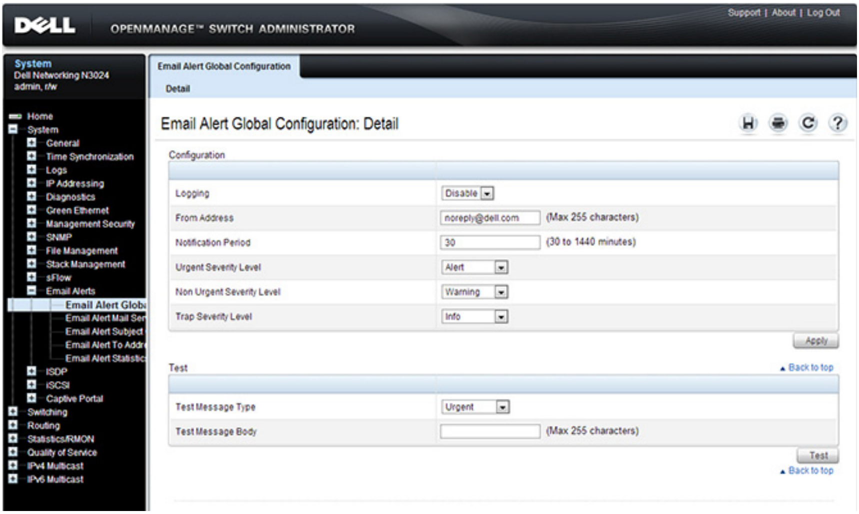


Email Alert Global Configuration

Use the **Email Alert Global Configuration** page to enable the email alerting feature and configure global settings so that system log messages can be sent to from the switch to one or more email accounts.

To display the **Email Alert Global Configuration** page, click **System** → **Email Alerts** → **Email Alert Global Configuration** in the navigation panel.

Figure 11-16. Email Alert Global Configuration

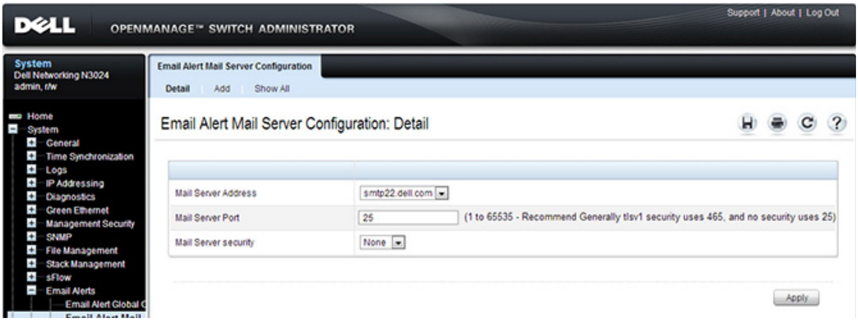


Email Alert Mail Server Configuration

Use the **Email Alert Mail Server Configuration** page to configure information about the mail server the switch uses for sending email alert messages.

To display the **Email Alert Mail Server Configuration** page, click **System** → **Email Alerts Email Alert Mail Server Configuration** in the navigation panel.

Figure 11-17. Email Alert Mail Server Configuration



Adding a Mail Server

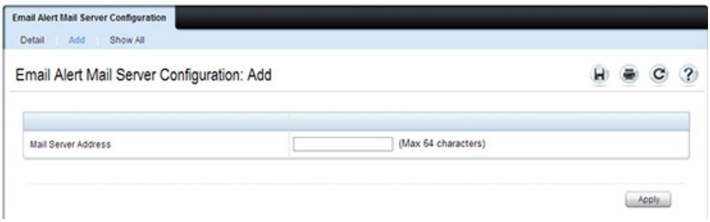
To add a mail server:

1 Open the **Email Alert Mail Server Configuration** page.

2 Click **Add** to display the **Email Alert Mail Server Add** page.

3 Specify the hostname of the mail server.

Figure 11-18. Add Mail Server

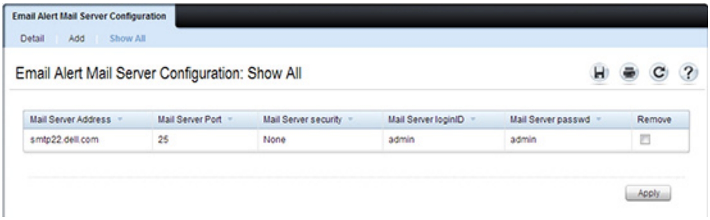


4 Click **Apply**.

5 If desired, click **Configuration** to return to the **Email Alert Mail Server Configuration** page to specify port and security settings for the mail server.

Click the **Show All** link to view or remove mail servers configured on the switch.

Figure 11-19. Show All Mali Servers

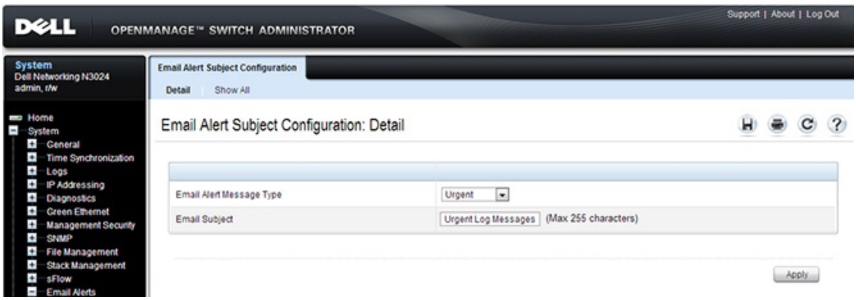


Email Alert Subject Configuration

Use the **Email Alert Subject Configuration** page to configure the subject line for email alerts that are sent by the switch. The subject for the message severity and entry status can customize be customized.

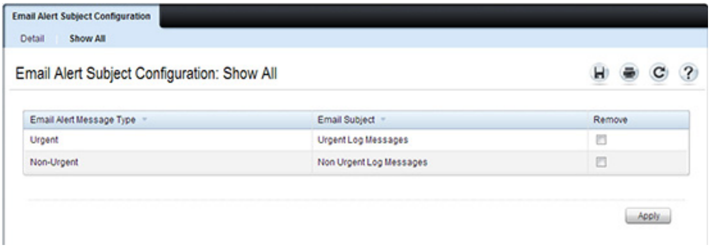
To display the **Email Alert Subject Configuration** page, click **System** → **Email Alerts Email Alert Subject Configuration** in the navigation panel.

Figure 11-20. Email Alert Subject Configuration



To view all configured email alert subjects, click the **Show All** link.

Figure 11-21. View Email Alert Subjects

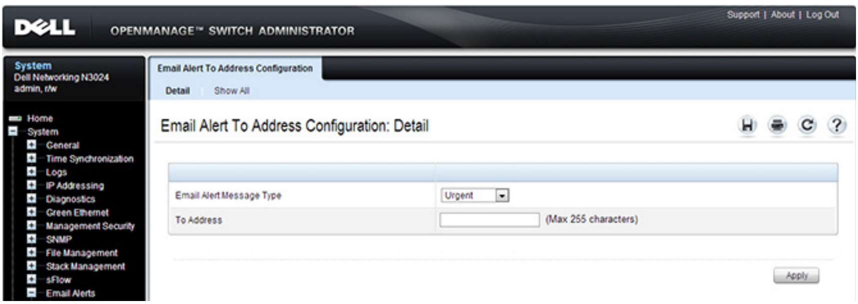


Email Alert To Address Configuration

Use the **Email Alert To Address Configuration** page to specify where the email alerts are sent. Multiple recipients can be configured and different message severity levels can be associated with different recipient addresses.

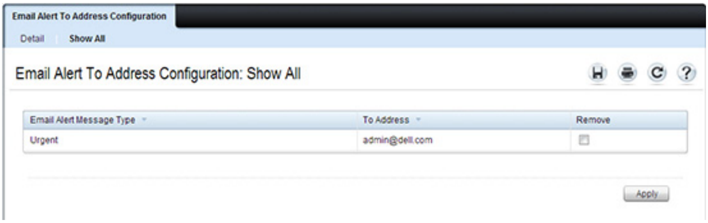
To display the **Email Alert To Address Configuration** page, click **System** → **Email Alerts** → **Email Alert To Address Configuration** in the navigation panel.

Figure 11-22. Email Alert To Address Configuration



To view configured recipients, click the **Show All** link.

Figure 11-23. View Email Alert To Address Configuration

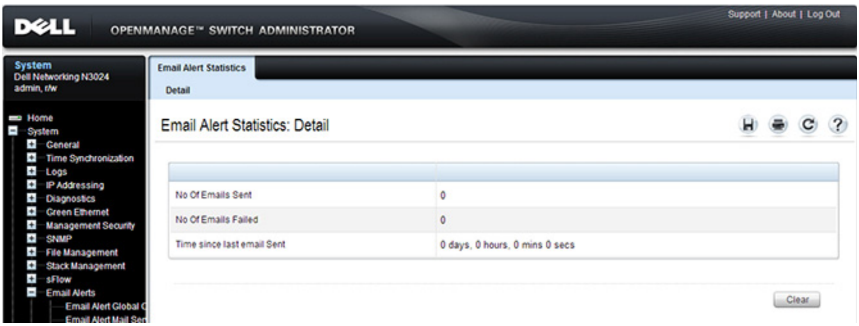


Email Alert Statistics

Use the **Email Alert Statistics** page to view the number of emails that were successfully and unsuccessfully sent, and when emails were sent.

To display the **Email Alert Statistics** page, click **System >Email Alerts Email Alert Statistics** in the navigation panel.

Figure 11-24. Email Alert Statistics



Monitoring System Information and Configuring Logging (CLI)

This section provides information about the commands used for configuring features for monitoring on the Dell Networking N1500, N2000, N3000, and N4000 Series switches. For more information about these commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Viewing System Information and Enabling the Locator LED

Beginning in Privileged EXEC mode, use the following commands to view system health and resource information and to enable the switch locator LED.

Command	Purpose
show system	Display various system information.
show system power	Display the power supply status.
show system temperature	Display the system temperature and fan status.
show memory cpu	Display the total and available RAM space on the switch.
show process cpu	Display the CPU utilization for each process currently running on the switch.
locate [**switch** *unit*] [**time** *time*]	Enable the switch locator LED located on the back panel of the switch. Optionally, the unit to identify within a switch stack and the length of time that the LED blinks can be specified.

Running Cable Diagnostics

Beginning in Privileged EXEC mode, use the following commands to run the cable diagnostic tests.

NOTE: Cable diagnostics may give misleading results if green mode is enabled on the port. Disable green mode prior to running any cable diagnostics.

Command	Purpose
test copper-port tdr *interface*	Perform the Time Domain Reflectometry (TDR) test to diagnose the quality and characteristics of a copper cable attached to the specified port. SFP, SFP+, and QSFP cables with passive copper assemblies are not capable of performing TDR tests. **CAUTION: Issuing the test copper-port tdr command will bring the interface down.** **NOTE:** To ensure accurate measurements, disable all Green Ethernet modes (EEE and energy-detect mode) on the port before running the test. The interface is specified in unit/slot/port format. For example 1/0/3 is GbE interface 3 on unit 1 of the stack.
show copper-ports tdr [*interface*]	Display the diagnostic information collected by the test **copper-port tdr** command for all copper interfaces or a specific interface.
show fiber-ports optical-transceiver [*interface*]	Display the optical transceiver diagnostics for all fiber ports. Include the *interface* option to show information for the specified port.

Configuring Local Logging

Beginning in Privileged EXEC mode, use the following commands to configure the type of messages that are logged and where the messages are logged locally.

Command	Purpose
configure	Enter Global Configuration mode.
logging on	Globally enables logging.
logging audit	Enable switch auditing.
logging cli-command	Enable CLI command logging
logging web-sessions	Enable logging of the switch management Web page visits.
logging snmp	Enable logging of SNMP **set** commands.
terminal monitor	Enable display of system messages on the console for Telnet/SSH sessions.
logging {buffered|console| file} [*severity*]	Enable logging to the specified file. Optionally, a logging discriminator can be defined to help filter log messages and set the severity of the messages to log. • **buffered** — Enables logging to the RAM file (cache). If the switch resets, the buffered logs are cleared. • **console** — Enables logging to the screen when the administrator is connected to the CLI through the console port. • **file** — Enables logging to the startup and operational log files on the flash. • **discriminator** *disc-name* — (Optional) Include a message discriminator to help filter log messages. The *disc-name* can contain up to eight alphanumeric characters. Spaces are not permitted. • *severity* — (Optional) Enter the number or name of the desired severity level. For information about severity levels, see Table 11-1.
logging facility *facility-type*	Set the facility for logging messages. Permitted *facility-type* values are local0, local1, local2, local3, local4, local5, local 6, local7

Command	Purpose
CTRL + Z	Exit to Privileged EXEC mode.
show logging	Displays the state of logging and the syslog messages stored in the internal buffer.
show logging file	View information about the flash (persistent) log file.
clear logging	Use to clear messages from the logging buffer.

Configuring Remote Logging

Beginning in Privileged EXEC mode, use the following commands to define a remote server to which the switch sends log messages.

Command	Purpose
configure	Enter Global Configuration mode.
logging {*ip-address*| *hostname*}	Define a remote log server and enter the configuration mode for the specified log server.
description *description*	Describe the log server. Use up to 64 characters. If the description includes spaces, surround it with quotation marks.
level *severity*	Specify the severity level of the logs that should be sent to the remote log server. For information about severity levels, see Table 11-1.
port *udp-port*	Specify the UDP port to use for sending log messages. The range is 1 to 65535, and the default is 514.
CTRL + Z	Exit to Privileged EXEC mode.
show syslog-servers	Verify the remote log server configuration.

Configuring Mail Server Settings

Beginning in Privileged EXEC mode, use the following commands to configure information about the mail server (SMTP host) on the network that will initially receive the email alerts from the switch and relay them to the correct recipient.

Command	Purpose
configure	Enter Global Configuration mode.
mail-server *ip-address*	Specify the IP address of the SMTP server on the network and enter the configuration mode for the mail server.
security {**tlsvl**|**none**}	(Optional) Specify the security protocol to use with the mail server.
port {25|465}	Configure the TCP port to use for SMTP, which can be 25 (SMTP) or 465 (SMTP over SSL).
username *username*	If the SMTP server requires authentication, specify the username to use for the switch. The same username and password settings must be configured on the SMTP host.
password *password*	If the SMTP server requires authentication from clients, specify the password to associate with the switch username.
CTRL + Z	Exit to Privileged EXEC mode.
show mail-server all config	View mail server configuration information for all configured mail servers.

Configuring Email Alerts for Log Messages

Beginning in Privileged EXEC mode, use the following commands to configure email alerts so that log messages are sent to the specified address.

Command	Purpose
configure	Enter Global Configuration mode.
logging email [*severity*]	Enable email alerting and determine which non-critical log messages should be emailed. Including the *severity* value sets the lowest severity for which log messages are emailed. These messages are collected and sent in a single email at the configured log duration. *severity* — (Optional) Enter the number or name of the severity level for non-critical messages. Log messages at or above this severity level are emailed. For information about severity levels, see Table 11-1. Log messages below the specified level are not emailed.
logging email urgent {*severity* | **none**}	Determine which log messages are critical and should be sent in a single email as soon as they are generated. *severity* — (Optional) Enter the number or name of the severity level for critical messages. For information about severity levels, see Table 11-1.
logging email logtime *minutes*	Specify how often to send the non-critical email alerts that have been collected. . The valid range is 30 - 1440 minutes.
logging email message-type {urgent | non-urgent | both} to-addr *email-address*	Specify the email address of the recipient for log messages.
logging email from-addr *email-address*	Specify the email address of the sender, which is the switch.
logging email message-type {urgent | non-urgent | both} subject *subject*	Specify the text that will appear in the subject line of email alerts sent by the switch.

Command	Purpose
logging email test message-type {urgent | non-urgent | both} message-body *body*	Send a test email to the configured recipient to verify that the feature is properly configured.
CTRL + Z	Exit to Privileged EXEC mode.
show logging email config	View the configured settings for email alerts.
show logging email statistics	View information about the number of emails sent and the time they were sent.
clear logging email statistics	Clear the email alerting statistics.

Logging Configuration Examples

This section contains the following examples:

- Configuring Local and Remote Logging
- Configuring Email Alerting

Configuring Local and Remote Logging

This example shows how to enable switch auditing and CLI command logging. Log messages with a severity level of Notification (level 5) and above are sent to the RAM (buffered) log. Emergency, Critical, and Alert (level 2) log messages are written to the log file on the flash drive. All log messages are displayed on the console and sent to a remote syslog server.

To configure the switch:

1 Enable switch auditing and CLI command logging.

```
console#configure
console(config)#logging audit
console(config)#logging cli-command
```

2 Specify where the logs are sent locally and what severity level of message is to be logged. The severity can be specified as the level number, as shown in the first two commands, or as the keyword, as shown in the third command.

```
console(config)#logging buffered 5
console(config)#logging file 2
console(config)#logging console debugging
```

3 Define the remote log server.

```
console(config)#logging 192.168.2.10
console(Config-logging)#description "Syslog Server"
console(Config-logging)#level debug
console(Config-logging)#exit
console(config)#exit
```

4 Verify the remote log server configuration.

```
console#show syslog-servers

IP Address/Hostname        Port    Severity        Description
------------------------- ------  --------------  ----------
192.168.2.10              514     debugging      Syslog Server
```

5 Verify the local logging configuration and view the log messages stored in the buffer (RAM log).

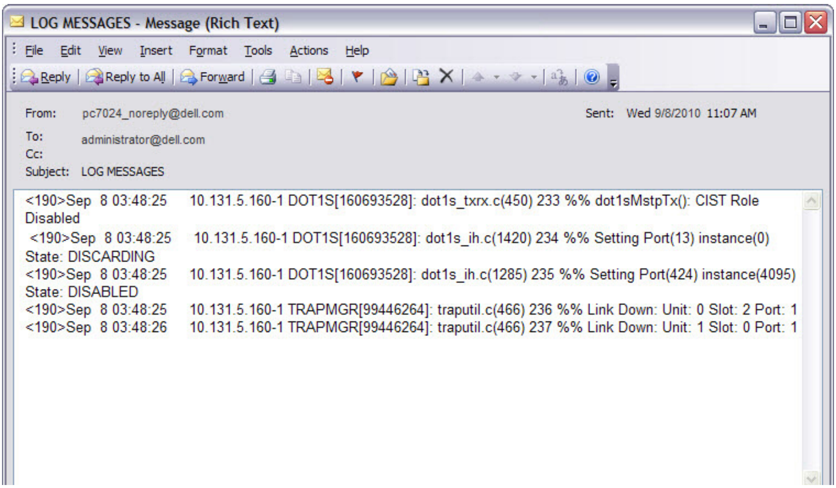
```
console#show logging

Logging is enabled
Console Logging: level debugging. Console Messages: 748
Dropped.
Buffer Logging: level notifications. Buffer Messages: 79
Logged,
File Logging: level critical. File Messages: 973 Dropped.
CLI Command Logging : enabled
Switch Auditing : enabled
Web Session Logging : disabled
SNMP Set Command Logging : disabled
Syslog server  192.168.2.10 logging: debug. Messages: 0 dropped
412 Messages dropped due to lack of resources.
Buffer Log:
<186> FEB 02 05:53:03 0.0.0.0-1 UNKN[1073741088]:
bootos.c(232) 1 %% Event(0xaaaaaaaa)
<189> FEB 02 05:53:03 0.0.0.0-1 UNKN[1073741088]:
bootos.c(248) 2 %% Starting code... BSP initialization
complete, starting application.
```

Configuring Email Alerting

The commands in this example define the SMTP server to use for sending email alerts. The mail server does not require authentication and uses the standard TCP port for SMTP, port 25, which are the default values. Only Emergency messages (severity level 0) will be sent immediately as individual emails, and messages with a severity of alert, critical, and error (levels 1-3) will be sent in a single email every 120 minutes. Warning, notice, info, and debug messages are not sent in an email.

The email the administrator will in the inbox has a format similar to the following:

Figure 11-25. Email Alert Message Format



For emergency-level messages, the subject is LOG MESSAGE - EMERGENCY. For messages with a severity level of alert, critical, and error, the subject is LOG MESSAGE.

To configure the switch:

1 Specify the mail server to use for sending messages.

```
console#configure
console(config)#mail-server ip-address 192.168.2.34
```

2 Configure the username and password that the switch must use to authenticate with the mail server.

```
console(Mail-Server)#username switchN3048
console(Mail-Server)#password passwordN3048
console(Mail-Server)#exit
```

3 Configure emergencies and alerts to be sent immediately, and all other messages to be sent in a single email every 120 minutes.

```
console(config)#logging email error
console(config)#logging email urgent emergency
console(config)#logging email logtime 120
```

4 Specify the email address of the sender (the switch).

```
console(config)#logging email from-addr N3048_noreply@dell.com
```

5 Specify the address where email alerts should be sent.

```
console(config)#logging email message-type both to-addr
administrator@dell.com
```

6 Specify the text that will appear in the email alert Subject line.

```
console(config)#logging email message-type urgent subject "LOG
MESSAGES - EMERGENCY"
console(config)#logging email message-type non-urgent subject
"LOG MESSAGES"
```

7 Verify the configuration.

```
console#show mail-server all config

Mail Servers Configuration:

No of mail servers configured................. 1

Email Alert Mail Server Address............... 192.168.2.34
Email Alert Mail Server Port.................. 25
Email Alert SecurityProtocol.................. none
Email Alert Username.......................... switchN3048
Email Alert Password.......................... passwordN3048

console#show logging email config

Email Alert Logging........................... enabled
Email Alert From Address......................
N3048_noreply@dell.com
Email Alert Urgent Severity Level............. 0
Email Alert Non Urgent Severity Level......... 3
```

```
Email Alert Trap Severity Level................ 6
Email Alert Notification Period................ 120 min

Email Alert To Address Table:

For Msg Type..........................1

Address1..............................administrator@dell.com

For Msg Type..........................2

Address1..............................administrator@dell.com

Email Alert Subject  Table      :

For Msg Type 1, subject is............LOG MESSAGES - EMERGENCY

For Msg Type 2, subject is............LOG MESSAGE
```

12

Managing General System Settings

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to set system information, such as the hostname, and time settings, and how to select the Switch Database Management (SDM) template to use on the switch.

For the Dell Networking N1500, N2000, and N3000 Series switches, this chapter also describes how to configure the Power over Ethernet (PoE) settings. For the Dell Networking N3000 Series switches, this chapter also describes how to view back-panel expansion slot information.

The topics covered in this chapter include:

- System Settings Overview
- Default General System Information
- Configuring General System Settings (Web)
- Configuring System Settings (CLI)
- General System Settings Configuration Examples

System Settings Overview

The system settings include the information described in Table 12-1. This information helps identify the switch.

Table 12-1. System Information

Feature	Description
System Name	The switch name (host name). If the system name is changed, the CLI prompt changes from `console` to the system name.
System contact	Identifies the person to contact for information regarding the switch.
System location	Identifies the physical location of the switch.
Asset tag	Uniquely identifies the switch. Some organizations use asset tags to identify, control, and track each piece of equipment.

Table 12-1. System Information (Continued)

Feature	Description
CLI Banner	Displays a message upon connecting to the switch or logging on to the switch by using the CLI.
SDM Template	Determines the maximum resources a switch or router can use for various features. For more information, see "What Are SDM Templates? " on page 391

The switch can obtain the time from a Simple Network Time Protocol (SNTP) server, or the time can be set manually. Table 12-2 describes the settings that help the switch keep track of time.

Table 12-2. Time Settings

Feature	Description
SNTP	Controls whether the switch obtains its system time from an SNTP server and whether communication with the SNTP server requires authentication and encryption. Information for up to eight SNTP servers can be configured. The SNTP client on the switch can accept updates from both IPv4 and IPv6 SNTP servers.
Real time clock (RTC)	If SNTP is disabled, the system time and date can be entered manually.
Time Zone	Specifies the offset from Coordinated Universal Time (UTC), which is also known as Greenwich Mean Time (GMT).
Summer Time	In some regions, the time shifts by one hour in the fall and spring. In the United States, this is called daylight saving time.

The Dell Networking N1524P/N1548P, N2024P/N2048P and N3024P/N3048P switch ports are IEEE 802.1at-2009-compliant (PoE Plus) and can provided up to 34.2W of power per port. For more information about PoE Plus support, see "What Are the Key PoE Plus Features for the Dell Networking N1524P/N1548P, N2024P/N2048P, and N3024P/N3048P Switches? " on page 396.

Why Does System Information Need to Be Configured?

Configuring system information is optional. However, it can be helpful in providing administrative information about the switch. For example, if an administrator manages several standalone Dell Networking N-Series switches and has Telnet sessions open with several different switches, the system name can help quickly identify the switch because the host name replaces `console` as the CLI command prompt.

The Banner can provide information about the switch status. For example, if multiple users connect to the switch, the message of the day (MOTD) banner might alert everyone who connects to the switch about a scheduled switch image upgrade.

What Are SDM Templates?

An SDM template is a description of the maximum resources a switch or router can use for various features. Different SDM templates allow different combinations of scaling factors, enabling different allocations of resources depending on how the device is used. In other words, SDM templates enable reallocating system resources to support a different mix of features based on network requirements.

Dell Networking N-Series switches support the following three templates:

- Dual IPv4 and IPv6 (default)
- IPv4 Routing
- IPv4 Data Center

Table 12-3 describes the parameters that are scaled for each template and the per-template maximum value of the parameter.

Table 12-3. SDM Template Parameters and Values

Parameter	Dual IPv4/IPv6	Dual IPv4/IPv6 Data Center	IPv4 Only	IPv4 Data Center
ARP entries				
N1500	4096	4096	1024	0
N2000	4096	4096	1024	6144
N3000	4096	4096	6144	4096
N4000	4096	4096	6144	

Table 12-3. SDM Template Parameters and Values (Continued)

Parameter	Dual IPv4/IPv6	Dual IPv4/IPv6 Data Center	IPv4 Only	IPv4 Data Center
IPv4 unicast routes				
N1500	512	512	1024	0
N2000	512	512	1024	0
N3000	8160	8160	12288	8160
N4000	8160	8160	12288	8160
IPv6 Neighbor Discovery Protocol (NDP) entries				
N1500	512	512	0	0
N2000	512	512	0	0
N3000	2560	2560	0	0
N4000	1024	1024	0	0
IPv6 unicast routes				
N1500	64	64	0	0
N2000	256	256	0	0
N3000	4096	4096	0	0
N4000	4096	4096	0	0
ECMP next hops				
N1500	1	1	1	1
N2000	1	1	1	1
N3000	4	16	16	16
N4000	4	16	4	16
IPv4 multicast routes				
N1500	0	0	0	0
N2000	0	0	0	0
N3000	1536	1536	2048	2048
N4000	512	512	1024	2048

Table 12-3. SDM Template Parameters and Values (Continued)

Parameter	Dual IPv4/IPv6	Dual IPv4/IPv6 Data Center	IPv4 Only	IPv4 Data Center
IPv6 multicast routes				
N1500	0	0	0	0
N2000	0	0	0	0
N3000	512	512	0	0
N4000	256	256	0	0

SDM Template Configuration Guidelines

When the switch is configured to use an SDM template that is not currently in use, the switch must be reloaded for the configuration to take effect.

NOTE: If a unit is attached to a stack and its template does not match the stack's template, then the new unit will automatically reboot using the template used by the management unit. To avoid the automatic reboot, you may first set the template to the template used by the management unit. Then power off the new unit, attach it to the stack, and power it on.

If the IPv4 Routing or IPv4 Data Center template is currently in use and the administrator attempts to configure IPv6 routing features without first selecting the Dual IPv4-IPv6 Routing template, the IPv6 commands do not take effect. IPv6 features are not available when an IPv4-only template is active.

Why is the System Time Needed?

The switch uses the system clock to provide time stamps on log messages. Additionally, some **show** commands include the time in the command output. For example, the **show users login-history** command includes a Login Time field. The system clock provides the information for the Login Time field.

How Does SNTP Work?

SNTP assures accurate switch clock time synchronization. Time synchronization is performed by a network SNTP server.

Time sources are established by Stratums. Stratums define the accuracy of the reference clock. The higher the stratum (where zero is the highest), the more accurate the clock. The switch is at a stratum that is one lower than its time source. For example, if the SNTP server in an internal network is a Stratum 3 device, the switch is a Stratum 4 device.

The switch can be configured to request the time from an SNTP server on the network or it can receive SNTP broadcasts.

Requesting the time from a unicast SNTP server is more secure. Use this method if you know the IP address of the SNTP server on your network. If you allow the switch to receive SNTP broadcasts, any clock synchronization information is accepted, even if it has not been requested by the device. This method is less secure than polling a specified SNTP server.

To increase security, authentication can be required between the configured SNTP server and the SNTP client on the switch. Authentication is provided by Message Digest 5 (MD5). MD5 verifies the integrity of the communication and authenticates the origin of the communication.

What Configuration Is Required for Plug-In Modules?

The Dell Networking N3000/N4000 Series switches support several different plug-in modules (also known as cards) for the expansion slots located on the back of the switch. For information about the slots and the supported modules, see "Hardware Overview " on page 99. The card type can be pre-configured prior to inserting it into the switch.

Hot-swap is supported on the Dell Networking N3000/N4000 Series switch modules.

Before inserting a new module into the expansion slot that was previously occupied by a different type of module, issue a **no slot** command from the CLI so that the switch can recognize the new module.

Once a module has been recognized by the switch, its configuration is stored locally on the switch as the switch default. The module configuration appears in the running-config for informational purposes.

What Are the Key PoE Plus Features for the Dell Networking N1524P/N1548P, N2024P/N2048P, and N3024P/N3048P Switches?

Table 12-4 describes some of the key PoE Plus features the switches support.

Table 12-4. PoE Plus Key Features

Feature	Description
Global Usage Threshold	Provides the ability to specify a power limit as a percentage of the maximum power available to PoE ports. Setting a limit prevents the PoE switch from reaching an overload condition.
Per-Port Power Prioritization	Provides the ability to assign a priority to each PoE port. When the power budget of the PoE switch has been exhausted, the higher-priority ports are given preference over the lower-priority ports. Lower priority ports are automatically stopped from supplying power in order to provide power to higher-priority ports.
Per-Port Power Limit	Configurable power limit for each PoE-Plus port.
Power Management Modes	Supports two power-management modes: • Static—Reserves a guaranteed amount of power for a PoE port. This is useful for powering up devices which draw variable amount of power and provide them an assured power range to operate within. • Dynamic—Power is not reserved for a given port at any point of time. The power available with the PoE switch is calculated by subtracting the instantaneous power drawn by all the ports from the maximum available power. Thus more ports can be powered at the same time. This feature is useful to efficiently power up more number of devices when the available power with the PoE switch is limited
Power Detection Mode	Sets the mode to legacy or 4-point 802.3at plus legacy detection.

Default General System Information

By default, no system information or time information is configured, and the SNTP client is disabled. The default SDM Template applied to the switch is the Dual IPv4-IPv6 template.

The following table shows the default PoE Plus settings for the Dell Networking N1524P/N1548P, N2024P /N2048P, and N3024P/N3048P switches.

Table 12-5. PoE Plus Key Features (Dell Networking N2024P, N2048P, N3024P, N3048P Only)

Feature	Description
Global Usage Threshold	0%
Per-Port Admin Status	Auto
Per-Port Power Prioritization	Enabled (globally, per-port priority is Low
Per-Port Power Limit	None
Power Management Mode	Dynamic
Power Detection Mode	802.3at plus legacy
Power Pairs	alternative-a

Power Management in Guard Band

The Dell Networking N1500P/N2000P/N3000P Series switches support a dynamic guard band, which means that the guard band used varies depending upon the following factors:

- Power Management Mode
- Class of the device being powered up.

Prior to a device being powered up, the switch calculates the following:

> [current power consumption + computed power draw of the new device – (switch power capacity – guard band)].

If this value is less than zero (which means powering up the new PD device will put the total power draw into the guard band or above the switch power capacity), then the switch does not power up the new device.

Dynamic/Static Power Management Mode

In this mode, the dynamic guard band is:

- Class 0 device: User-defined power limit for the port being powered up (this value can be found with the command **show power inline** *interface-id* **detailed**). By default, 32 Watts.
- Class 1 device: 4 Watts.
- Class 2 device: 7 Watts.
- Class 3 device: 15.4 Watts.
- Class 4 device: User defined power limit.

Class-Based Power Management Mode

In this mode, the dynamic guard band is:

- Class 0 device: User defined power limit.
- Class 1 device: 4 Watts.
- Class 2 device: 7 Watts.
- Class 3 device: 15.4 Watts.
- Class 4 AF/AT device: If AF device, 15.4 Watts. If AT device, 36 Watts.

Configuring General System Settings (Web)

This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring general system settings on the Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click at the top of the page.

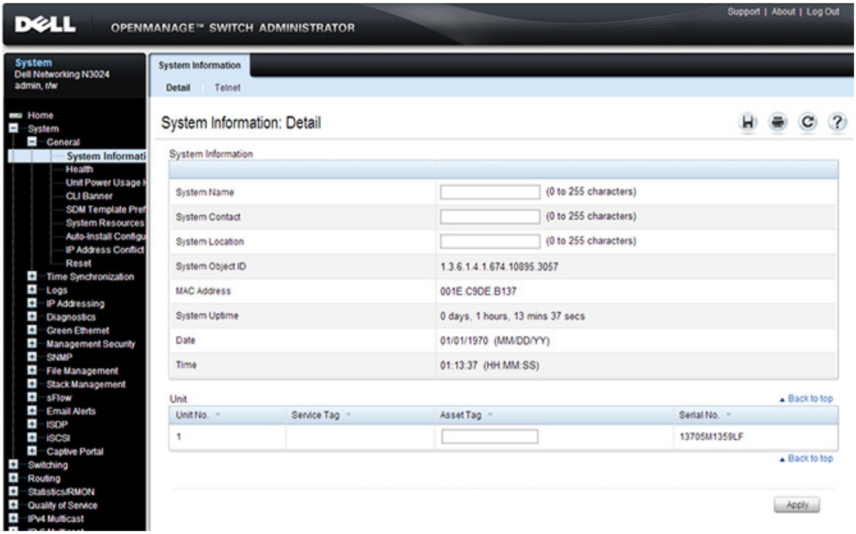
System Information

Use the **System Information** page to configure the system name, contact name, location, and asset tag.

NOTE: A Telnet session to the switch can also be initiated from the **System Information** page.

To display the **System Information** page, click **System** → **General** → **System Information** in the navigation panel.

Figure 12-1. System Information



Initiating a Telnet Session from the Web Interface

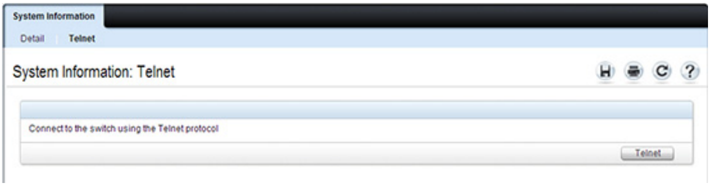
NOTE: The Telnet client feature does not work with Microsoft Windows Internet Explorer 7 and later versions. Initiating this feature from any browser running on a Linux operating system is not supported.

To launch a Telnet session:

1 From the **System** → **General** → **System Information** page, click the Telnet link.

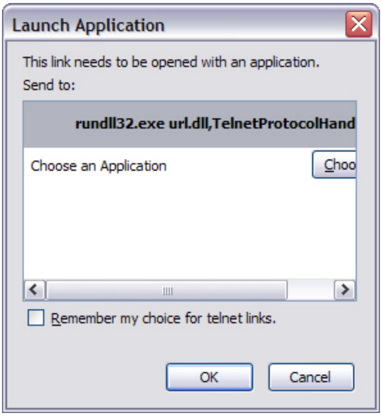
2 Click the **Telnet** button.

Figure 12-2. Telnet



3 Select the Telnet client, and click **OK**.

Figure 12-3. Select Telnet Client



The selected Telnet client launches and connects to the switch CLI.

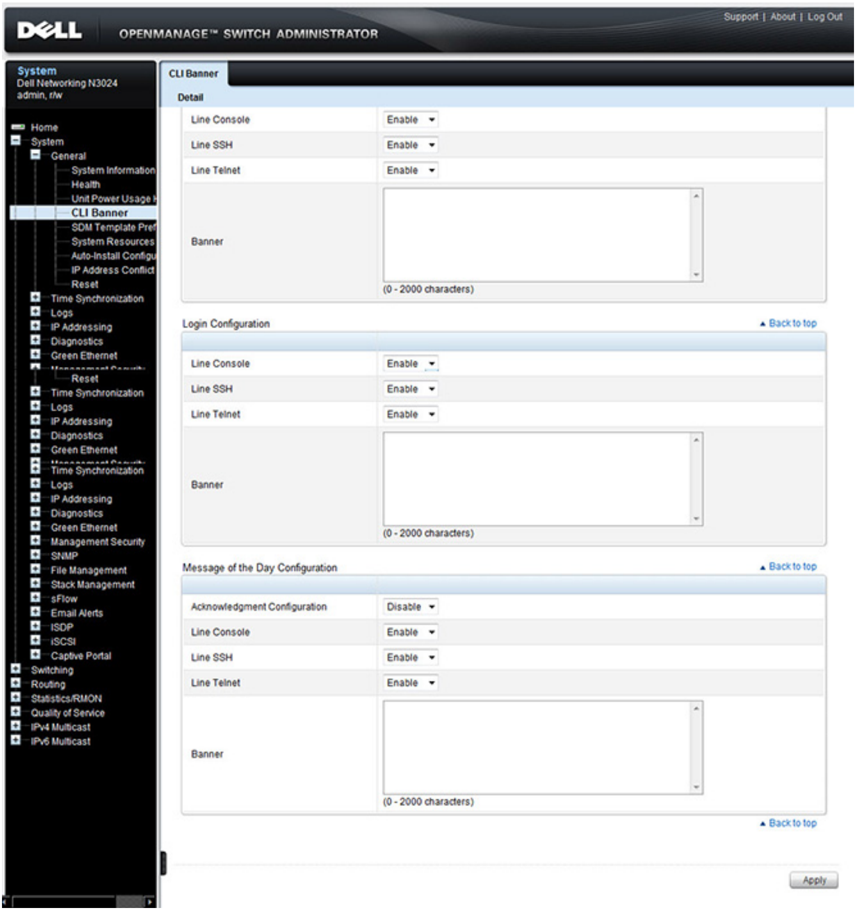
Figure 12-4. Telnet Session

CLI Banner

Use the **CLI Banner** page to configure a message for the switch to display when a user connects to the switch by using the CLI. Different banners can be configured for various CLI modes and access methods.

To display the **CLI Banner** page, click **System** → **General** → **CLI Banner** in the navigation panel.

Figure 12-5. CLI Banner

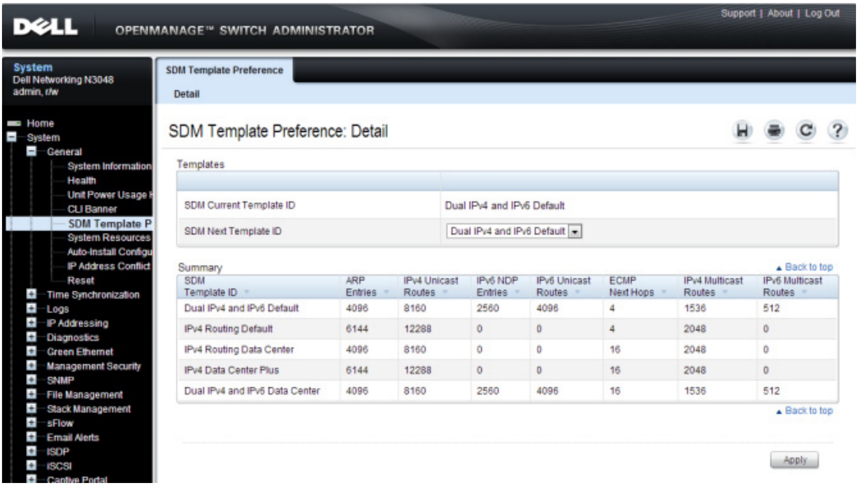


SDM Template Preference

Use the **SDM Template Preference** page to view information about template resource settings and to select the template that the switch uses. If a new SDM template is selected for the switch to use, the switch must be rebooted before the template is applied.

To display the **SDM Template Preference** page, click **System** → **General** → **SDM Template Preference** in the navigation panel.

Figure 12-6. SDM Template Preference

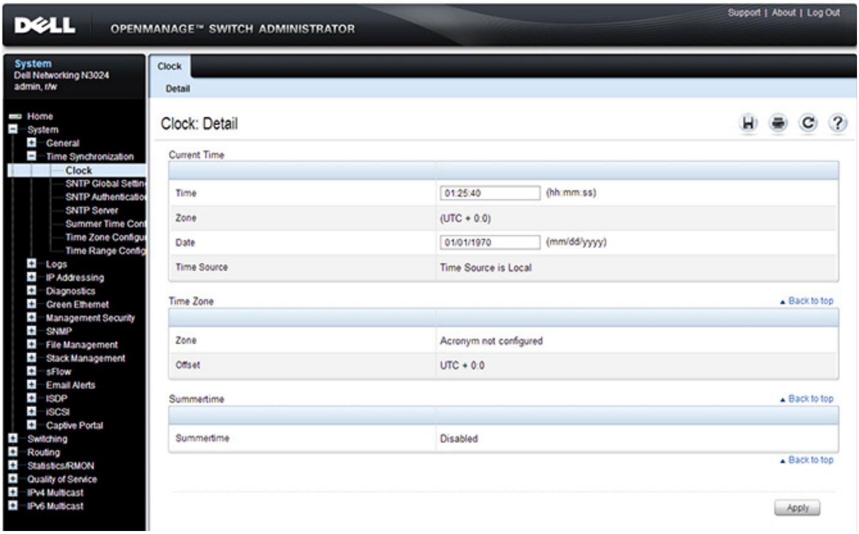


Clock

If the switch is not configured to obtain the system time from an SNTP server, the date and time can be manually set on the switch using the **Clock** page. The **Clock** page also displays information about the time settings configured on the switch.

To display the **Clock** page, click **System** → **Time Synchronization** → **Clock** in the navigation panel.

Figure 12-7. Clock



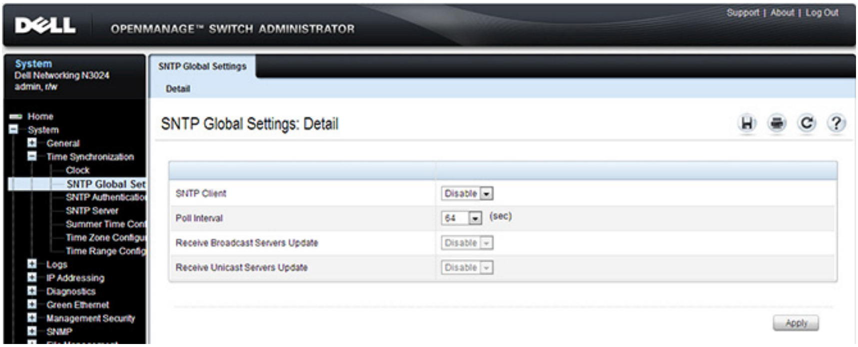
NOTE: The system time cannot be set manually if the SNTP client is enabled. Use the **SNTP Global Settings** page to enable or disable the SNTP client.

SNTP Global Settings

Use the **SNTP Global Settings** page to enable or disable the SNTP client, configure whether and how often the client sends SNTP requests, and determine whether the switch can receive SNTP broadcasts.

To display the SNTP Global Settings page, click **System** → **Time Synchronization** → **SNTP Global Settings** in the navigation panel.

Figure 12-8. SNTP Global Settings



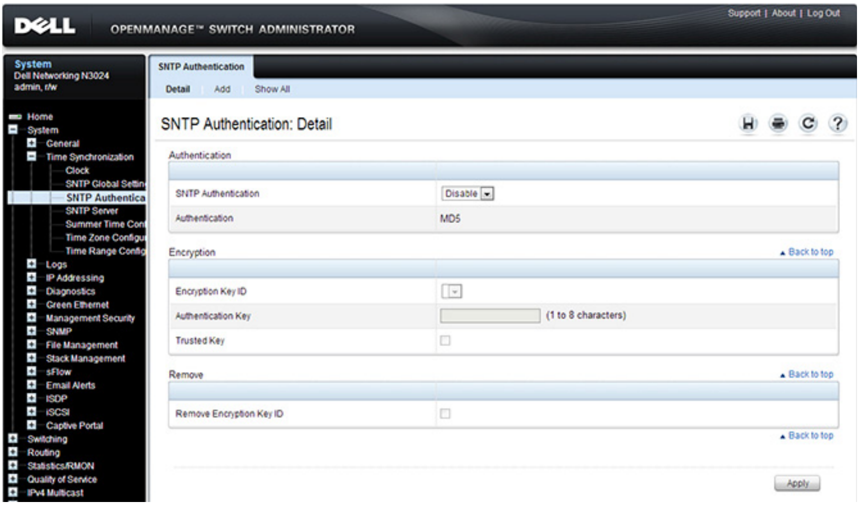
SNTP Authentication

Use the **SNTP Authentication** page to enable or disable SNTP authentication, to modify the authentication key for a selected encryption key ID, to designate the selected authentication key as a trusted key, and to remove the selected encryption key ID.

NOTE: The SNTP server must be configured with the same authentication information to allow time synchronization to take place between the two devices.

Click **System** → **Time Synchronization** → **SNTP Authentication** in the navigation panel to display the **SNTP Authentication** page.

Figure 12-9. SNTP Authentication



Adding an SNTP Authentication Key

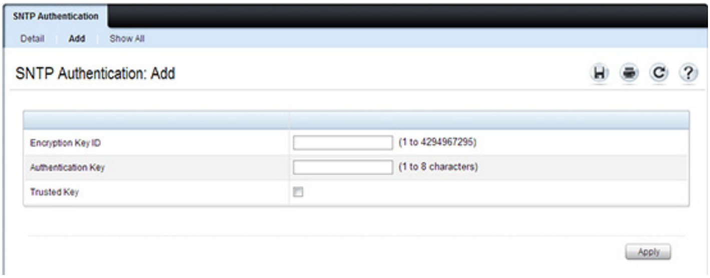
To configure SNTP authentication:

1 Open the **SNTP Authentication** page.

2 Click the **Add** link.

The **Add Authentication Key** page displays:

Figure 12-10. Add Authentication Key



3 Enter a numerical encryption key ID and an authentication key in the appropriate fields.

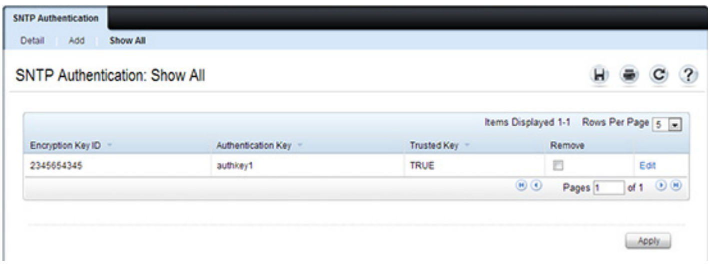
4 If the key is to be used to authenticate a unicast SNTP server, select the **Trusted Key** check box. If the check box is clear, the key is untrusted and cannot be used for authentication.

5 Click **Apply**.

The SNTP authentication key is added, and the device is updated.

To view all configured authentication keys, click the **Show All** link. The **Authentication Key Table** displays. The **Authentication Key Table** can also be used to remove or edit existing keys.

Figure 12-11. Authentication Key Table

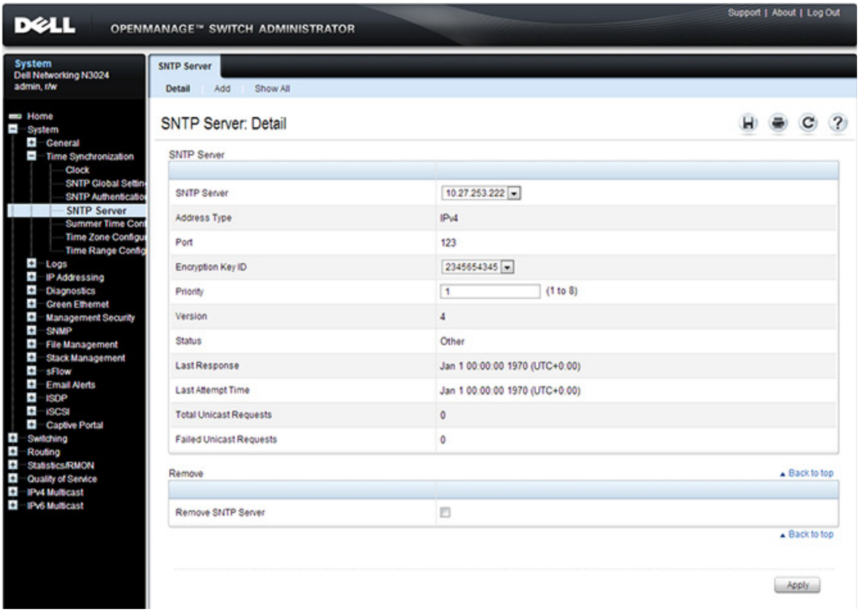


SNTP Server

Use the **SNTP Server** page to view and modify information about SNTP servers, and to add new SNTP servers that the switch can use for time synchronization. The switch can accept time information from both IPv4 and IPv6 SNTP servers.

To display the **SNTP Server** page, click **System** → **Time Synchronization** → **SNTP Server** in the navigation panel. If no servers have been configured, the fields in the following image are not displayed.

Figure 12-12. SNTP Servers



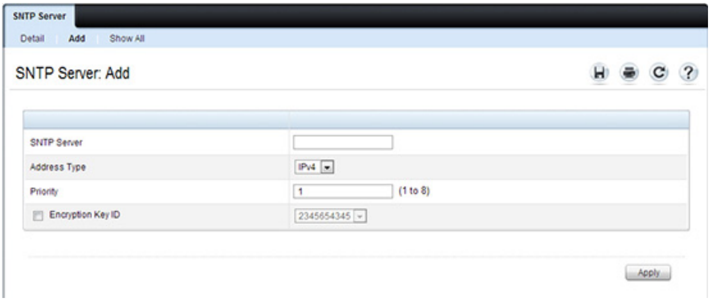
Defining a New SNTP Server

To add an SNTP server:

1. Open the **SNTP Servers** page.
2. Click **Add**.

   The **Add SNTP Server** page displays.

Figure 12-13. Add SNTP Server



3 In the **SNTP Server** field, enter the IP address or host name for the new SNTP server.

4 Specify whether the information entered in the **SNTP Server** field is an IPv4 address, IPv6 address, or a hostname (DNS).

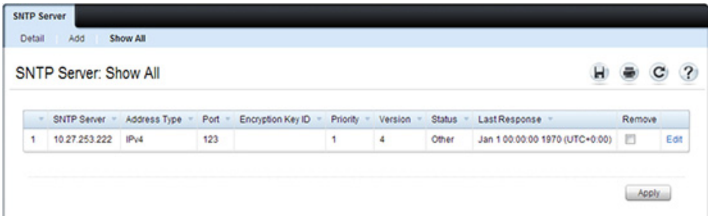
5 If authentication is required between the SNTP client on the switch and the SNTP server, select the Encryption Key ID check box, and then select the key ID to use.

To define a new encryption key, see "Adding an SNTP Authentication Key " on page 406.

NOTE: The SNTP server must be configured with the same authentication information to allow time synchronization to take place between the two devices.

To view all configured SNTP servers, click the **Show All** link. The SNTP **Server Table** displays. The SNTP **Server Table** page can also be used to remove or edit existing SNTP servers.

Figure 12-14. SNTP Servers Table

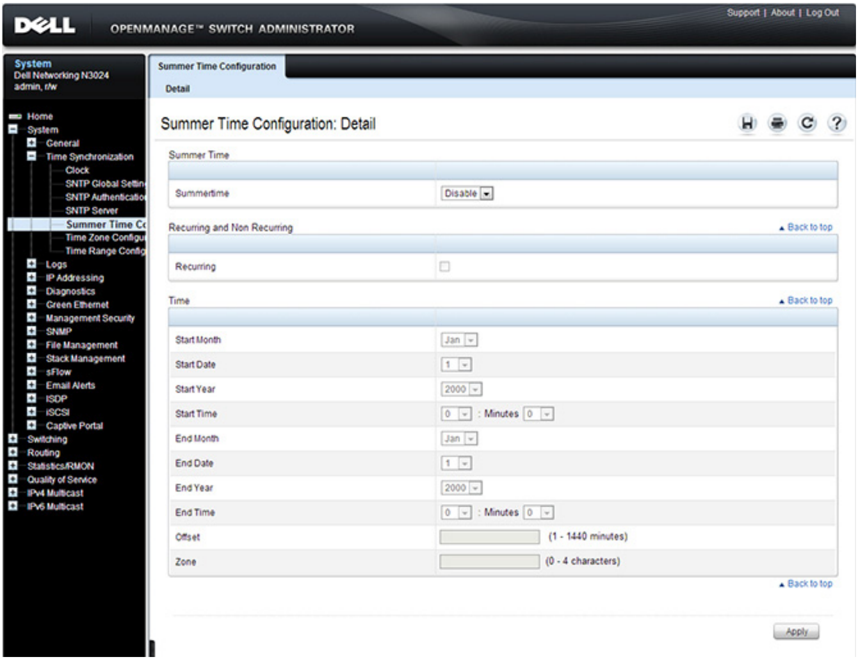


Summer Time Configuration

Use the **Summer Time Configuration** page to configure summer time (daylight saving time) settings.

To display the **Summer Time Configuration** page, click **System** → **Time Synchronization** → **Summer Time Configuration** in the navigation panel.

Figure 12-15. Summer Time Configuration



NOTE: The fields on the **Summer Time Configuration** page change when the Recurring check box is selected or cleared.

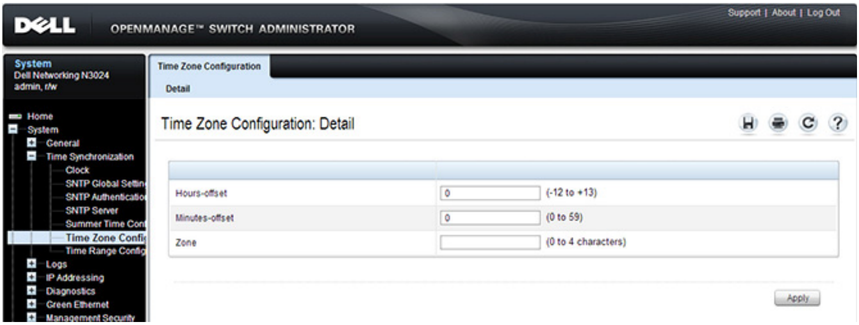
To use the pre-configured summer time settings for the United States or European Union, select the **Recurring** check box and specify USA or EU from the Location menu.

Time Zone Configuration

Use the **Time Zone Configuration** to configure time zone information, including the amount time the local time is offset from UTC and the acronym that represents the local time zone.

To display the **Time Zone Configuration** page, click **System** → **Time Synchronization** → **Time Zone Configuration** in the navigation panel.

Figure 12-16. Time Zone Configuration

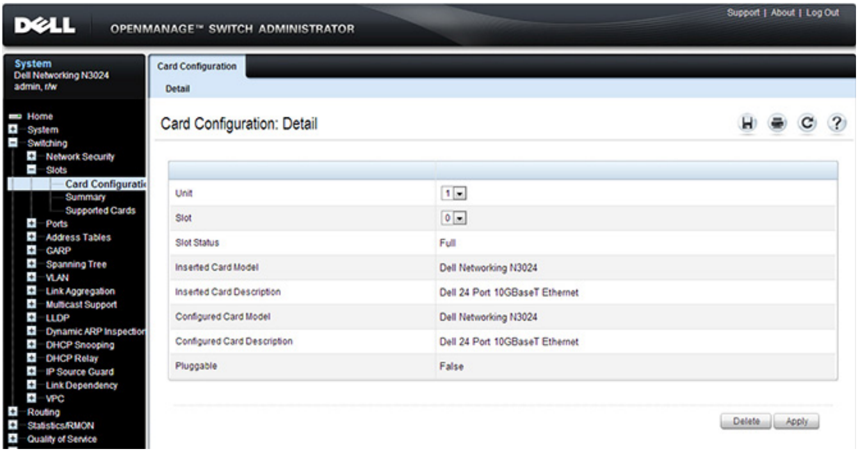


Card Configuration

Use the **Card Configuration** page to control the administrative status of the rear-panel expansion slots (Slot 1 or Slot 2) and to configure the plug-in module to use in the slot.

To display the **Card Configuration** page, click **Switching** → **Slots** → **Card Configuration** in the navigation panel.

Figure 12-17. Card Configuration

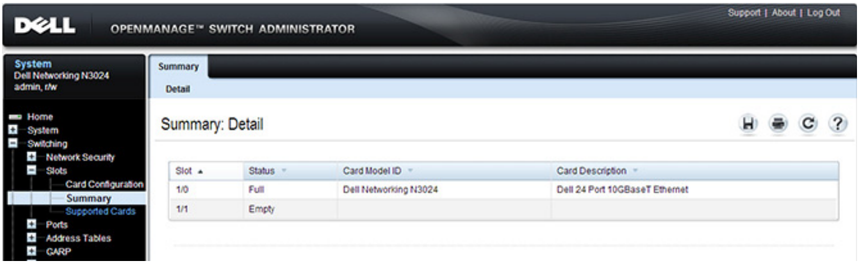


Slot Summary

Use the **Slot Summary** page to view information about the expansion slot status.

To display the **Slot Summary** page, click **Switching** → **Slots** → **Summary** in the navigation panel.

Figure 12-18. Slot Summary

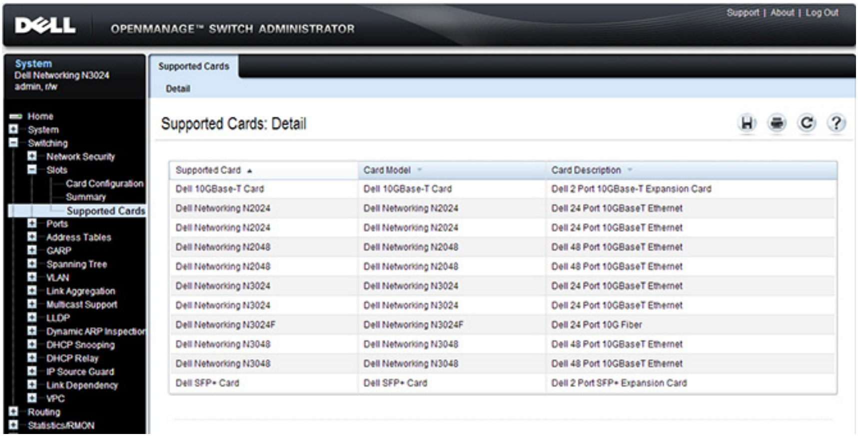


Supported Cards

Use the **Supported Cards** page to view information about the supported plug-in modules for the switch.

To display the **Supported Cards** page, click **Switching** → **Slots** → **Supported Cards** in the navigation panel.

Figure 12-19. Supported Cards

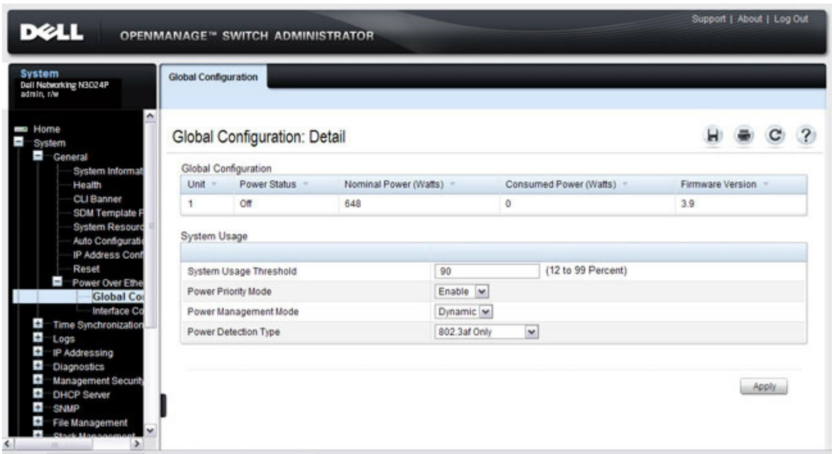


Power Over Ethernet Global Configuration (Dell Networking N1524P/N1548P, N2024P/N2048P, and N3024P/N3048P Only)

Use the PoE **Global Configuration** page to configure the PoE settings for the switch.

To display the **PoE Global Configuration** page, click **System** → **General** → **Power over Ethernet** → **Global Configuration** in the navigation panel.

Figure 12-20. PoE Global Configuration

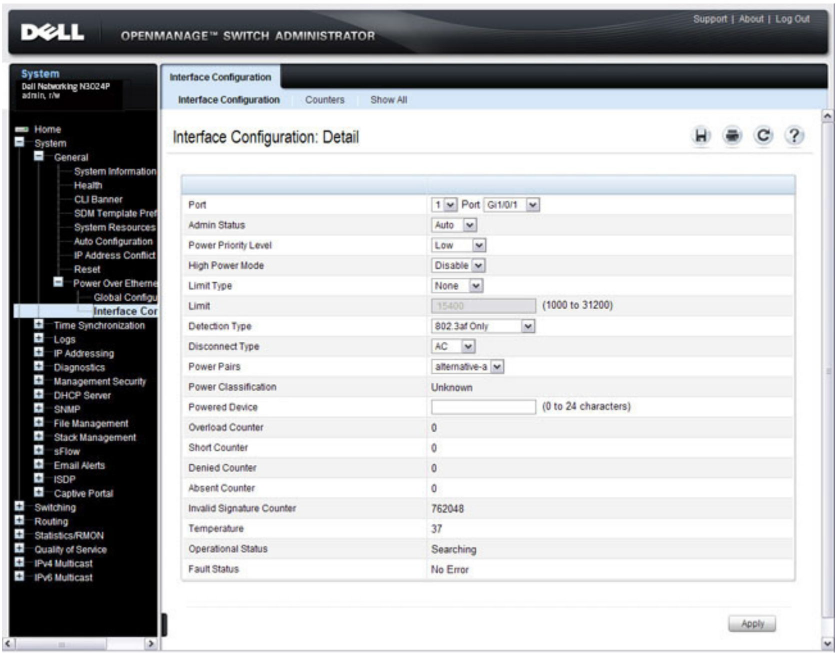


Power Over Ethernet Interface Configuration (Dell Networking N1524P/N1548P, N2024P/N2048P, and N3024P/N3048P Only)

Use the PoE **Interface Configuration** page to configure the per-port PoE settings. This page also provides access to the PoE Counters table and PoE Port Table. The PoE Port table allows viewing and configuring PoE settings for multiple ports on the same page.

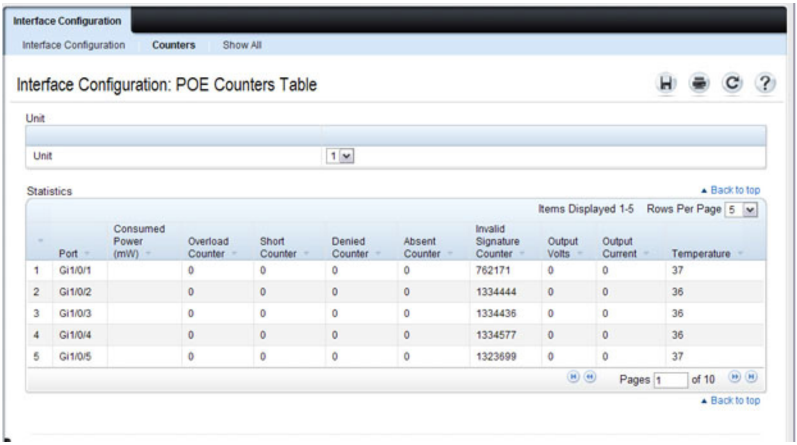
To display the PoE **Interface Configuration** page, click **System** → **General** → **Power over Ethernet** → **Interface Configuration** in the navigation panel.

Figure 12-21. PoE Interface Configuration



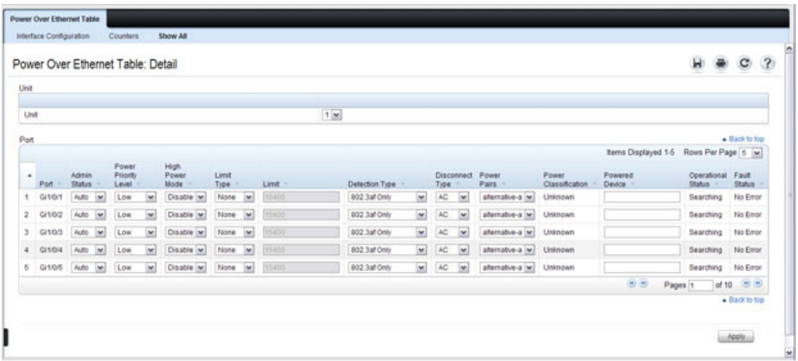
To view PoE statistics for each port, click **Counters.**

Figure 12-22. PoE Counters Table



To view the PoE Port Table, click **Show All**.

Figure 12-23. PoE Port Table



If you change any settings for one or more ports on the **PoE Port Table** page, click **Apply** to update the switch with the new settings.

Configuring System Settings (CLI)

This section provides information about the commands used for configuring system information and time settings on the Dell Networking N1500, N2000, N3000, and N4000 Series switches. For more information about these commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring System Information

Beginning in Privileged EXEC mode, use the following commands to configure system information.

Command	Purpose
configure	Enter Global Configuration mode.
hostname *name*	Configure the system name. The CLI prompt changes to the host name after you execute the command. The hostname is advertised in the LLDP system-name TLV.
snmp-server contact *name*	Configure the name of the switch administrator. If the name contains a space, use quotation marks around the name.
snmp-server location *location*	Configure the switch location.
asset-tag [**unit** *unit_id*] ***tag***	Configure the asset tag for the switch. Use the **unit** keyword to configure the asset tag for each unit in a stack of switches.
CTRL + Z	Exit to Privileged EXEC mode.
show system [**id**]	Display system information. Include the **id** keyword to display additional system information.

Configuring the Banner

Beginning in Privileged EXEC mode, use the following commands to configure the MOTD, login, or User EXEC banner. The switch supports the following banner messages:

- MOTD—Displays when a user connects to the switch.
- Login—Displays after the MOTD banner and before the login prompt.
- Exec—Displays immediately after the user logs on to the switch.

Command	Purpose
configure	Enter Global Configuration mode.
banner {**motd**|**login**|**exec**} *text*	Configure the banner message that displays when you connect to the switch (motd and login) or enter User EXEC mode (exec). Use quotation marks around a message if it includes spaces.
line {**telnet**|**ssh**|**console**}	Enter the terminal line configuration mode for Telnet, SSH, or the console.
motd-banner	Specify that the configured MOTD banner displays. To prevent the banner from displaying, enter **no motd-banner**.
exec-banner	Specify that the configured exec banner displays. To prevent the banner from displaying, enter **no exec-banner**.
login-banner	Specify that the configured login banner displays. To prevent the banner from displaying, enter **no login-banner**.
CTRL + Z	Exit to Privileged EXEC mode.
show banner	Display the banner status on all line terminals.

Managing the SDM Template

Beginning in Privileged EXEC mode, use the following commands to set the SDM template preference and to view information about the available SDM templates.

Command	Purpose
configure	Enter Global Configuration mode.
sdm prefer {dual-ipv4-and-ipv6 default| ipv4-routing {data-center | default}}	Select the SDM template to apply to the switch after the next boot.
CTRL + Z	Exit to Privileged EXEC mode.
show sdm prefer [*template*]	View information about the SDM template the switch is currently using. Use the *template* variable to view the parameters for the specified template.

Configuring SNTP Authentication and an SNTP Server

Beginning in Privileged EXEC mode, use the following commands to require the SNTP client to use authentication when communicating with the SNTP server. The commands also show how to configure an SNTP server.

Requiring authentication is optional. However, if you configure authentication on the switch SNTP client, the SNTP server must be configured with the same authentication information to allow time synchronization to take place between the two devices.

Command	Purpose
configure	Enter Global Configuration mode.
sntp authentication-key *key_id* **md5** *key_word*	Define an authentication key for SNTP. The variables are: • *key_id*— The encryption key ID, which is a number from 1–4294967295. • *key_word*—The authentication key, which is a string of up to eight characters.

Command	Purpose
sntp trusted-key *key_id*	Specify the authentication key the SNTP server must include in SNTP packets that it sends to the switch. The *key_id* number must be an encryption key ID defined in the previous step.
sntp authenticate	Require authentication for communication with the SNTP server. A trusted key must be configured before this command is executed.
sntp server {*ip_address* | *hostname*} [**priority** *priority*] [**key** *key_id*]	Define the SNTP server. • *ip_address*—The IP address (or host name) of the SNTP server to poll. The IP address can be an IPv4 or IPv6 address. • *priority*—(Optional) If multiple SNTP servers are defined, this number determines which server the switch polls first. The priority is 1–8, where 1 is the highest priority. If you do not specify a priority, the servers are polled in the order that they are entered. • *key_id*—(Optional) Enter an authentication key to use. The key must be previously defined by the **sntp authentication-key** command.
sntp {unicast|broadcast} client enable	This command enables the SNTP client and allows the switch to poll configured unicast SNTP servers for updates or receive broadcasts from any SNTP server.
sntp client poll timer *seconds*	Specify how often the SNTP client requests SNTP packets from the configured server(s). *seconds*—The poll interval can be 64, 128, 256, 512, or 1024 seconds.
CTRL + Z	Exit to Privileged EXEC mode.
show sntp configuration	Verify the SNTP configuration.
show sntp status	View information about the SNTP updates.

Setting the System Time and Date Manually

Beginning in Privileged EXEC mode, use the following commands to configure the time and date, time zone, and summer time settings.

Command	Purpose
clock set {*mm/dd/yyyy hh:mm:ss*} | {*hh:mm:ss mm/dd/yyyy*	Configure the time and date. Enter the time first and then the date, or the date and then the time. • *hh:mm:ss* —Time in hours (24-hour format, from 01-24), minutes (00-59), and seconds (00-59). • *mm/dd/yyyy* — Two digit month (1-12), two-digit date of the month (01-31), and four-digit year.
clock timezone hours-offset *hours-offset* [**minutes** *minutes-offset*] [**zone** *acronym*]	Configure the time zone settings. • *hours-offset* — Hours difference from UTC. (Range: –12 to +13) • *minutes-offset* — Minutes difference from UTC. (Range: 0–59) • *acronym* — The acronym for the time zone. (Range: Up to four characters)
clock summer-time recurring {**usa** | **eu** | {*week day month hh:mm week day month hh:mm*}} [**offset** *offset*] [**zone** *acronym*]	Use this command if the summer time starts and ends every year based on a set pattern. For switches located in the United States or European Union, use the **usa** or **eu** keywords to use the pre-configured values. Otherwise, configure the start and end times by using the following values: • *week* — Week of the month. (Range: 1–5, first, last) • *day* — Day of the week. (The first three letters by name) • *month* — Month. (The first three letters by name; jan, for example.) • *hh:mm* — Time in 24-hour format in hours and minutes. (Range: hh: 0–23, mm: 0–59) • *offset* — Number of minutes to add during the summertime. (Range:1–1440) • *acronym* — The acronym for the time zone to be displayed when summertime is in effect. (Up to four characters)

Command	Purpose
clock summer-time date {*date month* | *month date*} *year* *hh:mm* {*date month* | *month date*} *year* *hh:mm* [**offset** *offset*] [**zone** *acronym*]	Use this command if the summer time does not start and end every year according to a recurring pattern. Enter the month and then the date, or the date and then the month. • *date*— Day of the month. (Range: 1-31.) • *month* — Month. (Range: The first three letters by name) • *hh:mm* — Time in 24-hour format in hours and minutes. (Range: hh: 0–23, mm: 0–59) • *offset* — Number of minutes to add during the summertime. (Range:1–1440) • *acronym* — The acronym for the time zone to be displayed when summertime is in effect. (Range: Up to four characters)
CTRL + Z	Exit to Privileged EXEC mode.
show clock [**detail**]	View information about the time. Include the **detail** keyword to view information about the time zone and summer time.

Configuring the Expansion Slots (Dell Networking N3000 Series Only)

Beginning in Privileged EXEC mode, use the following commands to configure and view information about the expansion slots and plug-in modules (cards).

Command	Purpose
configure	Enter Global Configuration mode.
slot *unit/slot cardindex*	Configured the specified slot (1–2) to use the plug-in module identified by the *cardindex* number (CID). To view the CID associated with each plug-in module, use the **show supported cardtype** command.
CTRL + Z	Exit to Privileged EXEC mode.
show slot	Display status information about the expansion slots.
show supported cardtype	Display information about the plug-in modules the switch supports.

Viewing Slot Information (Dell Networking N4000 Series Only)

Use the following commands to view information about Slot 0 and its support.

Command	Purpose
show slot	Display status information about the expansion slots.
show supported cardtype	Display information about the modules the switch supports.

Configuring PoE Settings (Dell Networking N1524P/N1548P, N2024P/N2048P, and N3024P/N3048P Only)

Beginning in Privileged EXEC mode, use the following commands to configure PoE information.

Command	Purpose
configure	Enter Global Configuration mode.
power inline usage-threshold *threshold*	Specify the maximum usage for PoE power on the system. The *threshold* variable (range: 1–99%) is a percentage of total system power.
power inline management {class | static | dynamic}	Set the power-management mode for the switch.
power inline detection {dot3at | dot3at+legacy}	Set the power-management mode for the switch. • **802.3at-only**—IEEE 802.3at detection scheme is used. • **802.3at+legacy**—IEEE 802.3at 4point detection scheme is used and when it fails to detect a connected PD, legacy capacitive detection is used.
interface *interface*	Enter interface configuration mode for the specified port. The *interface* variable includes the interface type and number, for example **gigabitethernet 1/0/3**.

Command	Purpose
power inline {auto | never}	Set the PoE device discovery admin mode. • **auto** — Enables the device discovery protocol and, if found, supplies power to the device. • **never** — Disables the device discovery protocol and stops supplying power to the device.
power inline priority {critical | high | low}	Configures the port priority level for the delivery of power to an attached device.
power inline high-power	Configure the port high-power mode for connected-device compatibility. This setting is enabled by default.
power inline four-pair	Enable power feed on all pairs. High-power mode must also be enabled on the port.
power inline powered-device *type*	Provide a description to represent the type of device connected to the port.
power inline reset	(Optional) Reset the port. You might use this command if the port is stuck in an Error state.
CTRL + Z	Exit to Privileged EXEC mode.
show power inline	Display PoE information for the switch.
show power inline *interface*	Display PoE information for the specified interface.

General System Settings Configuration Examples

This section contains the following examples:

- Configuring System and Banner Information
- Configuring SNTP
- Configuring the Time Manually

Configuring System and Banner Information

In this example, an administrator configures the following system information:

- System name: N2048
- System contact: Jane Doe
- System location: RTP100
- Asset tag: 006429

The administrator then configures the MOTD banner to alert other switch administrators of the connected topology.

To configure the switch:

1 Configure the hosts name.

```
console#configure
console(config)#hostname N2048
```

2 Configure the contact, location, and asset tag. Notice that the prompt changed to the host name.

```
N2048(config)#snmp-server contact "Jane Doe"
N2048(config)#snmp-server location RTP100
N2048(config)#asset-tag 006429
```

3 Configure the message that displays when a user connects to the switch.

```
N2048(config)#banner motd "This switch connects users in
cubicles C121-C139."
N2048(config)#exit
```

4 View system information to verify the configuration.

```
N2048#show system
System Description: Dell Ethernet Switch
System Up Time: 0 days, 19h:36m:36s
System Contact: Jane Doe
System Name: N2048
```

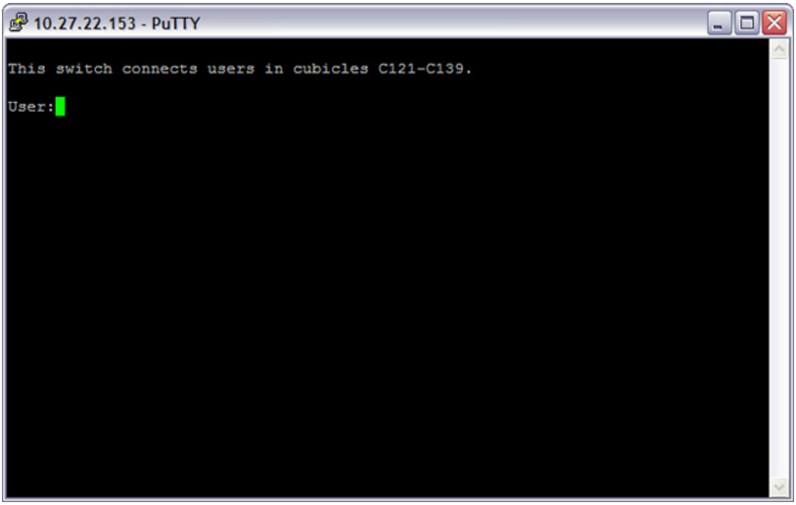
```
System Location: RTP100
Burned In MAC Address: 001E.C9AA.AA07
System Object ID: 1.3.6.1.4.1.674.10895.3035
System Model ID: N2048
Machine Type: Dell Networking N2048
Temperature Sensors:

Unit     Temperature (Celsius)    Status
----     ---------------------    ------
 1        43                      OK
Power Supplies:
Unit     Description    Status         Source
----     -----------    -----------    ------
 1       Main           OK             AC
 1       Secondary      Error          DC
```

5 View additional information about the system.

```
N2048#show system id
Service Tag:
Chassis Service Tag: N/A
Serial Number: 7048NX1011
Asset Tag: unit-1
Unit Service tag  Chassis Serv tag  Serial number  Asset tag
---- -----------  ----------------  -------------  ---------
1                 N/A               70498NX1011      unit-1
```

6 Initiate a new Telnet session to verify the MOTD.

Figure 12-24. Verify MOTD



Configuring SNTP

The commands in this example configure the switch to poll an SNTP server to synchronize the time. Additionally, the SNTP sessions between the client and server must be authenticated.

To configure the switch:

1. Configure the authentication information. The SNTP server must be configured with the same authentication key and ID.

```
console#configure
console(config)#sntp authentication-key 23456465 md5 sntpkey
console(config)#sntp trusted-key 23456465
console(config)#sntp authenticate
```

2. Specify the IP address of the SNTP server to poll and include the authentication key. This command automatically enables polling and sets the priority to 1.

```
console(config)#sntp server 192.168.10.30 key 23456465
console(config)#sntp unicast client enable
```

3. Verify the configuration.

```
console#show sntp configuration

Polling interval: 512 seconds
MD5 Authentication keys: 23456465
Authentication is required for synchronization.
Trusted keys: 23456465
Unicast clients: Enable

Unicast servers:
Server          Key         Polling   Priority
------------    ----------- --------- --------
192.168.10.30   23456465    Enabled   1
```

4 View the SNTP status on the switch.

```
console#show sntp status

Client Mode:          Unicast
Last Update Time:     MAR 01 09:12:43 2010

Unicast servers:
Server          Status       Last response
--------------- ------------ ---------------------
192.168.10.30   Other        09:12:43 Mar 1 2011
```

Configuring the Time Manually

The commands in this example manually set the system time and date. The time zone is set to Eastern Standard Time (EST), which has an offset of -5 hours. Summer time is enabled and uses the pre-configured United States settings.

To configure the switch:

1 Configure the time zone offset and acronym.

```
console#configure
console(config)#clock timezone -5 zone EST
```

2 Configure the summer time (daylight saving time) to use the pre-configured settings for the United States.

```
console(config)#clock summer-time recurring us
```

3 Set the local time and date.

```
console(config)#clock set 16:13.06 03/01/2010
```

4 Verify the time settings.

```
console#show clock detail

00:27:19 EST(UTC-5:00) Feb 3 2039
No time source

 Time zone:
 Acronym is EST
 Offset is UTC-5:00

 Summertime:
 Acronym not configured
 Recurring every year (USA)
 Begins on second Sunday of Mar at 02:00
 Ends on first Sunday of Nov at 02:00
 Offset is +60 minutes
```

13

SNMP

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

The topics covered in this chapter include:

- SNMP Overview
- Default SNMP Values
- Configuring SNMP (Web)
- Configuring SNMP (CLI)
- SNMP Configuration Examples

SNMP Overview

Simple Network Management Protocol (SNMP) provides a method for managing network devices. The Dell Networking N-Series switches support SNMP version 1, SNMP version 2, and SNMP version 3. Dell Networking switches support SNMP over IPv4 only.

What Is SNMP?

SNMP is a standard protocol that enables remote monitoring and management of a device through communication between an SNMP manager and an SNMP agent on the remote device. The SNMP manager is typically part of a Network Management System (NMS) that runs on an administrative host. The switch software includes Management Information Base (MIB) objects that the SNMP agent queries and modifies. The switch uses standard public MIBs and private MIBs.

A MIB acts as a structured road map for managed objects. A managed object is any feature or setting that can be configured or monitored on the switch. An Object Identifier (OID) is the unique number assigned to an object defined in a MIB. An OID is written as a sequence of subidentifiers in decimal notation.

The SNMP agent maintains a list of variables that are used to manage the switch. The variables are defined in the MIB. The MIB presents the variables controlled by the agent. The SNMP agent defines the MIB specification format, as well as the format used to access the information over the network. Access rights to the SNMP agent are controlled by access strings.

SNMP v3 also applies access control and a new traps mechanism to SNMPv1 and SNMPv2 PDUs. In addition, the User Security Model (USM) is defined for SNMPv3 and includes:

- Authentication — Provides data integrity and data origin authentication. Both MD5 and SHA authentication methods are supported.
- Privacy — Protects against disclosure of message content. DES, 3DES, or AES 128 Cipher-Bock-Chaining (CBC) is used for encryption. Either authentication is enabled on an SNMP message, or both authentication and privacy are enabled on an SNMP message. However privacy cannot be enabled without authentication.
- Timeliness — Protects against message delay or message redundancy. The SNMP agent compares incoming message to the message time information.
- Key Management — Defines key generation, key updates, and key use.

Authentication or Privacy Keys are modified in the SNMPv3 User Security Model (USM).

What Are SNMP Traps?

SNMP is frequently used to monitor systems for fault conditions such as temperature violations, link failures, and so on. Management applications can monitor for these conditions by polling the appropriate OIDs with the **get** command and analyzing the returned data. This method has its drawbacks. If it is done frequently, significant amounts of network bandwidth can be consumed. If it is done infrequently, the response to the fault condition may not occur in a timely fashion. SNMP traps avoid these limitations of the polling method.

An SNMP trap is an asynchronous event indicating that something significant has occurred. This is analogous to a pager receiving an important message, except that he SNMP trap frequently contains all the information needed to diagnose a fault.

Various features can be configured on the switch to generate SNMP traps that inform the NMS about events or problems that occur on the switch. Traps generated by the switch can also be viewed locally by using the web-based interface or CLI.

Why Is SNMP Needed?

Some network administrators prefer to use SNMP as the switch management interface. Settings that you view and configure by using the web-based Dell OpenManage Switch Administrator and the CLI are also available by using SNMP.

If you do not use NMS software to manage or monitor other devices on your network, it might not be necessary to configure SNMP on the switch.

Default SNMP Values

By default, SNMPv2 is automatically enabled on the device. SNMPv1 and SNMPv3 are disabled. To enable SNMPv3, you must define a local engine ID for the device. The local engineID is by default set to the switch MAC address, however when the switch operates in a stacking mode, it is important to manually configure the local engineID for the stack. This local engineID must be defined so that it is unique within the network. It is important to do this because the default engineID in a stack is the MAC address of the master unit, which may change if the master unit fails and another unit takes over the stack.

Table 13-1 summarizes the default values for SNMP.

Table 13-1. SNMP Defaults

Parameter	Default Value
SNMPv1	Disabled
SNMPv2	Enabled
SNMPv3	Disabled
SNMP traps	Enabled
SNMP trap receiver	None configured
Switch traps	Enabled

Table 13-1. SNMP Defaults

Parameter	Default Value
QoS traps	Enabled
Multicast traps	Disabled
Captive Portal traps	Disabled
OSPF traps	Disabled

Table 13-2 describes the two views that are defined by default.

Table 13-2. SNMP Default Views

View Name	OID Subtree	View Type
Default	iso	Included
	snmpVacmMIB	Excluded
	usmUser	Excluded
	snmpCommunityTable	Excluded
DefaultSuper	iso	Included

By default, three groups are defined. Table 13-3 describes the groups. The Read, Write, and Notify values define the pre-configured views that are associated with the groups.

Table 13-3. SNMP Default Groups

Group Name	Security Level	Read	Write	Notify
DefaultRead	No Auth No Priv	Default	–	Default
DefaultWrite	No Auth No Priv	Default	Default	Default
DefaultSuper	No Auth No Priv	DefaultSuper	DefaultSuper	DefaultSuper

Configuring SNMP (Web)

This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring the SNMP agent on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

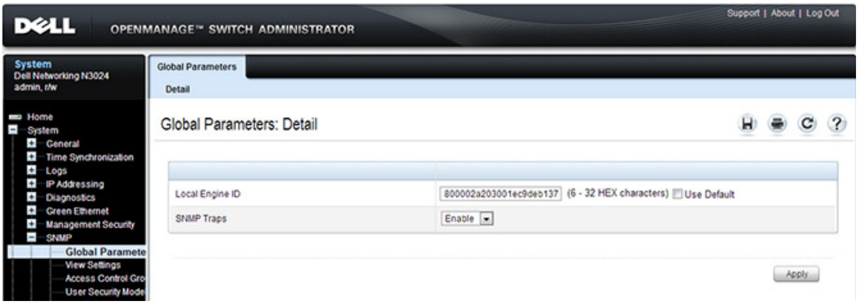
NOTE: For some features, the control to enable or disable traps is available from a configuration page for that feature and not from the **Trap Manager** pages that this chapter describes.

SNMP Global Parameters

Use the **Global Parameters** page to enable SNMP and Authentication notifications.

To display the **Global Parameters** page, click **System** → **SNMP** → **Global Parameters** in the navigation panel.

Figure 13-1. SNMP Global Parameters

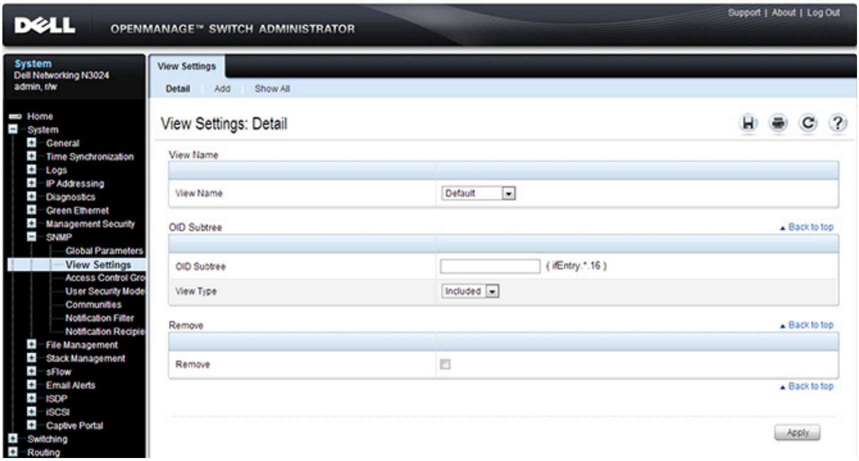


SNMP View Settings

Use the SNMP **View Settings** page to create views that define which features of the device are accessible and which are blocked. A view can be created that includes or excludes OIDs corresponding to interfaces.

To display the **View Settings** page, click **System** → **SNMP** → **View Settings** in the navigation panel.

Figure 13-2. SNMP View Settings



Adding an SNMP View

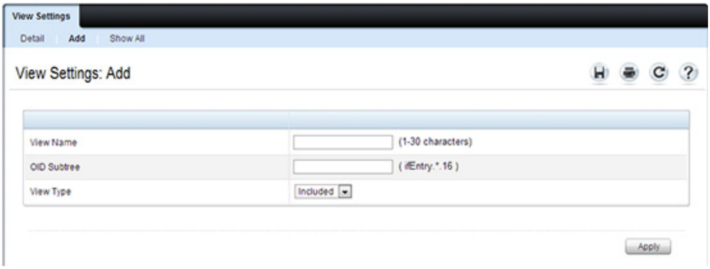
To add a view:

1 Open the **View Settings** page.

2 Click **Add**.

The **Add View** page displays:

Figure 13-3. Add View



3 Specify a name for the view and a valid SNMP OID string.
4 Select the view type.
5 Click **Apply**.
The SNMP view is added, and the device is updated.

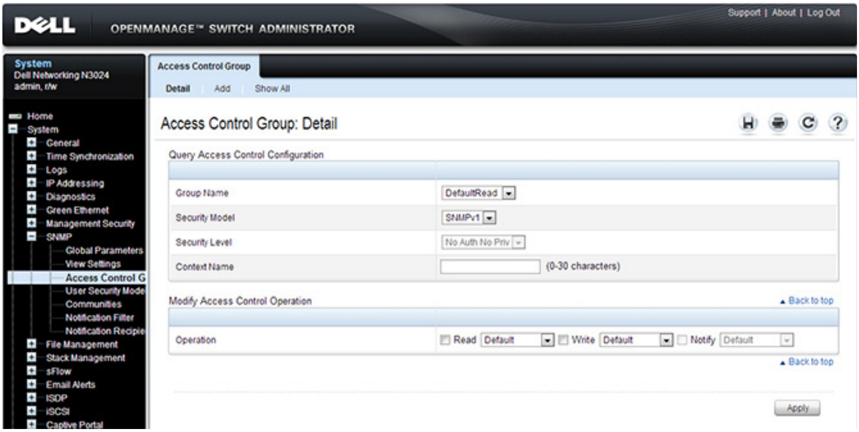
Click **Show All** to view information about configured SNMP Views.

Access Control Group

Use the **Access Control Group** page to view information for creating SNMP groups, and to assign SNMP access privileges. Groups allow network managers to assign access rights to specific device features or features aspects.

To display the **Access Control Group** page, click **System** → **SNMP** → **Access Control** in the navigation panel.

Figure 13-4. SNMP Access Control Group



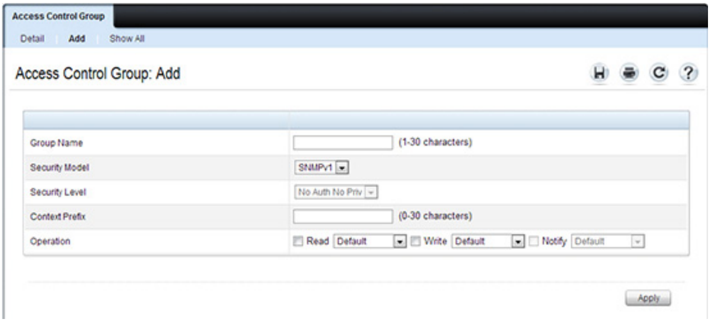
Adding an SNMP Group

To add a group:

1. Open the **Access Control Configuration** page.
2. Click **Add**.

   The **Add an Access Control Configuration** page displays:

Figure 13-5. Add Access Control Group



3 Specify a name for the group.

4 Select a security model and level

5 Define the context prefix and the operation.

6 Click **Apply** to update the switch.

Click **Show All** to view information about existing access control configurations.

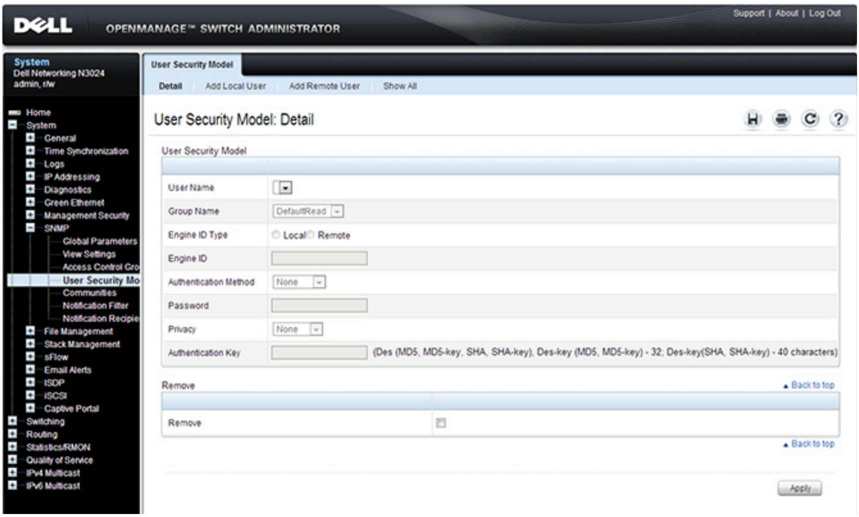
SNMPv3 User Security Model (USM)

Use the **User Security Model** page to assign system users to SNMP groups and to define the user authentication method.

NOTE: The Local User Database page under **Management Security** can also be used for configuring SNMPv3 settings for users. For more information, see "Authentication, Authorization, and Accounting " on page 229.

To display the **User Security Model** page, click **System** → **SNMP** → **User Security Model** in the navigation panel.

Figure 13-6. SNMPv3 User Security Model



Adding Local SNMPv3 Users to a USM

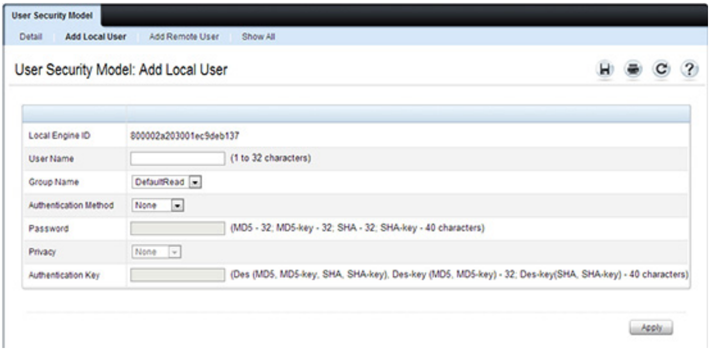
To add local users:

1 Open the **User Security Model** page.

2 Click **Add Local User**.

The **Add Local User** page displays:

Figure 13-7. Add Local Users



3 Define the relevant fields.

4 Click **Apply** to update the switch.

Click **Show All** to view the User Security Model Table, which contains information about configured Local and Remote Users.

Adding Remote SNMPv3 Users to a USM

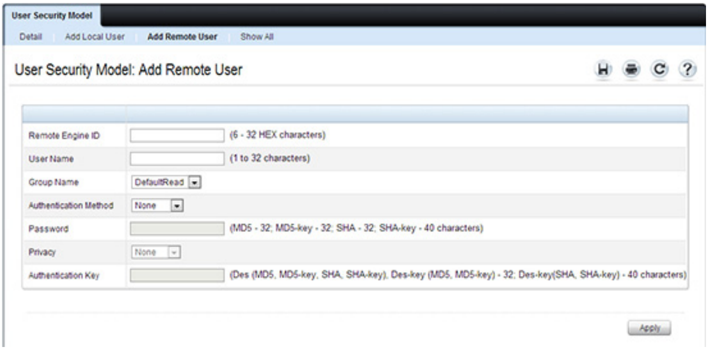
To add remote users:

1 Open the **SNMPv3 User Security Model** page.

2 Click **Add Remote User**.

The **Add Remote User** page displays:

Figure 13-8. Add Remote Users



3 Define the relevant fields.

4 Click **Apply** to update the switch.

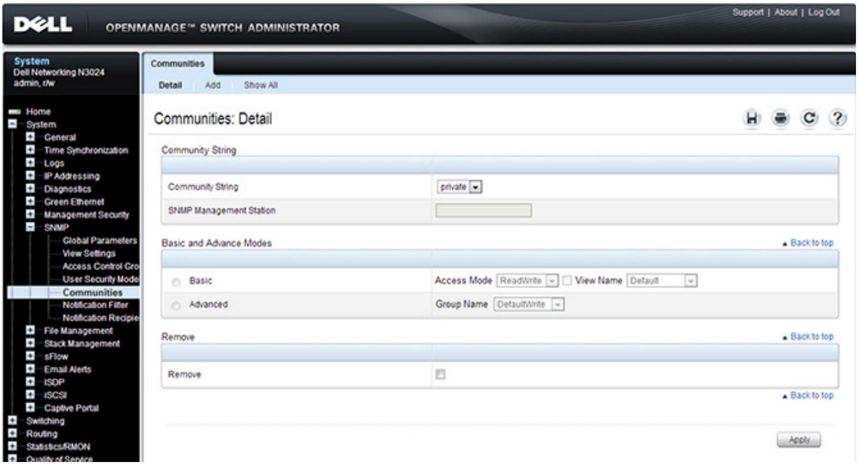
Click **Show All** to view the User Security Model Table, which contains information about configured Local and Remote Users.

Communities

Access rights for SNMPv1 and SNMPv2 are managed by defining communities **Communities** page. When the community names are changed, access rights are also changed. SNMP Communities are defined only for SNMP v1 and SNMP v2.

To display the **Communities** page, click **System** → **SNMP** → **Communities** in the navigation panel.

Figure 13-9. SNMP Communities



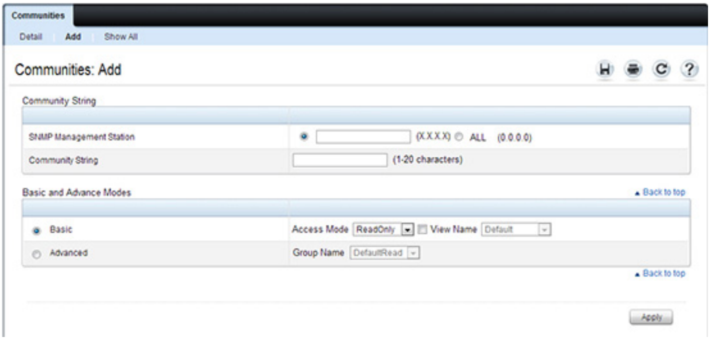
Adding SNMP Communities

To add a community:

1. Open the **Communities** page.
2. Click **Add**.

   The **Add SNMPv1,2 Community** page displays:

Figure 13-10. Add SNMPv1,2 Community



3 Specify the IP address of an SNMP management station and the community string to act as a password that will authenticate the management station to the SNMP agent on the switch.

4 Select the access mode.

5 Click **Apply** to update the switch.

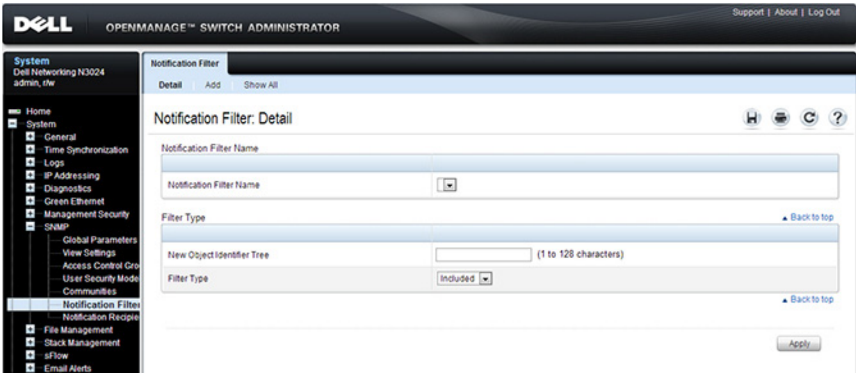
Click **Show All** to view the communities that have already been configured.

Notification Filter

Use the **Notification Filter** page to set filtering traps based on OIDs. Each OID is linked to a device feature or a feature aspect. The **Notification Filter** page also allows you to filter notifications.

To display the **Notification Filter** page, click **System** → **SNMP** → **Notification Filters** in the navigation panel.

Figure 13-11. SNMP Notification Filter



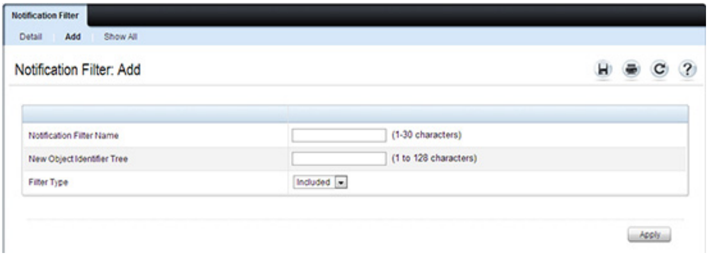
Adding a Notification Filter

To add a filter:

1. Open the **Notification Filter** page.
2. Click **Add**.

   The **Add Filter** page displays:

Figure 13-12. Add Notification Filter



3 Specify the name of the filter, the OID for the filter.

4 Choose whether to send (include) traps or informs to the trap recipient or prevent the switch from sending (exclude) the traps or informs.

5 Click **Apply** to update the switch.

Click **Show All** to view information about the filters that have already been configured.

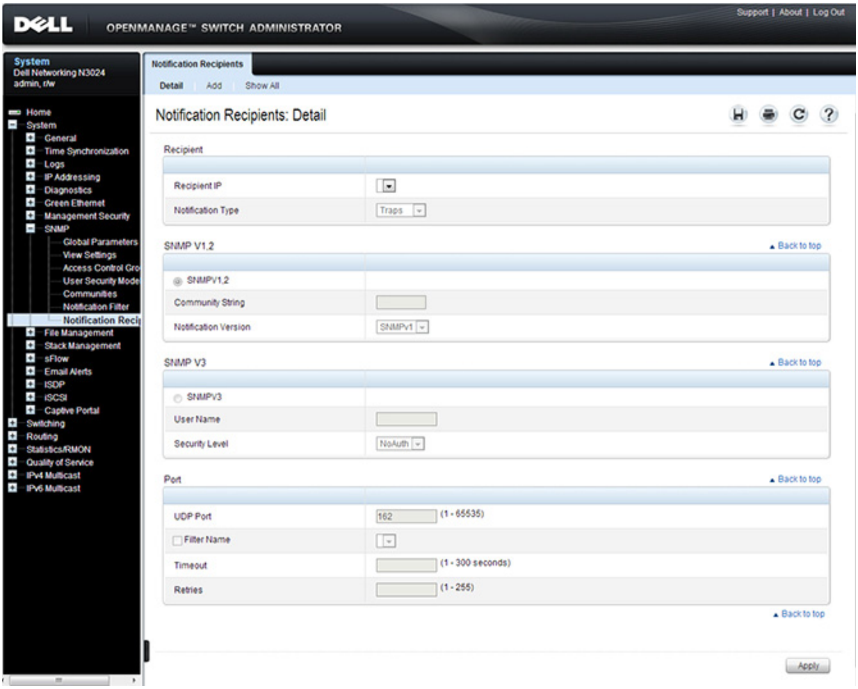
Notification Recipients

Use the **Notification Recipients** page to view information for defining filters that determine whether traps are sent to specific users, and the trap type sent. SNMP notification filters provide the following services:

- Identifying Management Trap Targets
- Trap Filtering
- Selecting Trap Generation Parameters
- Providing Access Control Checks

To display the **Notification Recipients** page, click **System** → **SNMP** → **Notification Recipient** in the navigation panel.

Figure 13-13. SNMP Notification Recipient



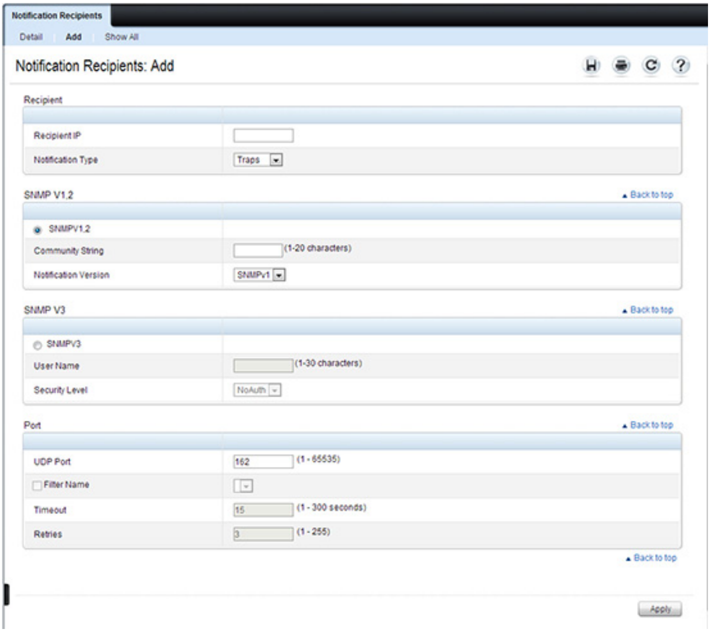
Adding a Notification Recipient

To add a recipient:

1. Open the **Notification Recipient** page.
2. Click **Add**.

   The **Add Recipient** page displays:

Figure 13-14. Add Notification Recipient



3 Specify the IP address or hostname of the host to receive notifications.

4 Select whether to send traps or informs to the specified recipient

5 Define the relevant fields for the SNMP version you use.

6 Configure information about the port on the recipient.

7 Click **Apply** to update the switch.

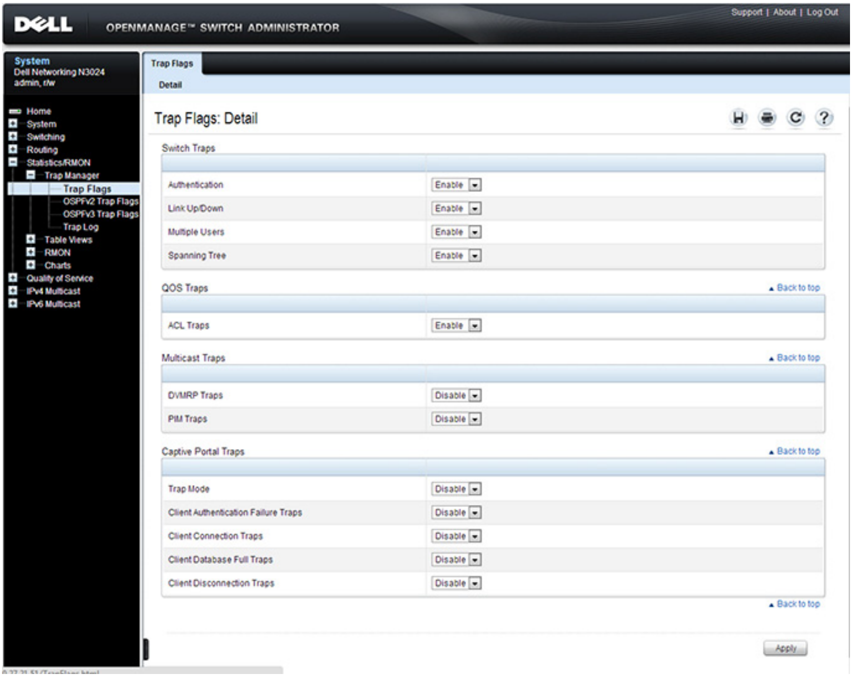
Click **Show All** to view information about the recipients that have already been configured.

Trap Flags

The **Trap Flags** page is used to specify which traps you want to enable or disable. When the condition identified by an active trap is encountered by the switch, a trap message is sent to any enabled SNMP Trap Receivers, and a message is written to the trap log.

To access the **Trap Flags** page, click **Statistics/RMON** → **Trap Manager** → **Trap Flags** in the navigation panel.

Figure 13-15. Trap Flags

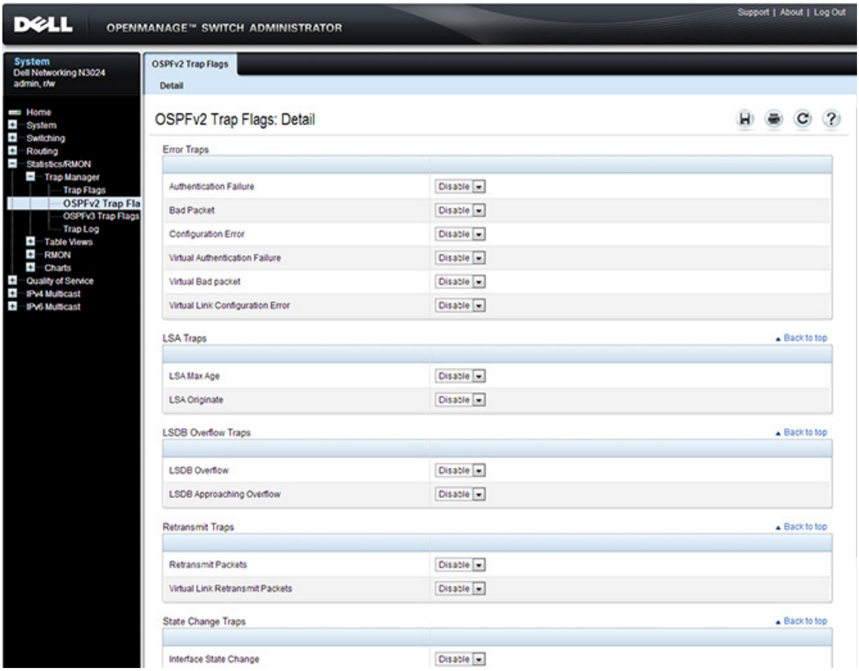


OSPFv2 Trap Flags

The **OSPFv2 Trap Flags** page is used to specify which OSPFv2 traps you want to enable or disable. When the condition identified by an active trap is encountered by the switch, a trap message is sent to any enabled SNMP Trap Receivers, and a message is written to the trap log.

To access the **OSPFv2 Trap Flags** page, click **Statistics/RMON** → **Trap Manager** → **OSPFv2 Trap Flags** in the navigation panel.

Figure 13-16. OSPFv2 Trap Flags

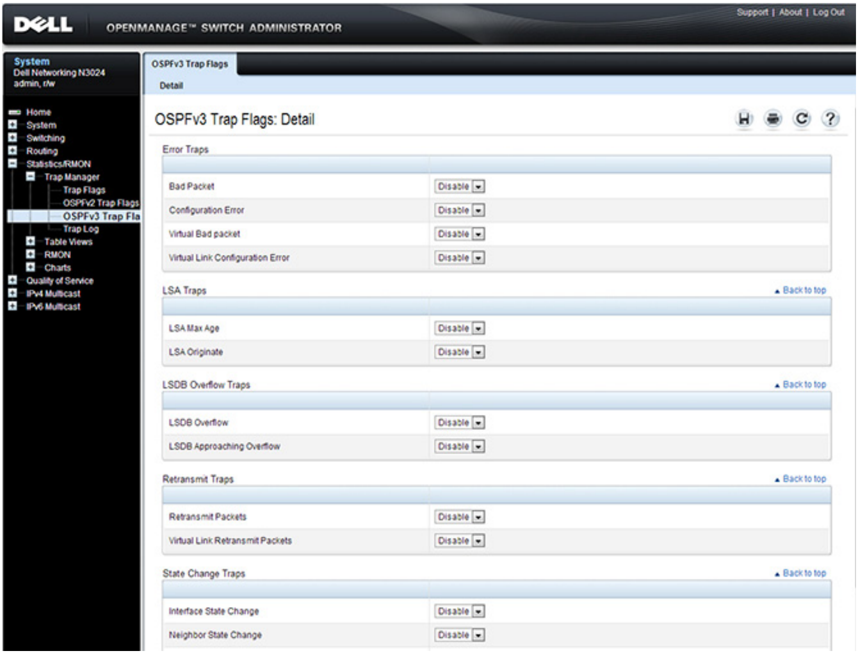


OSPFv3 Trap Flags

The **OSPFv3 Trap Flags** page is used to specify which OSPFv3 traps you want to enable or disable. When the condition identified by an active trap is encountered by the switch, a trap message is sent to any enabled SNMP Trap Receivers, and a message is written to the trap log.

To access the **OSPFv3 Trap Flags** page, click **Statistics/RMON** → **Trap Manager** → **OSPFv3 Trap Flags** in the navigation panel.

Figure 13-17. OSPFv3 Trap Flags

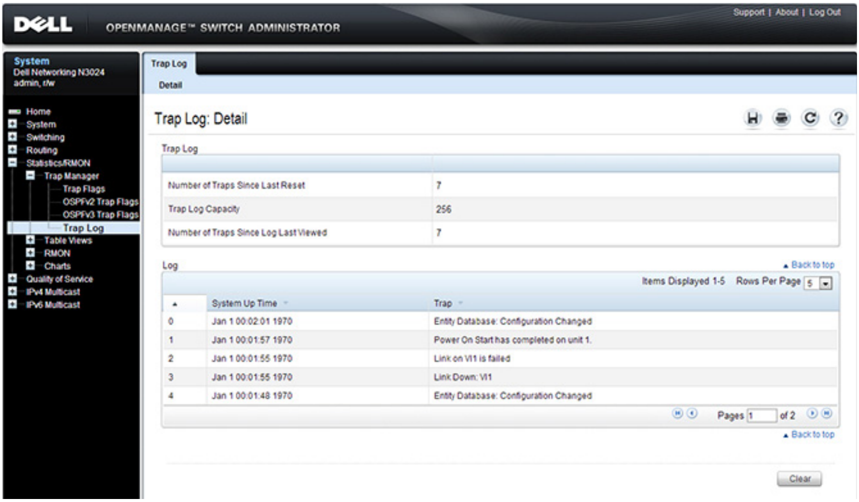


Trap Log

The **Trap Log** page is used to view entries that have been written to the trap log.

To access the **Trap Log** page, click **Statistics/RMON → Trap Manager → Trap Log** in the navigation panel.

Figure 13-18. Trap Logs



Click **Clear** to delete all entries from the trap log.

Configuring SNMP (CLI)

This section provides information about the commands you use to manage and view SNMP features on the switch. For more information about these commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring the SNMPv3 Engine ID

To use SNMPv3, the switch must have an engine ID configured. The default string that is generated using the MAC address of the switch can be used, or another value can be specified. If the SNMPv3 engine ID is deleted, or if the configuration file is erased, then SNMPv3 cannot be used. Since the EngineID should be unique within an administrative domain, Dell recommends that you use the default keyword to configure the Engine ID for stand-alone switches.

The following guidelines are recommended:

- For standalone switches use the default keyword to configure the Engine ID.
- For a stack of switches, configure your own EngineID, and verify that is unique within your administrative domain.

Changing the value of SNMP EngineID has important side-effects. A user's password (entered on the command line) is converted to an MD5 or SHA security digest. This digest is based on both the password and the local engine ID. The command line password is then destroyed, as required by RFC 2274. Because of this deletion, if the local value of engineID changes, the security digests of SNMPv3 users will be invalid, and the users will have to be reconfigured.

Beginning in Privileged EXEC mode, use the following commands to configure an engine ID for SNMP.

Command	Purpose
configure	Enter Global Configuration mode

Command	Purpose
snmp-server engineID local {*engineid-string* **| default}**	Configure the SNMPv3 Engine ID. • engineid-string — The character string that identifies the engine ID. The engine ID is a concatenated hexadecimal string. Each byte in hexadecimal character strings is two hexadecimal digits. Each byte can be separated by a period or colon. (Range: 6-32 characters) • default — The engineID is created automatically, based on the device MAC address.
exit	Exit to Privileged EXEC mode.
show snmp engineid	View the local SNMP engine ID.

Configuring SNMP Views, Groups, and Users

Beginning in Privileged EXEC mode, use the following commands to define SNMP views, and SNMP groups, and local and remote SNMPv3 users.

Command	Purpose
configure	Enter Global Configuration mode
snmp-server view *view-name oid-tree* **{included | excluded}**	Configure the SNMP view. When configuring groups, users, and communities, a view can be associated with the group, user, or community • *view-name* — Specifies the name of the view. (Range: 1-30 characters.) • *oid-tree* — Specifies the object identifier of the ASN.1 subtree to be included or excluded from the view. To identify the subtree, specify a text string consisting of numbers, such as `1.3.6.2.4`, or a word, such as `system`. Replace a single subidentifier with the asterisk (*) wildcard to specify a subtree family; for example `1.3.*.4`. • **included** — Indicates that the view type is included. • **excluded** — Indicates that the view type is excluded.

Command	Purpose
snmp-server group *groupname* {**v1** | **v2** | **v3** {**noauth** | **auth** | **priv**} [**notify** *view-name*]} [**context** *view-name*] [**read** *view-name*] [**write** *view-name*]	Specify the identity string of the receiver and set the receiver timeout value. • *groupname* — Specifies the name of the group. (Range: 1-30 characters.) • **v1** — Indicates the SNMP Version 1 security model. • **v2** — Indicates the SNMP Version 2 security model. • **v3** — Indicates the SNMP Version 3 security model. • **noauth** — Indicates no authentication of a packet. Applicable only to the SNMP Version 3 security model. • **auth** — Indicates authentication of a packet without encrypting it. Applicable only to the SNMP Version 3 security model. • **priv** — Indicates authentication of a packet with encryption. Applicable only to the SNMP Version 3 security model. • *view-name* — Specifies the view (defined in the previous step) to use for the context, notification, read, and write privileges for the group.

Command	Purpose
snmp-server user *username groupname* [**remote** *engineid-string*] [{**authmd5** *password* | **auth-sha** *password* | **auth-md5-key** *md5-key* | **auth-sha-key** *sha-key*} [**priv-des** *password* | **priv-des-key** *des-key* | **priv-3des** *password* | **priv-3des-key** *des-key* | **priv-aes128** *password* | **priv-aes128-key** *aes-key*]]	Configure a new SNMPv3 user. • *username* — Specifies the name of the user on the host that connects to the agent. (Range: 1-30 characters.) • *groupname* — Specifies the name of the group to which the user belongs. (Range: 1-30 characters.) • *engineid-string* — Specifies the engine ID of the remote SNMP entity to which the user belongs. The engine ID is a concatenated hexadecimal string. Each byte in the hexadecimal character string is two hexadecimal digits. The remote engine id designates the remote management station, and should be defined to enable the device to receive acknowledgements to "informs." (Range: 5-32 characters.) • **auth-md5** — The HMAC-MD5-96 authentication level. • **auth-sha** — The HMAC-SHA-96 authentication level. • *password* — A password. (Range: 1 to 32 characters.) • **auth-md5-key** — The HMAC-MD5-96 authentication level. Enter a pregenerated MD5 key. • **auth-sha-key** — The HMAC-SHA-96 authentication level. Enter a pregenerated SHA key. • *md5-key* — Character string—length 32 hex characters. • *sha-key* — Character string—length 48 characters. • *aes-key* — Character string—length 32 hex characters. • **priv-des** — The CBC-DES Symmetric Encryption privacy level. Enter a password. • **priv-des-key** — The CBC-DES Symmetric Encryption privacy level. The user must enter a pregenerated MD5 or SHA key depending on the authentication level selected. • **priv-3des** — The CBC-3DES Symmetric Encryption privacy level. Enter a password. • **priv-3des-key** — The CBC-3DES Symmetric Encryption privacy level. The user must enter a pregenerated MD5 or SHA key depending on the authentication level selected.

Command	Purpose
(continued)	• *des-key* — A pregenerated DES encryption key. Length is determined by authentication method selected: 32 hex characters if MD5 Authentication is selected, 48 hex characters if SHA Authentication is selected. • **priv-aes128** — The CBC-AES128 Symmetric Encryption privacy level. Enter a password. • **priv-aes128-key** — The CBC-AES128 Symmetric Encryption privacy level. The user must enter a pre-generated MD5 or SHA key depending on the authentication level selected.
exit	Exit to Privileged EXEC mode.
show snmp views	View SNMP view configuration information.
show snmp group [*group_name*]	View SNMP group configuration information.
show snmp user [*user_name*]	View SNMP user configuration information.

Configuring Communities

Beginning in Privileged EXEC mode, use the following commands to configure access rights for SNMPv1 and SNMPv2.

Command	Purpose
configure	Enter Global Configuration mode

<table>
<tr><th>Command</th><th>Purpose</th></tr>
<tr><td>snmp-server community string [ro | rw | su] [view view-name] [ipaddress ip_address]</td><td>Configure the community string and specify access criteria for the community.
• community-string — Acts as a password and is used to authenticate the SNMP management station to the switch. The string must also be defined on the NMS in order for the NMS to access the SNMP agent on the switch (Range: 1-20 characters)
• ro — Indicates read-only access
• rw — Indicates read-write access.
• view-name — Specifies the name of a previously defined MIB view.
• ip_address — Specifies the IPv4 address mask of the allowed management stations. If no IP address or an all-zeros IP address is specified, all management stations are permitted. The IPv4 address is a mask where '1' bits are AND'd with the connecting station address and the result is compared with the IPv4 address mask. If not equal, access is not allowed.</td></tr>
<tr><td>snmp-server community-group community-string group-name [ipaddress ip-address]</td><td>Map the internal security name for SNMP v1 and SNMP v2 security models to the group name.
• community-string — Community string that acts like a password and permits access to the SNMP protocol. (Range: 1-20 characters)
• group-name — Name of a previously defined group. The group defines the objects available to the community. (Range: 1-30 characters)
• ip-address — Management station IPv4 address mask. Default is all IP addresses are allowed access.</td></tr>
<tr><td>exit</td><td>Exit to Privileged EXEC mode.</td></tr>
<tr><td>show snmp</td><td>View SNMP settings and verify the configuration</td></tr>
</table>

Configuring SNMP Notifications (Traps and Informs)

Beginning in Privileged EXEC mode, use the following commands to allow the switch to send SNMP traps and to configure which traps are sent.

Command	Purpose
configure	Enter Global Configuration mode
snmp-server enable traps [**vrf** *vrf-name*] [**acl** | **all** | **auto-copy-sw** | **bgp** | **captive-portal** *cp-type* | **dot1q** | **dvrmp** | **link** | **port-security** | **multiple-users** | **ospf** *ospftype* | **ospfv3** *ospfv3type* | **pim** | **poe** | **snmp authentication** | **spanning-tree** | **stack** | **vrrp**]	Specify the traps to enable. The captive portal, OSPF and OSPFv3 traps include several different traps that can be enabled. For more information, use the CLI command help or see the CLI Command Reference.
snmp-server filter *filter-name oid-tree* {**included** | **excluded**}	Configure a filter for SNMP traps and informs based on OIDs. Each OID is linked to a device feature or a feature aspect. • *filter-name* — Specifies the label for the filter record that is being updated or created. The name is used to reference the record. (Range: 1-30 characters.) • *oid-tree* — Specifies the object identifier of the ASN.1 subtree to be included or excluded from the view. To identify the subtree, specify a text string consisting of numbers, such as `1.3.6.2.4`, or a word, such as `system`. Replace a single subidentifier with the asterisk (*) wildcard to specify a subtree family; for example, `1.3.*.4`. • **included** — Indicates that the filter type is included. • **excluded** — Indicates that the filter type is excluded.

Command	Purpose
snmp-server host *host-addr* [**informs** [**timeout** *seconds*] [**retries** *retries*] | **traps version** {1 | 2}]] *community-string* [**udp-port** *port*] [**filter** *filtername*]	For SNMPv1 and SNMPv2, identify the system to receive SNMP traps or informs. • *host-addr* — Specifies the IP address of the host (targeted recipient) or the name of the host. (Range:1-158 characters). • **informs** — Indicates that SNMPv2 informs are sent to this host • **timeout** *seconds* — Number of seconds to wait for an acknowledgment before resending informs. The default is 15 seconds. (Range: 1-300 characters.) • *retries* — Maximum number of times to resend an inform request. The default is 3 attempts. • **traps** — Indicates that SNMP traps are sent to this host – **version 1** — Indicates that SNMPv1 traps will be used – **version 2** — Indicates that SNMPv2 traps will be used • *community-string* — Specifies a password-like community string sent with the notification operation. (Range: 1-20 characters) • *port* — UDP port of the host to use. The default is 162. (Range: 1-65535 characters.) • *filtername* — A string that is the name of the filter that defines the filter for this host. If unspecified, does not filter anything (Range: 1-30 characters.)

Command	Purpose
snmp-server v3-host {*ip-address* | *hostname*} *username* {**traps** | **informs**} [**noauth** | **auth** | **priv**] [**timeout** *seconds*] [**retries** *retries*] [**udpport** *port*] [**filter** *filtername*]	For SNMPv3, identify the system to receive SNMP traps or informs. • *ip-address* — Specifies the IP address of the host (targeted recipient). • *hostname* — Specifies the name of the host. (Range: 1-158 characters.) • *username* — Specifies user name used to generate the notification. (Range: 1-25 characters.) • **traps** — Indicates that SNMP traps are sent to this host. • **informs** — Indicates that SNMPv2 informs are sent to this host. • **noauth** — Specifies sending of a packet without authentication. • **auth** — Specifies authentication of a packet without encrypting it • **priv** — Specifies authentication and encryption of a packet. • *seconds* — Number of seconds to wait for an acknowledgment before resending informs. This is not allowed for hosts configured to send traps. The default is 15 seconds. (Range: 1-300 seconds.) • *retries* — Maximum number of times to resend an inform request. This is not allowed for hosts configured to send traps. The default is 3 attempts. (Range: 0-255 retries.) • *port* — UDP port of the host to use. The default is 162. (Range: 1-65535.) • *filter-name* — Specifies the optional filter (defined with the **snmp-server filter** command) to use for the host. (Range: 1-30 characters.)
exit	Exit to Privileged EXEC mode.
show trapflags	View the status of the configurable SNMP traps.

SNMP Configuration Examples

This section contains the following examples:

- Configuring SNMPv1 and SNMPv2
- Configuring SNMPv3

Configuring SNMPv1 and SNMPv2

This example shows how to complete a basic SNMPv1/v2 configuration. The commands enable read-only access from any host to all objects on the switch using the community string *public*, and enable read-write access from any host to all objects on the switch using the community string *private*.

This example also shows how to allow the switch to generate traps for all features that produce traps. The traps are sent to the host with an IP address of 192.168.3.65 using the community string *public*.

To configure the switch:

1 Configure the public community string.

```
console#configure
console(config)#snmp-server community public ro
```

2 Configure the private community string.

```
console(config)#snmp-server community private rw
```

3 Enable all traps and specify the IP address of the host where the traps should be sent.

```
console(config)#snmp-server enable traps all
console(config)#snmp-server host 192.168.3.65 public
console(config)#exit
```

4 View the current SNMP configuration on the switch.

```
console#show snmp

Community-String     Community-Access View Name IP Address
-------------------- ---------------- --------- -------
private              Read/Write       Default    All
public               Read Only        Default    All


Community-String     Group Name          IP Address
-----------------    --------------      ------------
private              DefaultWrite        All
```

```
public               DefaultRead         All

Traps are enabled.
Authentication trap is enabled.
Version 1,2 notifications
Target Addr. Type  Community  Version UDP   Filter TO    Retries
                                      Port  Name   Sec
------------ ---- ---------  ----     ----- -----  ---   -------
192.168.3.65 Trap  public     1       162

Version 3 notifications
Target Addr. Type  Username   Security UDP   Filter TO    Retries
                              Level    Port  Name   Sec
------------ ---- ---------  -------  ----- -----  ---   -------

System Contact:
System Location:
```

Configuring SNMPv3

This example shows how to complete a basic SNMPv3 configuration. The commands create a view that includes objects from the *internet* MIB subtree (OID 1.3.6.1), which includes all objects on the switch.

The user named *admin* has read-write privileges to all objects within the view (in other words, all objects on the switch) after supplying the appropriate authentication credentials (secretkey).

To configure the switch:

1. Configure the view. *view_snmpv3* and specify the objects to include.

```
console#configure
console(config)#snmp-server view view_snmpv3 internet included
```

2. Create the group *group_snmpv3* and allow read-write access to the view configured in the previous step.

```
console(config)#snmp-server group group_snmpv3 v3 auth read
view_snmpv3 write view_snmpv3
```

3. Create the user *admin*, assign the user to the group, and specify the authentication credentials.

```
console(config)#snmp-server user admin group_snmpv3 auth-md5
secretkey
```

4. Specify the IP address of the host where traps are to be sent. Packet authentication using MD5-SHA is enabled for the traps.

```
console(config)#snmp-server v3-host 192.168.3.35 admin traps
auth
console(config)#exit
```

5 View the current SNMP configuration on the switch. The output includes the SNMPv1/2 configuration in the previous example.

```
console#show snmp

Community-String     Community-Access View Name IP Address
-------------------- ---------------- --------- -------
private              Read/Write       Default    All
public               Read Only        Default    All


Community-String     Group Name           IP Address
-----------------    --------------       ------------
private              DefaultWrite         All
public               DefaultRead          All

Traps are enabled.
Authentication trap is enabled.
Version 1,2 notifications
Target Addr. Type Community  Version UDP   Filter TO   Retries
                                     Port  Name   Sec
------------ ---- ---------  ----    ----- -----  ---  -------
192.168.3.65 Trap public     1       162

Version 3 notifications
Target Addr. Type Username   Security UDP   Filter TO   Retries
                             Level    Port  Name   Sec
------------ ---- ---------  -------  ----- -----  ---  -------
192.168.3.35 Trap admin      Auth-NoP 162          15   3

System Contact:
System Location:

console#show snmp views
Name                 OID Tree                 Type
------------------   ------------------------ ------------
Default              iso                      Included
Default              snmpVacmMIB              Excluded
Default              usmUser                  Excluded
Default              snmpCommunityTable       Excluded
view_snmpv3          internet                 Included
```

```
DefaultSuper         iso                      Included

console#show snmp group

Name         Context  Model  Security Read     Views  Notify
             Prefix          Level             Write
------------ -------- ------ -------- -------- ------ -------
DefaultRead  ""       V1     NoAuth-  Default  ""     Default
                             NoPriv
DefaultRead  ""       V2     NoAuth-  Default  ""     Default
                             NoPriv
DefaultSuper ""       V1     NoAuth-  DefaultSuDefault Default
                             NoPriv   per      Super  Super
DefaultSuper ""       V2     NoAuth-  DefaultSuDefault Default
                             NoPriv   per      Super  Super
DefaultWrite ""       V1     NoAuth-  Default  Default Default
                             NoPriv
DefaultWrite ""       V2     NoAuth-  Default  Default Default
                             NoPriv
group_snmpv3 ""       V3     Auth-    view_snmpview_sn ""
                             NoPriv   v3       mpv3

console#show snmp user

Name       Group Name     Auth   Priv   Remote Engine ID
                          Meth   Meth
---------  -----------    -----  -----  ----------------
admin      group_snmpv3   MD5           800002a203001ec9aaaa07
```

14

Images and File Management

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to upload, download, and copy files, such as firmware images and configuration files, on the switch. The topics covered in this chapter include:

- Image and File Management Overview
- N3000 Dual Images
- Managing Images and Files (Web)
- Managing Images and Files (CLI)
- File and Image Management Configuration Examples

NOTE: For information about the Auto Configuration feature that enables the switch to automatically upgrade the image or load a new configuration file during the boot process, see DHCP and USB Auto-Configuration.

Image and File Management Overview

What Files Can Be Managed?

Dell Networking N-Series switches maintain several different types of files on the flash file system. Table 14-1 describes the files that can be managed. The table also lists the type of action that can be taken on the file, which is one or more of the following:

- Download the file to the switch from a remote system (or USB flash drive).
- Upload the file from the switch to a remote system (or USB flash drive).
- Copy the file from one location on the file system to another location.

Table 14-1. Files to Manage

File	Action	Description
image	Download Upload Copy	Firmware for the switch. The switch can maintain two images: the active image and the backup image.
startup-config	Download Upload Copy	Contains the software configuration that loads during the boot process.
running-config	Download Upload Copy	Contains the current switch configuration.
backup-config	Download Upload Copy	An additional configuration file that serves as a backup.
Configuration script	Download Upload	Text file with CLI commands. When you activate a script on the switch, the commands are executed and added to the running-config.
Log files	Upload	Provides various information about events that occur on the switch. For more information, see Monitoring and Logging System Information.
SSH key files	Download	Contains information to authenticate SSH sessions. The switch supports the following files for SSH: • SSH-1 RSA Key File • SSH-2 RSA Key File (PEM Encoded) • SSH-2 Digital Signature Algorithm (DSA) Key File (PEM Encoded)

Table 14-1. Files to Manage

File	Action	Description
SSL certificate files	Download	Contains information to encrypt, authenticate, and validate HTTPS sessions. The switch supports the following files for SSL: • SSL Trusted Root Certificate File (PEM Encoded) • SSL Server Certificate File (PEM Encoded) • SSL Diffie-Hellman Weak Encryption Parameter File (PEM Encoded) • SSL Diffie-Hellman Strong Encryption Parameter File (PEM Encoded)
IAS Users	Download	List of Internal Authentication Server (IAS) users for IEEE 802.1X authentication. For more information, see "What is the Internal Authentication Server? " on page 287

Why Is File Management Needed?

This section provides some reasons why you might choose to manage various files.

Image Files

The switch can store two firmware images, but only one is active. The other image file is a backup image. By default, the switch has only one image. You might copy an image or download an image to the switch for the following reasons:

- To create a backup image
- To upgrade the firmware as new images become available

The Dell Networking N-Series switches are named as follows:

<switch name>v<version number>.stk

Where the switch name is:

N4000 — Dell Networking N4000 Series switch firmware for N4032, N4032F, N4064, N4064F.

N3000_N2000 — Dell Networking N2000/N3000 Series switch firmware for N2024, N2048, N2024P, N2048P, N3024, N3024P, N3024F, N3048, N3048P.

N3000_BGP — Dell Networking N3000 Series switch firmware for N3024, N3024P, N3024F, N3048, N3048P.

N1500 — Dell Networking N1500 Series switch firmware for N1524, N1524P, N1548, N1548P.

And the version number is:

Version Numbering Convention

Version number	Description
6 0 0 1	Four part version number
	Denotes the build number.
	Denotes a scheduled maintenance release of the firmware.
	Denotes a minor release of the firmware.
	Denotes a major release of the firmware.

- Major release numbers start at 6.
- Minor release numbers start at 0.
- Maintenance release numbers start at 0.
- Web release build numbers start at 1. A build number of 0 indicates a factory build, which should be upgraded using a web release build from www.dell.com/support.

Examples:

- N1500v6.2.5.0.stk — Dell Networking N1500 Series switch firmware release 6.2.5.0. This is the factory build for the fifth maintenance release of the second minor release of the 6.X major release family.

- N3000_N2000v6.0.1.3.stk — Dell Networking N3000/N2000 Series switch firmware version 6.0.1.3. This is the third build for the first maintenance release for the 6.0 major release.
- N4000v6.1.0.1.stk — Dell Networking N4000 Series switch firmware version 6.1.0.1. This is the first build for the first minor release after the 6.0 major release, i.e., release 6.1.

Configuration Files

Configuration files contain the CLI commands that change the switch from its default configuration. The switch can maintain three separate configuration files: startup-config, running-config, and backup-config. The switch loads the startup-config file when the switch boots. Any configuration changes that take place after the boot process completes are written to the running-config file. The backup-config file does not exist until you explicitly create one by copying an existing configuration file to the backup-config file or downloading a backup-config file to the switch.

Configuration scripts, which are text files that contains CLI commands, can also be created.

NOTE: You must use the CLI to manage configuration scripts. The configuration scripting feature is not available from the web interface.

When you apply (run) a configuration script on the switch, the commands in the script are executed in the order in which they are written as if you were typing them into the CLI. The commands that are executed in the configuration script are added to the running-config file.

You might upload a configuration file from the switch to a remote server for the following reasons:

- To create a backup copy
- To use the configuration file on another switch
- To manually edit the file

You might download a configuration file from a remote server to the switch for the following reasons:

- To restore a previous configuration
- To load the configuration copied from another switch

- To load the same configuration file on multiple switches

Use a text editor to open a configuration file and view or change its contents.

SSH/SSL Keys and Certificates

If you use OpenManage Switch Administrator to manage the switch over an HTTPS connection, you must import the appropriate certificate files to the switch (**crypto key import**). If you use the CLI to manage the switch over an SSH connection, you must import the appropriate key files to the switch or use the **crypto key** command to generate the key files locally.

What Methods Are Supported for File Management?

Any of the following protocols can be used to download files from a remote system to the switch or to upload files from the switch to a remote system:

- TFTP
- SFTP
- SCP
- FTP
- HTTP (Web only)
- HTTPS (Web only)

Files can also be copied between the file system on the internal flash and a USB flash drive that is connected to the external USB port.

What Factors Should Be Considered When Managing Files?

Uploading and Downloading Files

To use TFTP, SFTP, SCP, or FTP for file management, you must provide the IP address of the remote system that is running the appropriate server (TFTP, SFTP, SCP or FTP). Make sure there is a route from the switch to the remote system. The **ping** command in the CLI can be used to verify that a route exists between the switch and the remote system.

If you are downloading a file from the remote system to the switch, be sure to provide the correct path to the file and the correct file name.

Managing Images

When you download a new image to the switch, it overwrites the backup image, if it exists. To use the new image, it must be activated and reloaded on the switch. The image that was previously the active image becomes the backup image after the switch reloads. If the switch is upgraded to a newer image and the image is found to be incompatible with the network, the switch can revert to the original image.

If a new image is activated and reloaded on the switch, and the switch is unable to complete the boot process due to a corrupt image or other problem, the boot menu can be used to activate the backup image. The administrator must connect to the switch through the console port to access the boot menu. The image files may contain firmware for the PHY processors on the switch. The PHY firmware may be updated to the firmware version supported by the switch firmware during the boot process or, in the case of switches that support the hot swap of cards, when the card is inserted into the switch.

Editing and Downloading Configuration Files

Each configuration file contains a list of executable CLI commands. The commands must be complete and in a logical order, as if you were entering them by using the switch CLI.

When you download a startup-config or backup-config file to the switch, the new file replaces the previous version. To change the running-config file, you execute CLI commands either by typing them into the CLI or by applying a configuration script with the **script apply** command. The startup-config and backup-config files can also be applied to the running-config by using the **script apply** command.

Creating and Applying Configuration Scripts

When you use configuration scripting, keep the following considerations and rules in mind:

- The application of scripts is partial if the script fails. For example, if the script executes four of ten commands and the script fails, the script stops at four, and the final six commands are not executed.
- Scripts cannot be modified or deleted while being applied.
- Validation of scripts checks for syntax errors only. It does not validate that the script will run.

- The file extension must be `.scr`.
- A maximum of seven scripts are allowed on the switch.
- The combined size of all script files on the switch cannot exceed 2 MB.
- The maximum number of configuration file command lines is 2000.

Single-line annotations in the configuration file can be used to improve script readability. The exclamation point (!) character flags the beginning of a comment. The comment flag character can begin anywhere within a single line, and all input following this character to the end of the line is ignored. Any line in the file that begins with the "!" character is recognized as a comment line and ignored by the parser.

The following example shows annotations within a file (commands are bold):

```
! Script file for displaying management access
show telnet !Displays the information about remote connections
! Display information about direct connections
show serial
! End of the script file
```

Managing Files on a Stack

Image files downloaded to the master unit of a stack are automatically downloaded to all stack members. If you activate the backup image on the master, it is activated on all units as well so that when you reload the stack, all units use the same image.

The running-config, startup-config, and backup-config files, as well as all keys and certificates are synchronized across the stack when the running-config file is saved to the startup-config file.

Configuration scripts are not distributed across the stack and only exist on the unit that is the master unit at the time of the file download.

Uploading Configuration Files by Using SNMP

When you use SNMP to upload a configuration file to a TFTP server, the agentTransferUploadFileName object must be set to the local filename, which is either startup-config or backup-config.

How Is the Running Configuration Saved?

Changes you make to the switch configuration while the switch is operating are written to the running-config. These changes are not automatically written to the startup-config. When you reload the switch, the startup-config file is loaded. If you reload the switch (or if the switch resets unexpectedly), any settings in the running-config that were not explicitly saved to the startup-config are lost. You must save the running-config to the startup-config to ensure that the settings you configure on the switch are saved across a switch reset.

To save the running-config to the startup-config by using the web-based interface, click (the save icon), which is available at the top of each page.

To save the running-config to the startup-config from the CLI, use the **write** command.

N3000 Dual Images

The Dell Networking N3000 Series switches support two image types. Each image type has a distinct set of features intended to optimize the switch resources for the intended purpose. For each image type, the user interface has been optimized to eliminate access to the omitted features.

When changing roles, the configuration for the preceding role will show an error during switch reboot. The unsupported configuration lines are not placed into the running-config, and a message showing the unsupported configuration lines is shown. It is recommended that the startup-config be removed or modified prior to rebooting when changing roles.

Access Router/Switch Role

Standard naming format: N3000_N2000vA.B.C.D.stk.

This image operates on both Dell Networking N3000/N2000 Series switches and should be the only image downloaded to the N2000 and selected N3000 Series switches.

Table 14-2 shows the feature set differences for this image (features not shown are included in the image). This image is the default image present on switches shipped from the factory.

Table 14-2. Features Included/Excluded in Access Router/Switch Image

Includes	Excludes
iSCSI	BGP
DVLAN	
MVR	
MLAG	
GARP	
Auto-VoIP	
Web interface for all features	

Aggregation Router Role

New naming format: N3000_BGPvA.B.C.D.stk.

This image should only be downloaded to the Dell Networking N3000 Series switches.

Table 14-3 shows the feature set for this image (features not shown are included in the image).

Table 14-3. Features Included/Excluded in Aggregation Router Image

Includes	Excludes
BGP	iSCSI
	DVLAN
	MVR
	MLAG
	GARP
	Auto-VoIP
	Web interface

Managing Images and Files (Web)

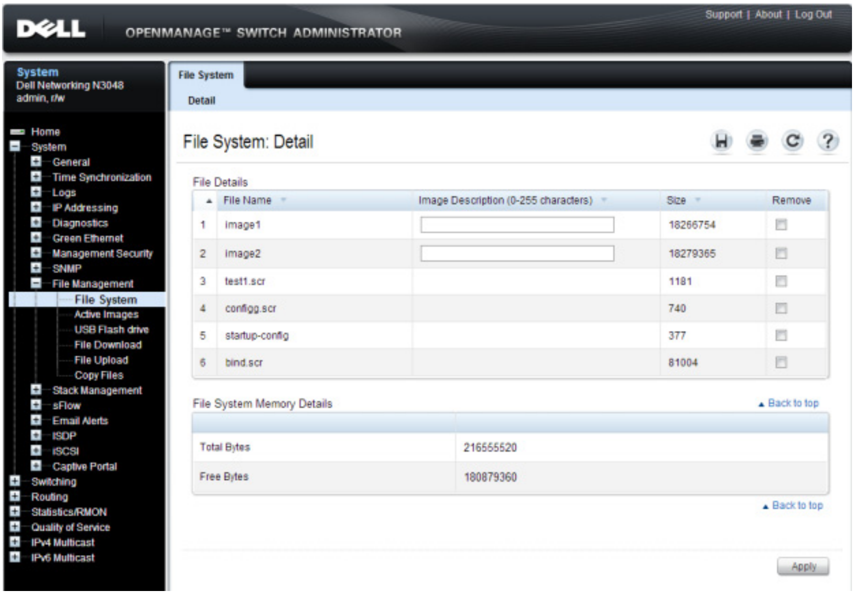
This section provides information about the OpenManage Switch Administrator pages to use to manage images and files on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

File System

Use the **File System** page to view a list of the files on the device and to modify the image file descriptions.

To display the **File System** page, click **System** → **File Management** → **File System** in the navigation panel.

Figure 14-1. File System



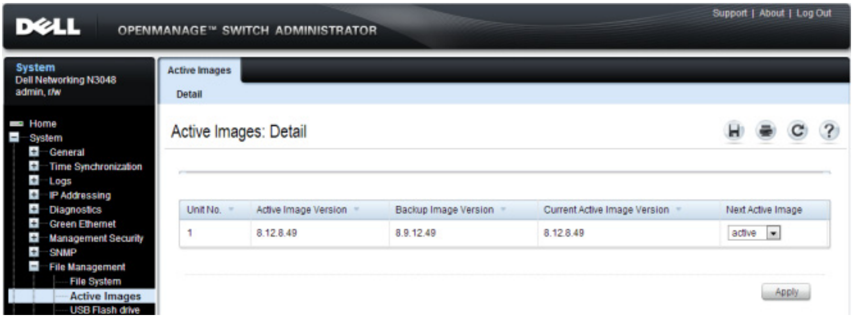
Active Images

Use the **Active Images** page to set the firmware image to use when the switch boots. If you change the boot image, it does not become the active image until you reset the switch.

On the Dell Networking N4000 Series switches, the images are named *active* and *backup*.

To display the **Active Images** page, click **System** → **File Management** → **Active Images** in the navigation panel.

Figure 14-2. Active Images



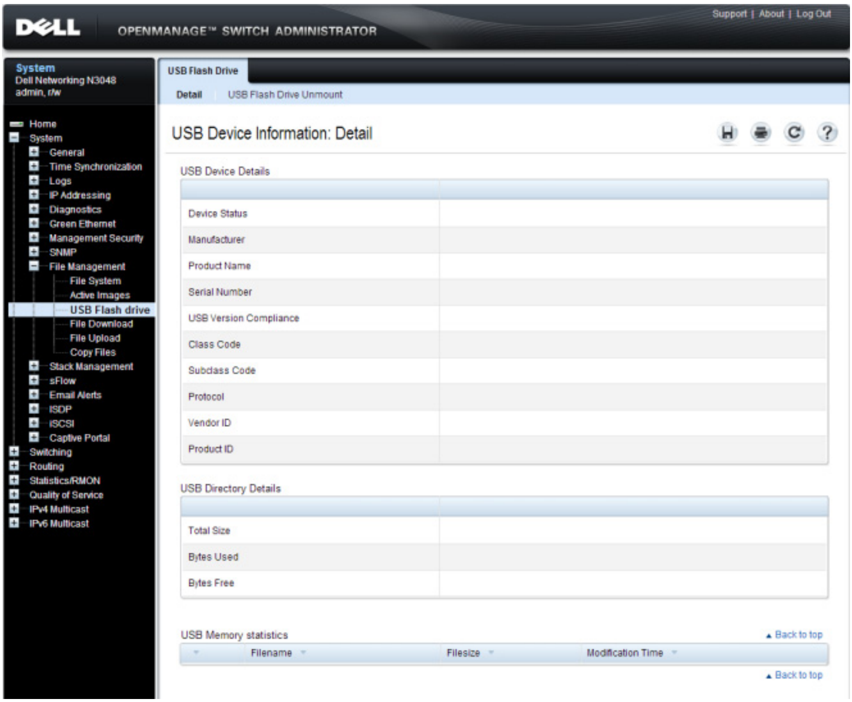
USB Flash Drive

Use the **USB Flash Drive** page to view information about a USB flash drive connected to the USB port on the front panel of the switch. The page also displays information about the files stored on the USB flash drive.

A USB flash drive must be un-mounted by the operator before removing it from the switch. If a new USB flash drive is installed without un-mounting the previous drive, the new flash drive may not be recognized. If a USB flash drive is removed without un-mounting it, un-mount the flash drive (i.e., use the command **unmount usb**) and remove and reinstall the USB flash drive in the switch.

To display the **USB Flash Drive** page, click **System** → **File Management** → **USB Flash Drive** in the navigation panel.

Figure 14-3. USB Flash Drive

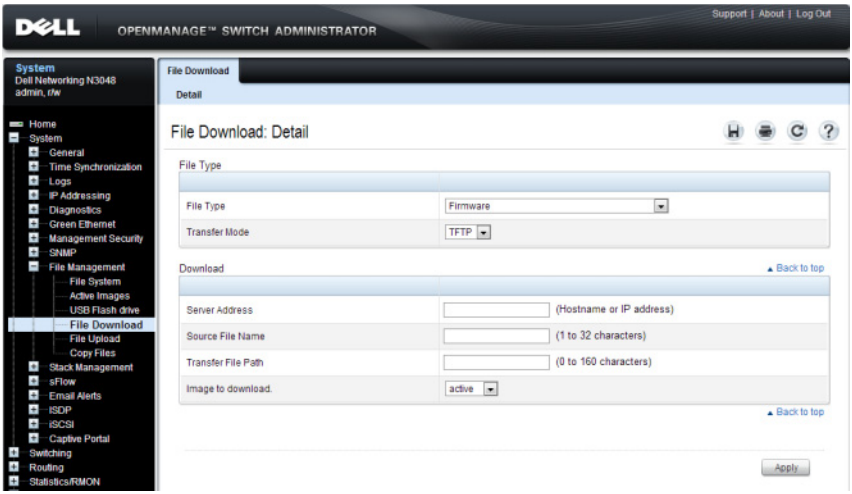


File Download

Use the **File Download** page to download image (binary) files, SSH and SSL certificates, IAS User files, and configuration (ASCII), files from a remote server to the switch.

To display the **File Download** page, click **System** → **File Management** → **File Download** in the navigation panel.

Figure 14-4. File Download



Downloading Files

To download a file to the switch:

1 Open the **File Download** page.

2 Select the type of file to download to the switch.

3 Select the transfer mode.

If you select a transfer mode that requires authentication, additional fields appear in the Download section. If you select HTTP as the download method, some of the fields are hidden.

NOTE: If you are using HTTPS to manage the switch, the download method will be HTTPS.

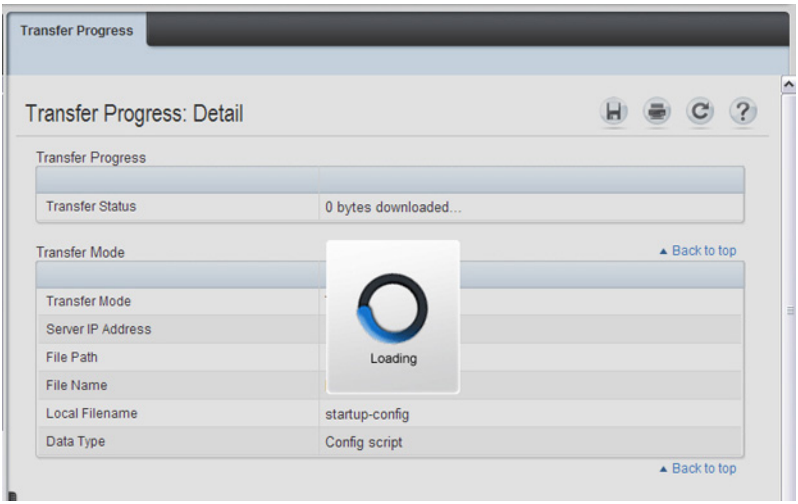
4 To download using HTTP, click **Browse** and select the file to download, then click **Apply**.

5 To download using any method other than HTTP, enter the IP address of the server that contains the file to download, the name of the file and the path on the server where it is located. For SFTP and SCP, provide the user name and password.

6 Click **Apply** to begin the download.

NOTE: After you start a file download, the page refreshes and a transfer status field appears to indicate the number of bytes transferred. The web interface is blocked until the file download is complete.

Figure 14-5. File Download in Progress



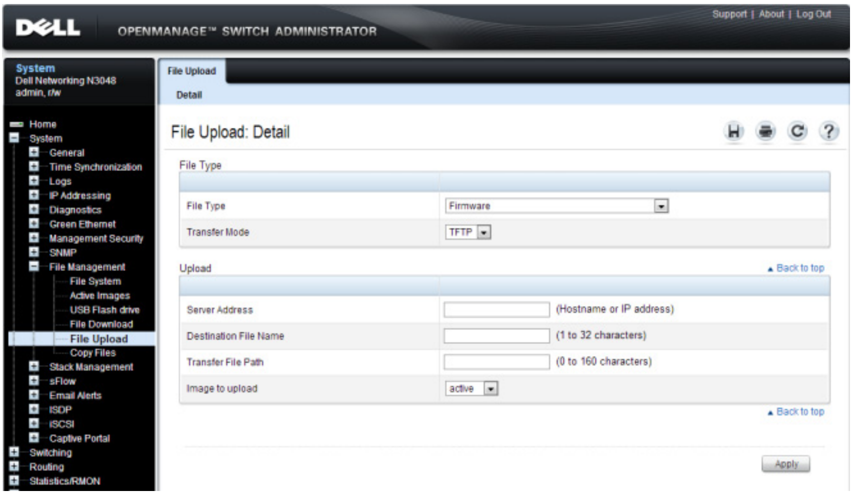
7 The file is downloaded to the switch.

File Upload

Use the **File Upload to Server** page to upload configuration (ASCII), image (binary), IAS user, operational log, and startup log files from the switch to a remote server.

To display the **File Upload to Server** page, click **System** → **File Management** → **File Upload** in the navigation panel.

Figure 14-6. File Upload



Uploading Files

To upload a file from the switch to a remote system:

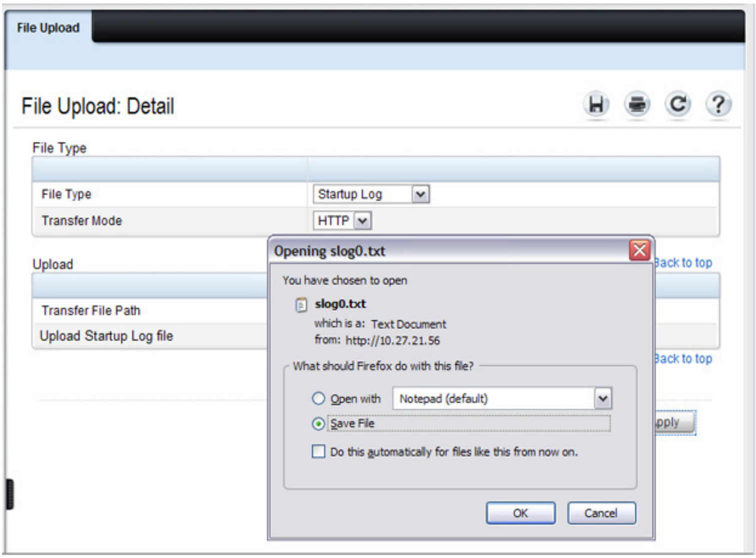
1. Open the **File Upload** page.
2. Select the type of file to download to the remote server.
3. Select the transfer mode.

   If you select a transfer mode that requires authentication, additional fields appear in the Upload section. If you select HTTP as the upload method, some of the fields are hidden.

NOTE: If you are using HTTPS to manage the switch, the download method will be HTTPS.

4 To upload by using HTTP, click **Apply**. A dialog box opens to allow you to open or save the file.

Figure 14-7. File Upload



5 To upload by using any method other than HTTP, enter the IP address of the server and specify a name for the file. For SFTP and SCP, provide the user name and password.

6 Click **Apply** to begin the upload.

NOTE: For some file uploads and methods, the page refreshes and a transfer status field appears to indicate the number of bytes transferred. The web interface is blocked until the file upload is complete.

7 The file is uploaded to the specified location on the remote server.

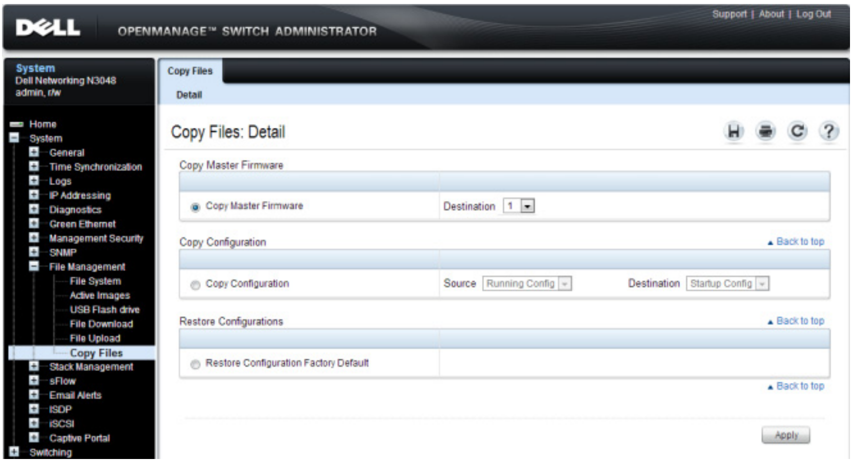
Copy Files

Use the **Copy Files** page to:

- Copy the active firmware image to one or all members of a stack.
- Copy the running, startup, or backup configuration file to the startup or backup configuration file.
- Restore the running configuration to the factory default settings.

To display the **Copy Files** page, click **System** → **File Management** → **Copy Files** in the navigation panel.

Figure 14-8. Copy Files



Managing Images and Files (CLI)

This section provides information about the commands you use to upload, download, and copy files to and from the Dell Networking N1500, N2000, N3000, and N4000 Series switches. For more information about these commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support. It also describes the commands that control the Auto Configuration feature.

NOTE: Upload, download, and copy functions use the **copy** command. The basic syntax for the command is **copy** *source destination*. This section shows several different ways to use the **copy** command.

Downloading and Activating a New Image (TFTP)

Beginning in Privileged EXEC mode, use the following commands to download a new firmware image to the switch and to make it the active image. This example shows how to use TFTP to download the image.

Command	Purpose
copy tftp://{*ip-address*| *hostname*}/*path*/*file-name* **image**	Use TFTP to download the firmware image at the specified source to the non-active image. If the image file is in the TFTP file system root (download path), you do not need to specify the path in the command. On Dell Networking N4000 Series switches, use the following command: **copy tftp**://{*ip-address*| *hostname*}/*path*/*file-name* {**active** | **backup**}
show version	View information about the currently active image.
filedescr {**image1** | **image2**} *description*	Add a description to the image files. On Dell Networking N4000 Series switches, use the following command: **filedescr** {**active** | **backup**} *description*

Command	Purpose
boot system {image1 | image2}	Set the image to use as the boot (active) image after the switch resets. Images on the N4032/N4064 are named *active* and *backup*. For Dell Networking N4000 Series switches, use the following command: **boot system {active | backup}**
reload	Reboot the switch to make the new image the active image. You are prompted to verify that you want to continue.

Managing Files in Internal Flash

Beginning in Privileged EXEC mode, use the following commands to copy, rename, delete and list the files in the internal flash.

Command	Purpose
dir	List the files in the flash file system.
copy flash://*filename* **usb://***filename*	Copy a file from the internal flash to a USB flash drive. Use the **dir** command to see a list of the files that can be copied from the internal flash. Make sure a flash drive has been inserted in the USB port on the front panel before executing the command.
rename *current_name new_name*	Rename a file in flash.
delete *filename*	Remove the specified file.
erase {**startup-config** | **backup-image** | **backup-config** | **application**}	Erase the startup configuration, the backup configuration, the backup image, or a manufacturer-supplied application.
copy startup-config backup-config	Save the startup configuration to the backup configuration file.
copy running-config startup-config	Copy the current configuration to the startup configuration. This saves the current configuration to NVRAM.
show startup-config	View the contents of the startup-config file
show running-config	View the contents of the running-config file

Managing Files on a USB Flash Device

Beginning in Privileged EXEC mode, use the following commands to manage files that are on a USB device that is plugged into the USB flash port on the front panel of the switch.

Command	Purpose
show usb device	Display USB flash device details
dir usb	Display USB device contents and memory statistics
copy usb://*filename* {**backup-config** | **image** | **running-config** | **script** *filename*| **startup-config** | *filename*	Copy the specified file from the USB flash device to the specified file in internal flash.
unmount usb	Make the USB flash device inactive.

Uploading a Configuration File (SCP)

Beginning in Privileged EXEC mode, use the following commands to upload a configuration file from the switch to a remote system by using SCP.

Command	Purpose
copy *file* **scp://***user*@{*ip-address*| *hostname*}/*path* /*file-name*	Adds a description to an image file. The file can be one of the following files: • backup-config • image • operational-log • running-config • script *file-name* • startup-config • startup-log
Password entry	After you enter the **copy** command, the CLI prompts you for the password associated with the username.

Managing Configuration Scripts (SFTP)

Beginning in Privileged EXEC mode, use the following commands to download a configuration script from a remote system to the switch, validate the script, and activate it.

NOTE: The startup-config and backup-config files are essentially configuration scripts and can be validated and applied by using the commands in this section.

Command	Purpose
copy sftp://*user*@{*ip-address*|*hostname*}/*path* /*file-name* **script** *dest-name*	Downloads the specified script from the remote server to the switch.
Password entry	After you enter the **copy** command, the CLI prompts you for the password associated with the username.
script validate *script-name*	Checks the specified script for syntax errors. The script is automatically validated when you download it to the switch. This command can be used to validate it again.
script list	View the list of available scripts.
script activate *script-name*	Executes the commands within the script in order. The configuration changes in the script are applied to the running configuration.
script show *script-name*	View the contents of the specified script.

File and Image Management Configuration Examples

This section contains the following examples:

- Upgrading the Firmware
- Managing Configuration Scripts

Upgrading the Firmware

This example shows how to download a firmware image to the switch and activate it. The TFTP server in this example is PumpKIN, an open source TFTP server running on a Windows system.

- TFTP server IP address: 10.27.65.103
- File path: \image
- File name: dell_0308.stk

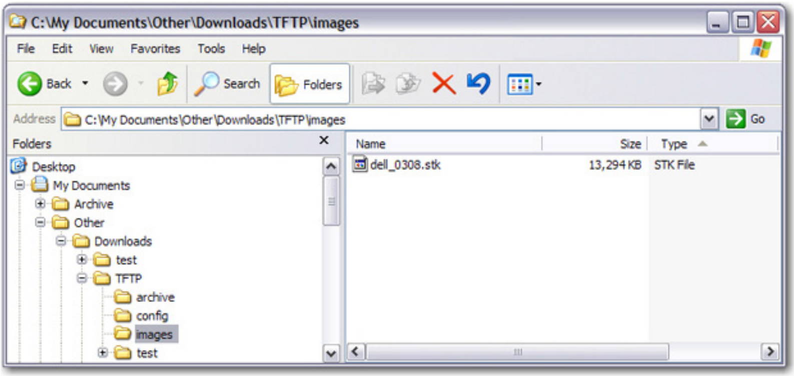
Use the following steps to prepare the download, and then download and upgrade the switch image.

1 Check the connectivity between the switch and the TFTP server.

```
console#ping 10.27.65.103
Pinging 10.27.65.103 with 0 bytes of data:

Reply From 10.27.65.103: icmp_seq = 0. time <10 msec.
Reply From 10.27.65.103: icmp_seq = 1. time <10 msec.
Reply From 10.27.65.103: icmp_seq = 2. time <10 msec.
Reply From 10.27.65.103: icmp_seq = 3. time <10 msec.

----10.27.65.103 PING statistics----
4 packets transmitted, 4 packets received, 0% packet
loss
round-trip (msec) min/avg/max = <10/<10/<10
```

2 Copy the image file to the appropriate directory on the TFTP server. In this example, the TFTP root directory is `C:\My Documents\Other\Downloads\TFTP`., so the file path is `images`.

Figure 14-9. Image Path



3 View information about the current image.

```
console#show version
Image Descriptions

 image1 :default image
 image2 :

Images currently available on Flash

------- ------------ ------------ --------------- --------------
unit    image1       image2       current-active  next-active
------- ------------ ------------ --------------- --------------
1       4.1.0.7      5.0.0.8      image1          image1
```

4 Download the image to the switch. After you execute the **copy** command, you must verify that you want to start the download.

Use either the **active** or **backup** keyword to select the specified image to replace (which takes effect only after a reboot). In the following example, the active image is replaced.

```
console#copy tftp://10.27.65.103/images/dell_0308.stk active

Mode............................................... TFTP
Set TFTP Server IP................................. 10.27.65.103
TFTP Path.......................................... images/
TFTP Filename...................................... dell_0308.stk
Data Type.......................................... Code
```

```
Destination Filename......................... image

Management access will be blocked for the duration of the
transfer
Are you sure you want to start? (y/n)y
```

5 Activate the new image (backup) so that it becomes the active image after the switch resets. This step is not necessary if downloading to the active image.

Use either the **active** or **backup** keyword, depending on which image you selected for replacement in step 4. It is possible to activate the backup image and subsequently reactivate the original active image prior to a reboot.

```
console#boot system backup
Activating image backup..
```

6 View information about the current image.

```
console#show bootvar
Image Descriptions

 active :
 backup :


 Images currently available on Flash

unit  active    backup    current-active    next-active
----- --------- --------- ----------------- -----------------
1     6.2.5.1   6.2.5.0   6.2.5.1           6.2.5.1
```

7 Copy the running configuration to the startup configuration to save the current configuration to NVRAM.

```
console#copy running-config startup-config

This operation may take a few minutes.
Management interfaces will not be available during this time.

Are you sure you want to save? (y/n)y

Configuration Saved!
```

8 Reset the switch to boot the system with the new image.

```
console#reload
```

```
Are you sure you want to continue? (y/n)y

Reloading all switches...
```

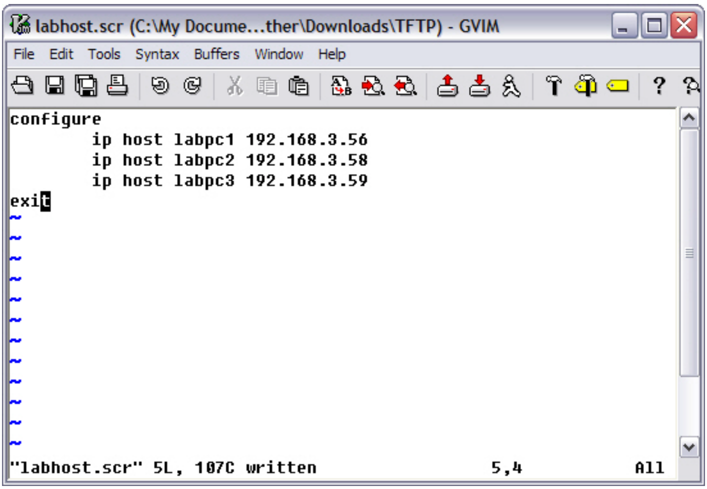
Managing Configuration Scripts

This example shows how to create a configuration script that adds three hostname-to-IP address mappings to the host table.

To configure the switch:

1 Open a text editor on an administrative computer and type the commands as if you were entering them by using the CLI.

Figure 14-10. Create Config Script



2 Save the file with an *.scr extension and copy it to the appropriate directory on your TFTP server.

3 Download the file from the TFTP server to the switch.

```
console#copy tftp://10.27.65.103/labhost.scr script
labhost.scr

Mode....................................... TFTP
Set TFTP Server IP......................... 10.27.65.103
TFTP Path.................................. ./
```

```
TFTP Filename................................ labhost.scr
Data Type.................................... Config Script
Destination Filename......................... labhost.scr

Management access will be blocked for the duration of the
transfer
```

4 After you confirm the download information and the script successfully downloads, it is automatically validated for correct syntax.

```
Are you sure you want to start? (y/n) y

135 bytes transferred

Validating configuration script...
configure
exit
configure
ip host labpc1 192.168.3.56

ip host labpc2 192.168.3.58

ip host labpc3 192.168.3.59

Configuration script validated.
File transfer operation completed successfully.
```

5 Run the script to execute the commands.

```
console#script apply labhost.scr

Are you sure you want to apply the configuration script? (y/n)y

configure
exit
configure
ip host labpc1 192.168.3.56

ip host labpc2 192.168.3.58

ip host labpc3 192.168.3.59

Configuration script 'labhost.scr' applied.
```

6 Verify that the script was successfully applied.

```
console#show hosts

Host name: test
Name/address lookup is enabled
Name servers (Preference order): 192.168.3.20
Configured host name-to-address mapping:
Host                      Addresses
------------------------  ------------------------
labpc1                    192.168.3.56
labpc2                    192.168.3.58
labpc3                    192.168.3.59
```

Managing Files by Using the USB Flash Drive

In this example, the administrator copies the backup image to a USB flash drive before overwriting the backup image on the switch with a new image. The administrator also makes a backup copy of the running-config by uploading it to a USB flash drive. After the backups are performed, the administrator downloads a new image from the USB flash drive to the switch to prepare for the upgrade.

This example assumes the new image is named new_img.stk and has already been copied from an administrative host onto the USB flash drive.

To configure the switch:

1 Insert the USB flash drive into the USB port on the front panel of the switch. The USB flash drive is automatically mounted.

2 Copy the backup image from the switch to the USB flash drive.

```
console#copy backup usb://img_backup.stk

Mode................................ unknown
Data Type........................... Code

Management access will be blocked for the duration of the
transfer
Are you sure you want to start? (y/n) y
```

3 Copy the running-config to the USB flash drive.

```
console#copy running-config usb://rc_backup.scr

Mode............................ unknown
Data Type....................... Config Script
Source Filename................. temp-config.scr
```

```
Management access will be blocked for the duration of the
transfer
Are you sure you want to start? (y/n) y
```

4 Download the new image from the USB flash drive to the switch. The image overwrites the backup image.

```
console#copy usb://new_image.stk backup

Mode................................... unknown
Data Type.............................. Code

Management access will be blocked for the duration of the
transfer
Are you sure you want to start? (y/n) y
```

5 To activate the new image after it has been successfully downloaded to the switch, follow the procedures described in Upgrading the Firmware, starting with step 5.

15

DHCP and USB Auto-Configuration

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

The topics covered in this chapter include:

- Auto Configuration Overview
- What Are the Dependencies for DHCP Auto Configuration?
- Default Auto Configuration Values
- Managing Auto Configuration (Web)
- Managing Auto Configuration (CLI)
- Auto Configuration Example

Auto Configuration Overview

The Auto Configuration feature can automatically update the firmware image and obtain configuration information when the switch boots. Auto Configuration begins the automatic download and installation process when the switch or stack master is initialized and no configuration file (startup-config) is found, or when the switch boots and loads a saved configuration that has Auto Configuration enabled. Auto Configuration is enabled by default. Allow downgrade is also enabled by default.

The Auto Configuration feature includes two components:

- USB Auto Configuration
- DHCP Auto Install

If no saved startup configuration file is found on the switch and the Auto Configuration feature is enabled (which it is by default), the Auto Configuration process begins. If a USB device is connected to the Dell Networking N-Series switch USB port and contains the appropriate file(s), the switch uses the USB Auto Configuration feature to update the configuration or image. If the USB Auto Configuration fails - either because it is disabled, no USB storage device is present, or no configuration or images files are present on the USB storage device, the switch uses the DHCP Auto Install process.

NOTE: Neither USB Configuration nor Auto Install is invoked if a saved startup configuration file is on the switch.

What Is USB Auto Configuration?

The USB Auto Configuration feature can be used to configure or upgrade one or more switches that have not been previously configured, such as when new switches are deployed. Before deploying the switch, the following steps must be performed:

1. Create a text file that contains IP addresses and netmasks (and, optionally, MAC addresses) and file names that are parsed and used to configure the switch. The optional MAC address used to identify the switch is the MAC address of the base MAC address of the switch, although the feature will accept any of the switch MAC addresses (see "Switch MAC Addresses " on page 142 for further information). The IP address and netmask are the minimum required fields in the configuration file. Refer to the example below for an explanation of the file format.
2. Copy the file onto a USB device, along with any desired switch firmware and configuration files.
3. Insert the USB device into the front-panel USB port on the Dell Networking N-Series switch.

When the Auto Configuration process starts and no saved startup-config file is present on the switch, the feature automatically searches a plugged-in USB device for information.

What Files Does USB Auto Configuration Use?

The USB Auto Configuration feature uses the following file types:

- *.setup file for initial switch configuration
- *.text file for configuration information
- *.stk file for software image installation

The Auto Configuration feature first searches the USB device for a file with a *.setup extension. If only one .setup file is present, the switch uses the file. When multiple *.setup files are present, the switch uses only the dellswitch.setup file. If no dellswitch.setup file is available, the switch checks for a file with a *.text configuration file and a *.stk image file. If multiple .text files exist, the switch uses the dellswitch.text file. If only a *.stk file is present, the switch checks the .stk file version and loads it into the backup image if the version is later than the current active image. If multiple *.stk files are present, the switch checks the image with the highest (most recent) version. Finally, if no *.setup, *.text, or *.stk files are found, the switch proceeds to the DHCP Auto Configuration process.

How Does USB Auto Configuration Use the Files on the USB Device?

The *.setup file can include the following information:

- MAC address of the switch (optional)
- IP address and netmask (mandatory)
- Config file (optional)
- Firmware Image file (optional)

MAC Address Lookup

The switch MAC address should be on the same line as the IP address and configuration file and/or image file name to allow a specific switch (identified by its MAC address) to be associated with a specific config file or image. The IP address on the line is assigned as the switch management IP address.

IP Address Lookup

If the switch MAC address is not found within the .setup text file, the first line that contains an IP address and no MAC address and is not marked in-use will be used by the switch to assign the management IP address/netmask. This method allows a group of IP addresses to be handed out without regard to the specific switch identified by the MAC address. A switch will mark a line as invalid if it is read and failed to properly parse if, for example, it contains an invalid configuration, a duplicate IP address or an image file name that is not available. Once a switch selects an IP address from the file, it adds its MAC address to the line, marks the line as in-use, and updates the file on the USB device.

If the *.setup file configuration line contains an IP address but no configuration or image file names, the management IP address will be assigned, and then the feature will search the USB device for files with the .text and .stk extensions, which indicates that all switches will be using the same configuration file and/or image on the USB device. This method allows different IP addresses to be assigned, but the same configuration file or image is downloaded to multiple switches. Alternatively, the line may contain a specific configuration or image file name, or both.

After the current switch has been configured and/or upgraded and the completion message is displayed on the switch, the current line in the *.setup text file will be marked as used. This allows using the *.setup file for additional switches without manually changing the file. The USB device can then be removed and inserted into the next switch to begin the process again. Also, the switch MAC address of a switch that has been automatically configured is added to the beginning of the line (if no MAC address was specified in the file) for lines using the IP address lookup method so that the MAC and IP address combinations are recorded within the *.setup file for future use bindings.

At the start of the next USB auto download, if all lines in the *.setup file are marked as already “in-use” or “invalid,” and there is no MAC address match for a switch, the process will halt, and a message similar to the following is displayed on the console:

```
<###> APR 22 08:32:43 Error: Auto Configuration has terminated
due to there being no more lines available for use within the
USB file "XXXXX.setup".
```

Configuration File

The *.text configuration file identified in the *.setup file contains the running-config to be loaded on to the switch. The configuration file may be specified on a line in the .setup file to assign specific configuration to specific switches, or, if it is desired to assign a single configuration to all switches, the configuration file need not be specified in the .setup file as long as it is present on the USB device. The configuration file specified in the *.setup file should exist on the USB device. It must have a .text file name extension. No other file name extension is allowed. For information about the format and contents of the *.text file, see Editing and Downloading Configuration Files.

Image File

If the Auto Configuration process includes a switch image upgrade, the name of the image file may optionally be included in the *.setup file. If it is desired to assign a specific image to a specific set of switches. If it is desired to use a single image for all switches being upgraded, it is not necessary to include the image file name in the .setup file as long as it is present on the USB device. The specified image file should exist on the USB device.

What Is the Setup File Format?

The setup file must have a *.setup extension or this part of the Auto Configuration process will never begin. If there are multiple .setup files located on the USB device, the dellswitch.setup file will be utilized. If no dellswitch.setup file is present and there are multiple .setup files present, the Auto Configuration process does not start.

The general format of the configuration file lines is as follows. The IP address and subnet mask are always required on each line of the .setup file. The MAC address, configuration file, and image file name entries are optional.

```
MAC_address   IP_Address   Subnet_Mask   Config_File   Image_File
```

The following example shows a *.setup example for two switches:

```
2180.c200.0010 192.168.0.10 255.255.255.0 switch-A.text N2000v6.0.1.3.stk
3380.c200.0011 192.168.0.11 255.255.255.0 switch-B.text N2000v6.0.1.3.stk
```

After a line has been read and implemented by the Auto Configuration feature, it automatically adds “in-use” to the end of the line to ensure that the information is not utilized for the next switch. To replicate the entire USB auto configuration process, the “in-use” statements from the .setup file need to be removed while leaving the inserted MAC address information. Then, if the process is restarted, the MAC address/IP address combinations will be ensured for any switch that has previously attempted upgrade and all other switch upgrades can take place as if for the first time.

What Is the DHCP Auto Configuration Process?

If the USB Auto Configuration fails or is not used, the switch can use a DHCP server to obtain configuration information from a TFTP server.

DHCP Auto Configuration is accomplished in three phases:

1. Assignment or configuration of an IP address for the switch
2. Assignment of a TFTP server
3. Obtaining a configuration file for the switch from the TFTP server

Auto Configuration is successful when an image or configuration file is downloaded to the switch or stack master from a TFTP server.

NOTE: The downloaded configuration file is not automatically saved to startup-config. You must explicitly issue a save request (**copy running-config startup-config**) in order to save the configuration. If the downloaded configuration is not saved to the startup-config, DHCP auto configuration will be done every time the DHCP lease expires.

Obtaining IP Address Information

DHCP is enabled by default on the Out-of-Band (OOB) interface on Dell Networking N3000 and N4000 Series switches. DHCP is enabled by default on VLAN 1 on the Dell Networking N1500 and N2000 Series switches. If an IP address has not been assigned, the switch issues requests for an IP address assignment.

A network DHCP server returns the following information:

- IP address and subnet mask to be assigned to the interface
- IP address of a default gateway, if needed for IP communication

After an IP address is assigned to the switch, if a hostname is not already assigned, Auto Configuration issues a DNS request for the corresponding hostname. This hostname is also displayed as the CLI prompt (as in response to the **hostname** command).

Obtaining Other Dynamic Information

The following information is also processed and may be returned by a BOOTP or DHCP server:

- Name of configuration file (the *file* field in the DHCP header or option 67) to be downloaded from the TFTP server.
- Identification of the TFTP server providing the file. The TFTP server can be identified by name or by IP address as follows:
  – hostname: DHCP option 66 or the *sname* field in the DHCP header)
  – IP address: DHCP option 150 or the *siaddr* field in the DHCP header
- NTP Server - Option 42
- SYSLOG Server - Option 7
- Client Host Name - Option 12
- Domain Name - Option 15

When a DHCP OFFER identifies the TFTP server more than once, the DHCP client selects one of the options in the following order: *sname*, option 66, option 150, *siaddr*. If the TFTP server is identified by hostname, a DNS server is required to translate the name to an IP address.

The DHCP client on the switch also processes the name of the text file (option 125, the V-I vendor-specific Information option) which contains the path to the image file as noted below.

Obtaining the Image

Auto Configuration attempts to download an image file from a TFTP server only if no configuration file was found in the internal flash or a USB drive, or even with a saved configuration file that has Auto Configuration enabled.

The network DHCP server may return a DHCP OFFER message with option 125. When configuring the network DHCP server for image downloads, you must include Option 125 and specify the Dell Enterprise Number, 674. Within the Dell section of option 125, sub-option 5 must specify the path and name of a file on the TFTP server. This file is not the image file itself, but rather a text file that contains the path and name of the image file. Upon receipt of option 125 with sub-option 5, the switch downloads the text file from the TFTP server, reads the name of the image file, and downloads the image file from the TFTP server.

For example, one might enter the following information:

- A2-02-00-00 — Dell Enterprise number 674. It should be written from right to left. The number 674 in decimal notation is 02 a2 00 00.
- 0c — Data Length (12 decimal)
- 05 — Sub-option code 5
- 0a — Sub-option length (10 decimal)
- Conversion of the file name from ACSII to hexadecimal:

  "Config.txt" – 43-6F-6E-66-69-67-2E-74-78-74

The Config.txt file should contain the full name of the image file on the TFTP server, including the path name, e.g.,

N3000_N2000v6.3.0.1.stk

or

mytftpserverpath/N3000_N2000v6.3.0.1.stk

Option 125 also supports sub-option 6, which is the path to a configuration file on the TFTP server. Only the path name is relevant. Configure the DHCP server to use vendor ID 674 and the required sub-option 6 and a hexadecimal encoded ASCII path value. If sub-option 6 is specified, the switch attempts to download the configuration file *<hostname>*.cfg using the DHCP-supplied host name (DHCP option 12). If that file is not found on the TFTP server, the switch attempts to download the "host.cfg" file. The configuration file consists of a series of CLI commands in ASCII text which are executed by the switch in Privileged Exec mode.

For example, one might enter the following information:

- A2-02-00-00 — Dell Enterprise number 674. It should be written from right to left. The number 674 in decimal notation is 02 a2 00 00.
- 08 — Data Length (8 decimal)
- 06 — Sub-option code 6
- 06 — Sub-option length (10 decimal)
- Conversion of the file name from ACSII to hexadecimal:

  "mypath" – 6D-79-70-61-74-68

After the switch successfully downloads and installs the new image, it automatically reboots. The download or installation might fail for one of the following reasons:

- The path or filename of the image on the TFTP server does not match the information specified in DHCP option 125.
- The downloaded image is the same as the current image.
- The validation checks, such as valid CRC Checksum, fails.

If the download or installation was unsuccessful, a message is logged.

NOTE: In stack of switches, the downloaded image is pushed to all members attached to the stack at the time of download. For members who join the stack after the download, the Stack Firmware Synchronization feature, if enabled, will push the latest image to all members.

Obtaining the Configuration File

If the DHCP OFFER identifies a configuration file, either as option 67 or in the *file* field of the DHCP header, the switch attempts to download the configuration file.

NOTE: The configuration file is required to have a file name that matches the following pattern: "*.cfg"

The TFTP client makes three unicast requests. If the unicast attempts fail, or if the DHCP OFFER did not specify a TFTP server address, the TFTP client makes three broadcast requests.

If the DHCP server does not specify a configuration file or download of the configuration file fails, the Auto Configuration process attempts to download a configuration file with the name `dell-net.cfg`. The switch unicasts or broadcasts TFTP requests for a network configuration file in the same manner as it attempts to download a host-specific configuration file.

The default network configuration file consists of a set of IP address-to-hostname mappings, using the command **ip host hostname** *address*. The switch finds its own IP address, as learned from the DHCP server, in the configuration file and extracts its hostname from the matching command. If the default network configuration file does not contain the switch's IP address, the switch attempts a reverse DNS lookup to resolve its hostname.

A sample `dell-net.cfg` file follows:

```
config
  ...
  ip host switch1 192.168.1.10
  ip host switch2 192.168.1.11
  ... <other hostname definitions>
exit
```

Once a hostname has been determined, the switch issues a TFTP request for a file named *`hostname`*`.cfg` , where *`hostname`* is the first thirty-two characters of the switch's hostname.

If the switch is unable to map its IP address to a hostname, Auto Configuration sends TFTP requests for the default configuration file `host.cfg.`

Table 15-1 summarizes the config files that may be downloaded and the order in which they are sought.

Table 15-1. Configuration File Possibilities

Order Sought	File Name	Description	Final File Sought
1	`bootfile.cfg`	Host-specific config file, ending in a *.cfg file extension	Yes
2	`dell-net.cfg`	Default network config file	No
3	`hostname.cfg`	Host-specific config file, associated with hostname.	Yes
4	`host.cfg`	Default config file	Yes

Table 15-2 displays the determining factors for issuing unicast or broadcast TFTP requests.

Table 15-2. TFTP Request Types

TFTP Server Address Available	Host-specific Switch Config Filename Available	TFTP Request Method
Yes	Yes	Issue a unicast request for the host-specific router config file to the TFTP server
Yes	No	Issue a unicast request for a default network or router config file to the TFTP server
No	Yes	Issue a broadcast request for the host-specific router config file to any available TFTP server
No	No	Issue a broadcast request for the default network or router config file to any available TFTP server

Monitoring and Completing the DHCP Auto Configuration Process

When the switch boots and triggers an Auto Configuration, a message displays on the console screen to indicate that the process is starting. After the process completes, the Auto Configuration process writes a log message. When Auto Configuration has successfully completed, the **show running-config** command can be used to validate the contents of configuration.

Saving a Configuration

The Auto Configuration feature includes an AutoSave capability that allows the downloaded configuration to be automatically saved; however, AutoSave is disabled by default. If AutoSave has not been enabled, you must explicitly save the downloaded configuration in nonvolatile memory on the stack master. This makes the configuration available for the next reboot. In the CLI, this is performed by issuing a **write** command or **copy running-config startup-config** command and should be done after validating the contents of saved configuration.

If the downloaded configuration is not saved to the startup-config, the configuration will be reloaded by the switch every time the DHCP lease expires.

Stopping and Restarting the Auto Configuration Process

The Auto Configuration process can be terminated at any time before the image or configuration file is downloaded. This is useful when the switch is disconnected from the network. Termination of the Auto Configuration process ends further periodic requests for a host-specific configuration file.

The Auto Configuration process automatically starts after a reboot if the startup-config file is not found on the switch. The configuration file will not be found if it has never been saved on the switch, or if you issue a command to erase the configuration file (**erase startup-config**).

Managing Downloaded Config Files

The configuration files downloaded by Auto Configuration are stored in the nonvolatile memory as .scr files. The files may be managed (viewed or deleted) along with files downloaded by the configuration scripting utility.

A file is not automatically deleted after it is downloaded. The file does not take effect upon a reboot unless you explicitly save the configuration (the saved configuration takes effect upon reboot). If you do not save the configuration downloaded by the Auto Configuration feature, the Auto Configuration process occurs again on a subsequent reboot or when the DHCP lease expires. This may result in one of the previously downloaded files being overwritten.

What Are the Dependencies for DHCP Auto Configuration?

The Auto Configuration process from TFTP servers depends upon the following network services:

- A DHCP server must be configured on the network with appropriate services.
- An image file and a text file containing the image file name for the switch must be available from a TFTP server if DHCP image download is desired.
- A configuration file (either from bootfile (or) option 67 option) for the switch must be available from a TFTP server.
- The switch must be connected to the network and have a layer-3 interface that is in an UP state.
- A DNS server must contain an IP address to hostname mapping for the TFTP server if the DHCP server response identifies the TFTP server by name.
- A DNS server must contain an IP address to hostname mapping for the switch if a `<hostname>.cfg` file is to be downloaded.
- If a default gateway is needed to forward TFTP requests, an IP helper address for TFTP needs to be configured on the default gateway.

Default Auto Configuration Values

Table 15-3 describes the Auto Configuration defaults.

Table 15-3. Auto Configuration Defaults

Feature	Default	Description
Auto Install Mode	Enabled	When the switch boots and no saved configuration is found, the Auto Configuration automatically begins.
Retry Count	3	When the DHCP or BootP server returns information about the TFTP server and bootfile, the switch makes three unicast TFTP requests for the specified bootfile. If the unicast attempts fail or if a TFTP server address was not provided, the switch makes three broadcast requests to any available TFTP server for the specified bootfile.
AutoSave	Disabled	If the switch is successfully auto-configured, the running configuration is not saved to the startup configuration.
AutoReboot	Enabled	After an image is successfully downloaded during the Auto Configuration process, the switch automatically reboots and makes the downloaded image the active image.

Managing Auto Configuration (Web)

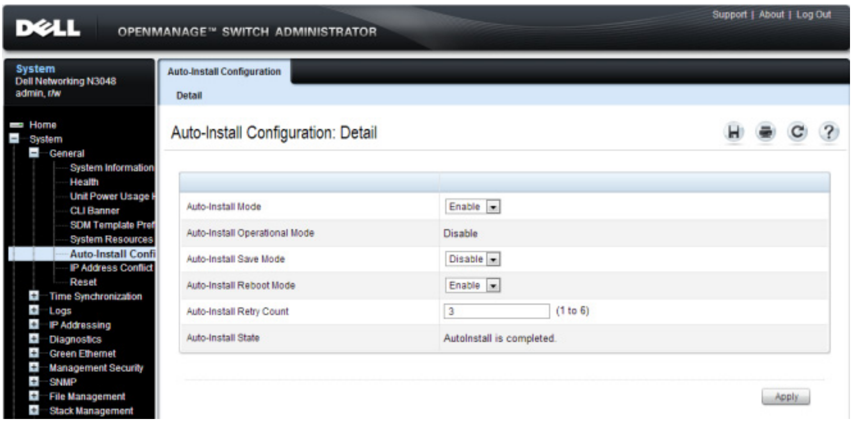
This section provides information about the OpenManage Switch Administrator pages to use to manage images and files on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

Auto-Install Configuration

Use the **Auto-Install Configuration** page to allow the switch to obtain network information (such as the IP address and subnet mask) and automatically download a host-specific or network configuration file during the boot process if no startup-config file is found.

To display the **Auto Configuration** page, click **System** → **General** → **Auto-Install Configuration** in the navigation panel.

Figure 15-1. Auto-Install Configuration



Managing Auto Configuration (CLI)

This section provides information about the commands you manage the Auto-Install Configuration feature on the switch. For more information about these commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Managing Auto Configuration

Beginning in Privileged EXEC mode, use the following commands to manually activate the Auto Configuration process and download a configuration script from a remote system to the switch, validate the script, and activate it.

NOTE: The Auto Configuration feature begins automatically when the switch is booted and no startup-config file is found or if the system boots and finds the **boot host dhcp** command in the startup-config file.

Command	Purpose
configure	Enter Global Configuration mode.
boot host dhcp	Enable Auto Configuration for the next reboot cycle. The command does not change the current behavior of Auto Configuration, but it does save the command to NVRAM.
boot host auto-save	Allow the switch to automatically save the configuration file downloaded to the switch by the Auto Configuration feature.
boot host retry-count *retries*	Specify the number of attempts to download the file (by sending unicast TFTP requests, and if unsuccessful, broadcast TFTP requests) specified in the response from the DHCP server. The range for *retries* is 1–3.
boot host auto-reboot	Allow the switch to automatically reboot when the image is successfully downloaded through the Auto Configuration feature.
exit	Exit to Privileged Exec mode.
show boot	Displays the current status of the Auto Configuration process.

Auto Configuration Example

A network administrator is deploying three Dell Networking N-Series switches and wants to quickly and automatically install the latest image and a common configuration file that configures basic settings such as VLAN creation and membership, RADIUS server settings, and 802.1X information. The configuration file also contains the command **boot host autosave** so that the downloaded configuration is automatically saved to the startup config.

This section describes two ways to enable automatic configuration file download:

- Enabling USB Auto Configuration and Auto Image Download
- Enabling DHCP Auto Configuration and Auto Image Download

Enabling USB Auto Configuration and Auto Image Download

This example describes how to deploy three switches and automatically install a custom configuration file on the switch and upgrade each switch with the latest software image by using the USB Auto Configuration feature. The switches have the following MAC addresses:

- Switch A: 001E.C9AA.AC17
- Switch B: 001E.C9AA.AC20
- Switch C: 001E.C9AA.AC33

To configure each switch with a pre-assigned IP address, include the switch MAC address in the .setup file on the line containing the corresponding IP address/netmask. An IP address and netmask are the minimum mandatory elements of each configuration line in the configuration file.

To use USB auto configuration:

1. Create a default config file for each switch (or one configuration file for all switches). For example, the configuration files may be named `switchA.txt`, `switchB.txt`, and `switchC.txt`. For information about creating configuration files, see Images and File Management.
2. Copy the configuration files to a USB device.
3. Copy the image file to the USB device. In this example, the image file that each switch will download is named `N2000v6.1.0.1.stk and N2000v6.2.0.1.stk`.

4 Create a setup file named `dellswitch.setup`. The setup file contains the following lines:

```
192.168.0.1 255.255.255.0 switchA.txt N2000v6.1.0.1.stk
192.168.0.2 255.255.255.0 switchB.txt N2000v6.2.0.1.stk
192.168.0.3 255.255.255.0 switchC.txt N2000v6.2.0.1.stk
```

5 Copy the `dellswitch.setup` file to the USB device.

6 Connect the USB device to Switch A.

7 Insert the USB device into the USB port on the front panel of Switch A.

8 Power on Switch A. If no startup-config file is found, the Easy Startup wizard will begin. Press N to skip the Easy Startup wizard and the USB Auto Configuration process will begin. If necessary, delete the startup-config file and reboot the switch.

   The configuration in `switchA.txt` file is downloaded to the switch, and the management interface acquires network information. After the process completes, a message displays to indicate the status. The `dellswitch.setup` file is updated to add the term `in-use` to the end of the line. The `N2000v6.1.0.1.stk` image is also downloaded to the switch and activated.

9 Remove the USB device from Switch A and insert it into Switch B.

10 Repeat the process to connect a port to the network. Power on the switch to begin the USB Auto Configuration process on Switch B.

11 Remove the USB device from Switch B after the process completes, and repeat the steps to perform the USB Auto Configuration process on Switch C. Note that switch A will use the 6.1.0.1 firmware and switches B and C will be loaded with 6.2.0.1 firmware.

Enabling DHCP Auto Configuration and Auto Image Download

If no USB device is connected to the USB port on the Dell Networking N-Series switch and no configuration file is found during the boot process, the Auto Configuration feature uses the DHCP Auto Configuration process to download the configuration file to the switch. This example describes the procedures to complete the configuration.

To use DHCP auto configuration:

1 Create a default config file for the switches named host.cfg. The host.cfg file contains the path and name of the image file on the TFTP server (option 125, sub-option 5). For information about creating configuration files, see Images and File Management.

2 Upload the host.cfg file to the TFTP server.

3 Upload the image file to the TFTP server.

4 Configure an address pool on the DHCP server to contain the following information:

   a The IP address (*yiaddr*) and subnet mask (option 1) to be assigned to the interface

   b The IP address of a default gateway (option 3)

   c DNS server address (option 6)

   d Name of config file for each host

   e Identification of the TFTP server by hostname (DHCP option 66 or the *sname* field in the DHCP header) or IP address (DHCP option 150 or the *siaddr* field in the DHCP header)

   f Name of the text file (option 125, the V-I vendor-specific Information option) that contains the image file name and path.

   For example, one might enter the following information:

   - A2-02-00-00 — Enterprise number 674. It should be written from right to left. The number 674 in decimal notation is 02 a2 00 00.
   - 0a — Data Length (10 decimal)
   - 05 — Sub option code 5
   - 08 — Sub option length (18 decimal)
   - Conversion of the file name from ACSII to hexadecimal

     "host.cfg" - 68-6F-73-74-46-63-66-67

5 Connect a port (OOB port for out-of-band management or any switch port that is a member of VLAN 1 for in-band management) on each switch to the network.

6 Boot the switches.

Easy Firmware Upgrade via USB

If a USB device is detected during bootup and there is an image on the USB device (and no .setup files and no .text files), and the switch has no saved startup config file, then the image version on the USB device is checked against the active image version on the switch. If a newer image version is found on the USB device, the image is copied to the switch backup image and the switch reloads using the new image.

1 Copy the startup-config file to the backup-config,; e.g., **copy startup-config backup-config**.
2 Delete the startup-config file; e.g., **del startup-config**.
3 Put the new image on a cleanly formatted USB stick and insert the USB stick into the stack master.
4 Reboot the stack master and skip the Easy Startup configuration wizard by pressing N when prompted.
5 After the upgrade completes, copy the backup-config to the startup-config, remove the USB stick, and reload the stack. The startup configuration is migrated to the new syntax when loaded into the running-config. Check the running-config, make any necessary adjustments and then save the running-config into the startup-config.

16

Monitoring Switch Traffic

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes sFlow features, Remote Monitoring (RMON), and Port Mirroring features.

The topics covered in this chapter include:

- Traffic Monitoring Overview
- Default Traffic Monitoring Values
- Monitoring Switch Traffic (Web)
- Monitoring Switch Traffic (CLI)
- Traffic Monitoring Examples

Traffic Monitoring Overview

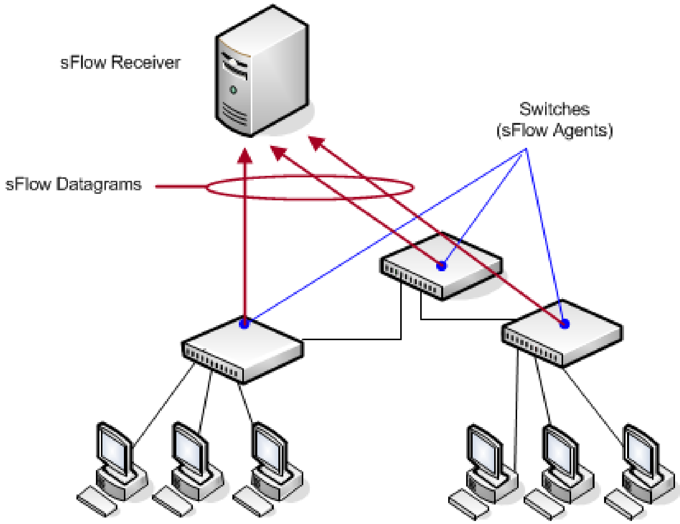
The switch maintains statistics about network traffic that it handles. It also has embedded technology that collects and sends information about traffic to other devices. Dell Networking N-Series switches include support for flow-based monitoring through sFlow and Remote Network Monitoring (RMON) agents.

What is sFlow Technology?

sFlow is an industry standard technology for monitoring high-speed switched and routed networks. Dell Networking N1500, N2000, N3000, and N4000 Series switches software has a built-in sFlow agent that can monitor network traffic on each port and generate sFlow data to an sFlow receiver (also known as a collector). sFlow helps to provide visibility into network activity, which enables effective management and control of network resources. sFlow is an alternative to the NetFlow network protocol, which was developed by Cisco Systems. The switch supports sFlow version 5.

As illustrated in Figure 16-1, the sFlow monitoring system consists of sFlow Agents (such as Dell Networking N-Series switches) and a central sFlow receiver. sFlow Agents use sampling technology to capture traffic statistics from monitored devices. sFlow datagrams forward sampled traffic statistics to the sFlow Collector for analysis. Up to eight different sFlow receivers can be specified to which the switch sends sFlow datagrams.

Figure 16-1. sFlow Architecture



The advantages of using sFlow are:

- It is possible to monitor all ports of the switch continuously, with no impact on the distributed switching performance.
- Minimal memory/CPU is required. Samples are not aggregated into a flow-table on the switch; they are forwarded immediately over the network to the sFlow receiver.
- The sFlow system is tolerant to packet loss in the network because statistical modeling means the loss is equivalent to a slight change in the sampling rate.
- sFlow receiver can receive data from multiple switches, providing a real-time synchronized view of the whole network.
- The receiver can analyze traffic patterns based on protocols found in the headers (e.g., TCP/IP, IPX, Ethernet, AppleTalk…). This alleviates the need for a layer-2 switch to decode and understand all protocols.

sFlow Sampling

The sFlow Agent in the Dell Networking software uses two forms of sampling:

- Statistical packet-based sampling of switched or routed Packet Flows
- Time-based sampling of counters

Packet Flow Sampling and Counter Sampling are performed by sFlow Instances associated with individual Data Sources within an sFlow Agent. Both types of samples are combined in sFlow datagrams. Packet Flow Sampling creates a steady, but random, stream of sFlow datagrams that are sent to the sFlow Collector. Counter samples may be taken opportunistically to fill these datagrams.

To perform Packet Flow Sampling, an sFlow Sampler Instance is configured with a Sampling Rate. Packet Flow sampling results in the generation of Packet Flow Records. To perform Counter Sampling, an sFlow Poller Instance is configured with a Polling Interval. Counter Sampling results in the generation of Counter Records. sFlow Agents collect Counter Records and Packet Flow Records and send them as sFlow datagrams to sFlow Collectors.

Packet Flow Sampling

Packet Flow Sampling, carried out by each sFlow instance, ensures that any packet observed at a Data Source has an equal chance of being sampled, irrespective of the Packet Flow(s) to which it belongs.

Packet Flow Sampling is accomplished as follows:

- A packet arrives on an interface.
- The Network Device makes a filtering decision to determine whether the packet should be dropped.
- If the packet is not filtered (dropped) a destination interface is assigned by the switching/routing function.
- A decision is made on whether or not to sample the packet.

  The mechanism involves a counter that is decremented with each packet. When the counter reaches zero a sample is taken.

- When a sample is taken, the counter indicating how many packets to skip before taking the next sample is reset. The value of the counter is set to a random integer where the sequence of random integers used over time is the Sampling Rate.

Counter Sampling

The primary objective of Counter Sampling is to efficiently, periodically export counters associated with Data Sources. A maximum Sampling Interval is assigned to each sFlow instance associated with a Data Source.

Counter Sampling is accomplished as follows:

- sFlow Agents keep a list of counter sources being sampled.
- When a Packet Flow Sample is generated the sFlow Agent examines the list and adds counters to the sample datagram, least recently sampled first. Counters are only added to the datagram if the sources are within a short period, 5 seconds say, of failing to meet the required Sampling Interval.
- Periodically, say every second, the sFlow Agent examines the list of counter sources and sends any counters that must be sent to meet the sampling interval requirement.

The set of counters is a fixed set.

What is RMON?

Like sFlow, RMON is a technology that enables the collection and analysis of a variety of data about network traffic. Dell Networking N1500, N2000, N3000, and N4000 Series switches software includes an RMON probe (also known as an RMON agent) that collect information and analyze packets. The data that is collected is defined in the RMON MIB, RFC 2819.

RMON is defined in an Internet Engineering Task Force (IETF) specification and is an extension of the SNMP MIB. The RMON information can be viewed locally on the switch or by using a generic RMON console on a network management station (NMS). SNMP does not need to be configured on the switch to view the RMON data locally. However, if you use a management station to view the RMON data that the switch collects and analyzes, you must configure the following SNMP settings:

- Set up the SNMP community string to be used by the SNMP manager at a given IP address.
- Specify the network management system IP address or permit management access from all IP addresses.

For more information about configuring SNMP, see "SNMP " on page 433.

The RMON agent in the switch supports the following groups:

- Group 1—Statistics. Contains cumulative traffic and error statistics.
- Group 2—History. Generates reports from periodic traffic sampling that are useful for analyzing trends.
- Group 3 —Alarm. Enables the definition and setting of thresholds for various counters. Thresholds can be passed in either a rising or falling direction on existing MIB objects, primarily those in the Statistics group. An alarm is triggered when a threshold is crossed and the alarm is passed to the Event group. The Alarm requires the Event Group.
- Group 9 —Event. Controls the actions that are taken when an event occurs. RMON events occur when:
  - A threshold (alarm) is exceeded
  - There is a match on certain filters.

NOTE: The switch supports RMON1.

What is Port Mirroring?

Port mirroring is used to monitor the network traffic that a port sends and receives. The Port Mirroring feature creates a copy of the traffic that the source port handles and sends it to a destination port. The source port is the port that is being monitored. The destination port is where a network protocol analyzer (probe) is connected. Dell Networking N-Series switches also support RSPAN destinations where traffic can be tunneled across the operational network over dedicated links.

A port monitoring session includes one or more source ports that mirror traffic to a single destination port (also known as a probe port). Sources can include VLANs, physical interfaces, port-channels, the internal CPU port, and IP or MAC ACL flows. Certain sources are not supported; i.e., physical members of a port-channel, VLANs that contain a LAG member, etc. Up to four monitoring sessions, each with a unique destination port, may be configured. Destination ports, once configured, no longer participate in spanning tree, IGMP/MLD snooping, or GVRP; do not learn MAC addresses (learned MAC addresses are purged); do not participate in routing (route entries are purged); and do not utilize any static filter configuration. The original configuration of a destination port is restored when the port is no longer configured as a destination port.

Each source port can be configured whether to mirror ingress traffic (traffic the port receives, or RX), egress traffic (traffic the port sends, or TX), or both ingress and egress traffic.

NOTE: A DiffServ policy class definition or an ACL can be created that mirrors specific types of traffic to a destination port. For more information, see "Differentiated Services " on page 1445 or "Access Control Lists " on page 629.

The packet that is copied to the destination port is normally in the same format as the original packet on the wire, except as noted in the following section: Port Mirroring Behaviors. This means that if the mirror port is copying a received packet, the copied packet is VLAN tagged or untagged as it was received on the source port. If the mirror is copying a transmitted packet, the copied packet is VLAN tagged or untagged as it is transmitted on the source port. Destinations include physical interfaces and RSPAN VLANs.

After configuring the port mirroring session, enable or disable the administrative mode of the session to start or stop the probe port from receiving mirrored traffic.

Port Mirroring Behaviors

The following behaviors are applicable to monitor ports:

- The following source port types may be configured in more than one session in support of M:N mirroring:
  - Physical ports
  - LAGs
  - CPU

  VLANs and RSPAN may not be configured as mirroring sources in more than a single session.

- The destination (probe) port loses its VLAN configuration when port mirroring is enabled. The VLAN configuration is restored when the port is no longer configured for a monitor session. The mirrored source and the transit ports retain their VLAN configuration. Transit ports must be members of the RSPAN VLAN.
- When port mirroring is enabled, all MAC address entries associated with destination ports are purged. This prevents transmitting packets out of the port that are not seen on the mirrored port. If spanning tree is enabled, this is treated as a topology change.
- The spanning tree protocol is disabled on destination ports such that frames are always received from or transmitted out of the port as soon as the port is up (spanning tree status is forwarding and role is disabled). This is analogous to always setting the spanning tree state of the port to forwarding. When a port is no longer configured to be the destination port, spanning tree is re-enabled for that port, if configured. Note that the disabling of spanning tree on a destination port means that administrators must only connect the destination port to directly attached probes to avoid the possibility of a network loop.
- GVRP is disabled on destination ports such that GVRP PDUs are never received from or transmitted to the port. Dynamic registrations are not allowed on a destination port. The GVRP configuration at the port is maintained and is reapplied when the port is no longer part of the SPAN.
- All static filters, both source and destination, are disabled on destination ports.
- If routing is enabled on a destination port or an RSPAN VLAN, all route entries associated with that port are purged. From a routing perspective, the interface is marked as down.
- Generally, the configuration of the source port is undisturbed so that its behavior remains the same as if it was not mirrored.
- Packets locally generated by the switch and transmitted over a source port are not copied in a mirroring session.
- The internal CPU port is allowed as a source port for local monitoring sessions only (not allowed for RSPAN). If the internal CPU port is mirrored, packets received and generated by the CPU for all ports are mirrored.

- On ingress, the port mirroring logic stage is after the VLAN tag processing stage in the hardware. This means that mirrored packets may not appear the same as they do on the wire if VLAN tag processing occurs. Examples of VLAN tag processing are DVLAN tunneling (QinQ) or VLAN rewriting. Likewise, on egress, the port mirroring logic stage is before the VLAN tag processing stage. This means that on egress, packets may not appear as they do on the wire if processing such as VLAN or CoS value rewriting is programmed.

Remote Capture

The Remote Capture feature enables mirroring packets transmitted and received by the switch CPU to a remote client for packet analysis using the Wireshark tool. This feature can be used to help diagnose switch behavior or monitor traffic sent to the switch CPU. The capture feature can also be configured to capture to a local file or to an in-memory buffer.

Why is Traffic Monitoring Needed?

Monitoring the traffic that the switch handles, as well as monitoring all traffic in the network, can help provide information about network performance and utilization. This information can be useful in network planning and resource allocation. Information about traffic flows can also help troubleshoot problems in the network.

Default Traffic Monitoring Values

The sFlow agent is enabled by default, but sampling and polling are disabled on all ports. Additionally, no sFlow receivers (collectors) are configured. Table 16-1 contains additional default values for the sFlow feature.

Table 16-1. sFlow Defaults

Parameter	Default Value
Receiver timeout for sampling	0
Receiver port	6343
Receiver Maximum Datagram Size	1400 bytes
Maximum header size	128 bytes

RMON is enabled by default, but no RMON alarms, events, or history statistic groups are configured.

Port mirroring is disabled, and no ports are configured as source or destination ports. After you configure a port mirroring session, the administrative mode is disabled until you explicitly enable it.

Monitoring Switch Traffic (Web)

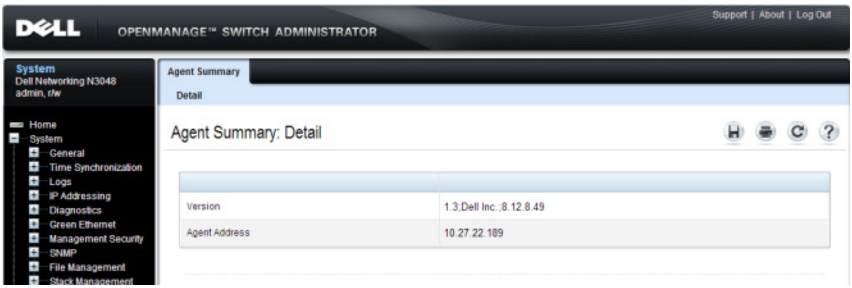
This section provides information about the OpenManage Switch Administrator pages to use to monitor network traffic on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

sFlow Agent Summary

Use the sFlow **Agent Summary** page to view information about sFlow MIB and the sFlow Agent IP address.

To display the **Agent Summary** page, click **System** → **sFlow** → **Agent Summary** in the navigation panel.

Figure 16-2. sFlow Agent Summary

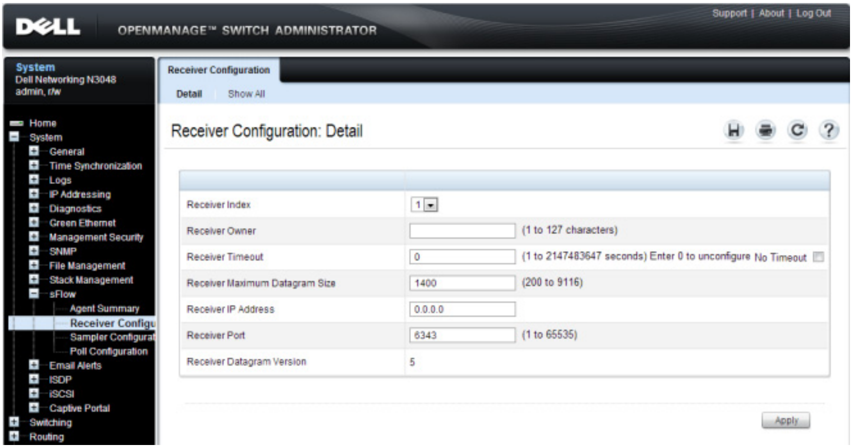


sFlow Receiver Configuration

Use the sFlow **Receiver Configuration** page to configure settings for the sFlow receiver to which the switch sends sFlow datagrams. Up to eight sFlow receivers can be configured to receive datagrams.

To display the **Receiver Configuration** page, click **System** → **sFlow** → **Receiver Configuration** in the navigation panel.

Figure 16-3. sFlow Receiver Configuration



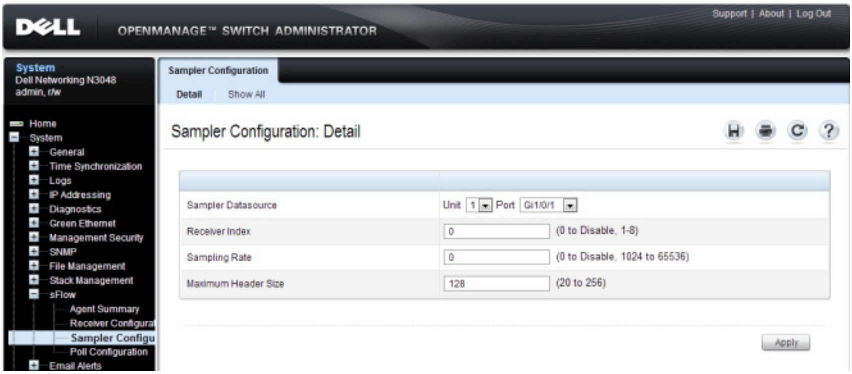
Click **Show All** to view information about configured sFlow receivers.

sFlow Sampler Configuration

Use the sFLow **Sampler Configuration** page to configure the sFlow sampling settings for switch ports.

To display the **Sampler Configuration** page, click **System** → **sFlow** → **Sampler Configuration** in the navigation panel.

Figure 16-4. sFlow Sampler Configuration



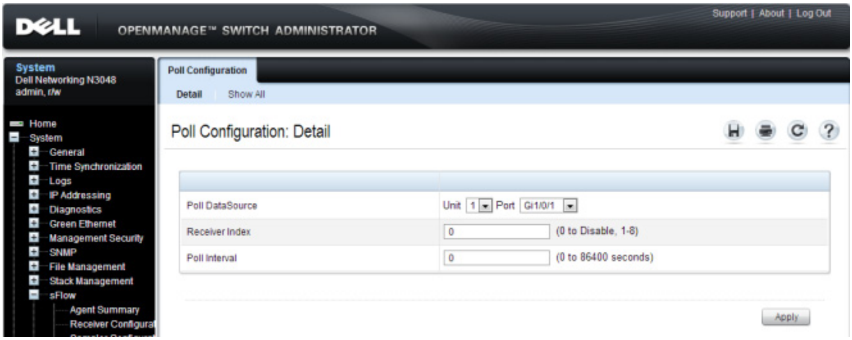
Click **Show All** to view information about configured sampler data sources.

sFlow Poll Configuration

Use the sFLow **Poll Configuration** page to configure how often a port should collect counter samples.

To display the **Sampler Configuration** page, click **System** → **sFlow** → **Sampler Configuration** in the navigation panel.

Figure 16-5. sFlow Poll Configuration



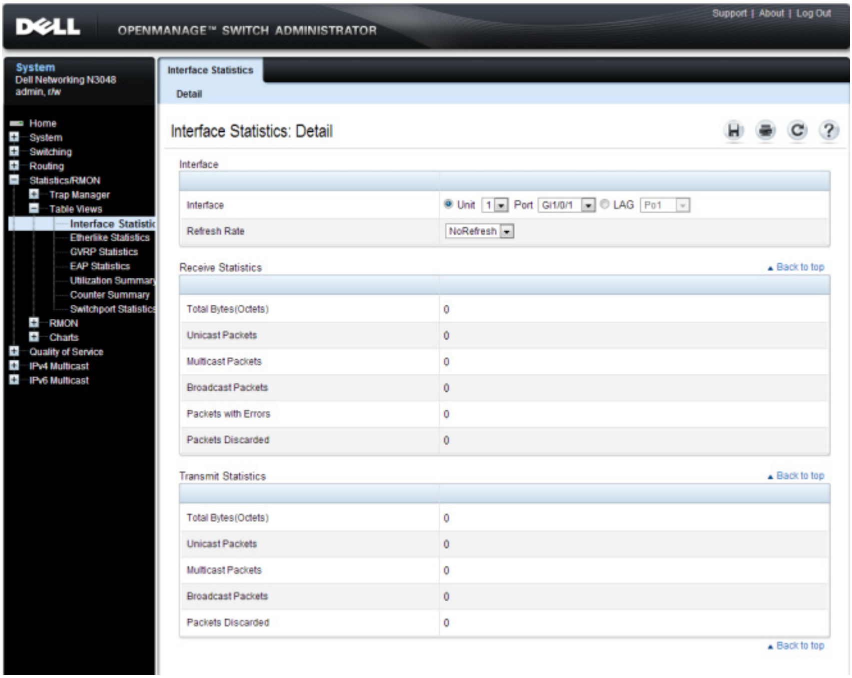
Click **Show All** to view information about the ports configured to collect counter samples.

Interface Statistics

Use the **Interface Statistics** page to display statistics for both received and transmitted packets. The fields for both received and transmitted packets are identical.

To display the page, click **Statistics/RMON** → **Table Views** → **Interface Statistics** in the navigation panel.

Figure 16-6. Interface Statistics

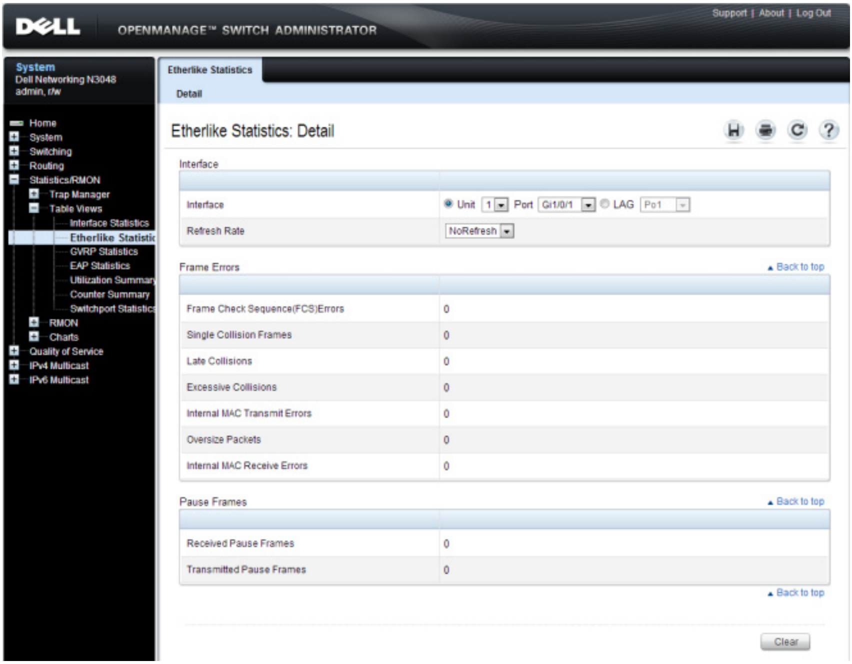


Etherlike Statistics

Use the **Etherlike Statistics** page to display interface statistics.

To display the page, click **Statistics/RMON** → **Table Views** → **Etherlike Statistics** in the navigation panel.

Figure 16-7. Etherlike Statistics

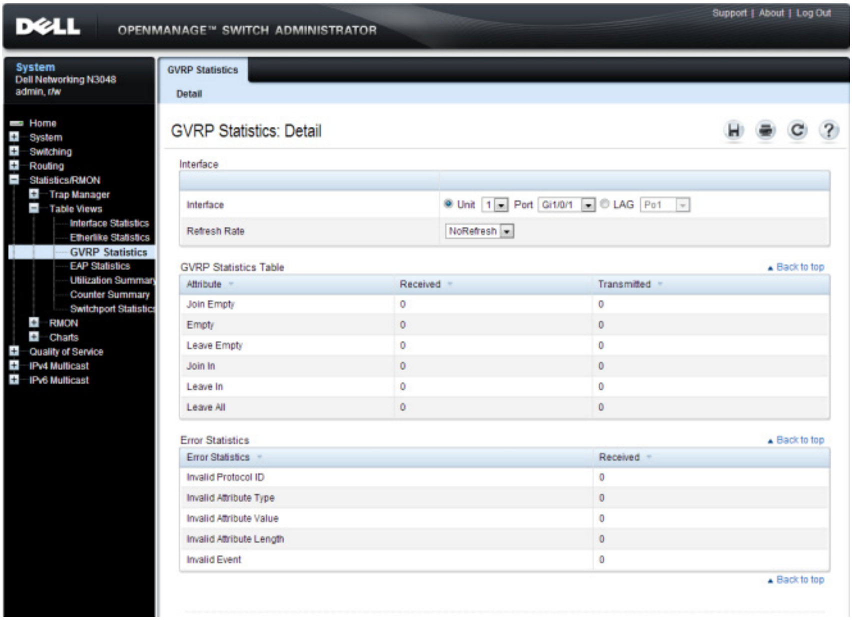


GVRP Statistics

Use the **GVRP Statistics** page to display switch statistics for GVRP.

To display the page, click **Statistics/RMON → Table Views → GVRP Statistics** in the navigation panel.

Figure 16-8. GVRP Statistics

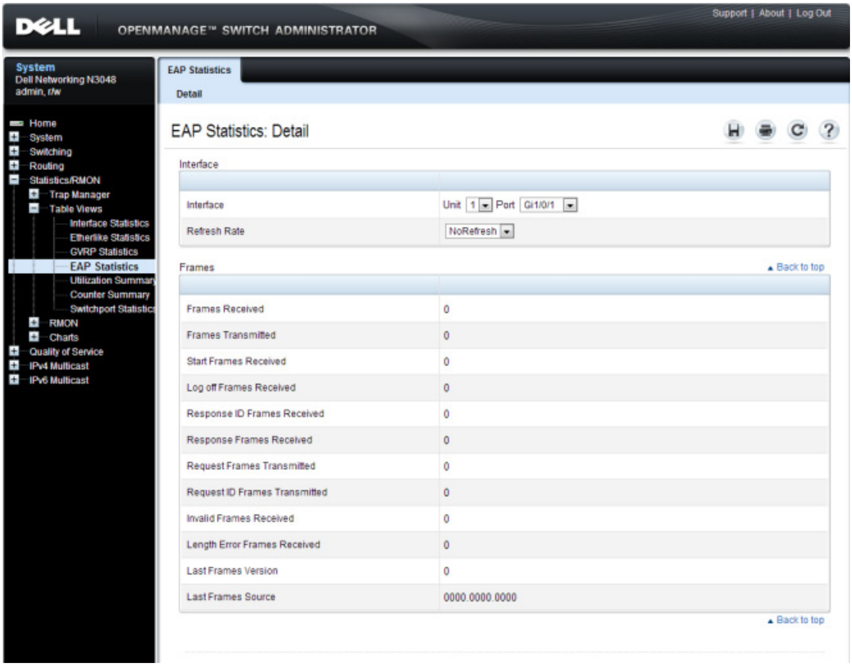


EAP Statistics

Use the **EAP Statistics** page to display information about EAP packets received on a specific port. For more information about EAP, see "Port and System Security " on page 623.

To display the **EAP Statistics** page, click **Statistics/RMON** → **Table Views** → **EAP Statistics** in the navigation panel.

Figure 16-9. EAP Statistics

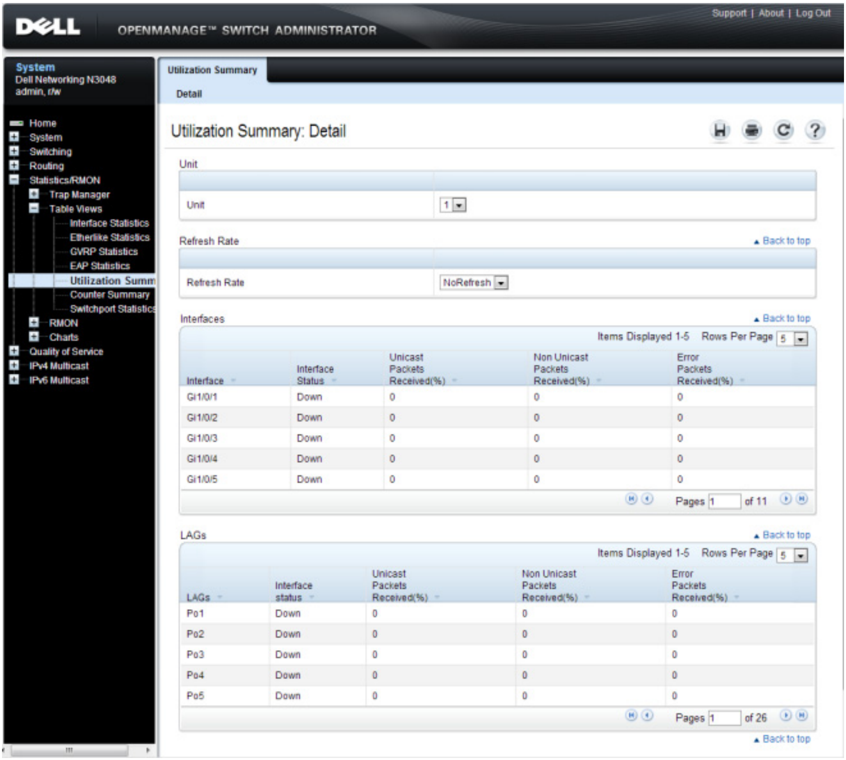


Utilization Summary

Use the **Utilization Summary** page to display interface utilization statistics.

To display the page, click **Statistics/RMON** → **Table Views** → **Utilization Summary** in the navigation panel.

Figure 16-10. Utilization Summary

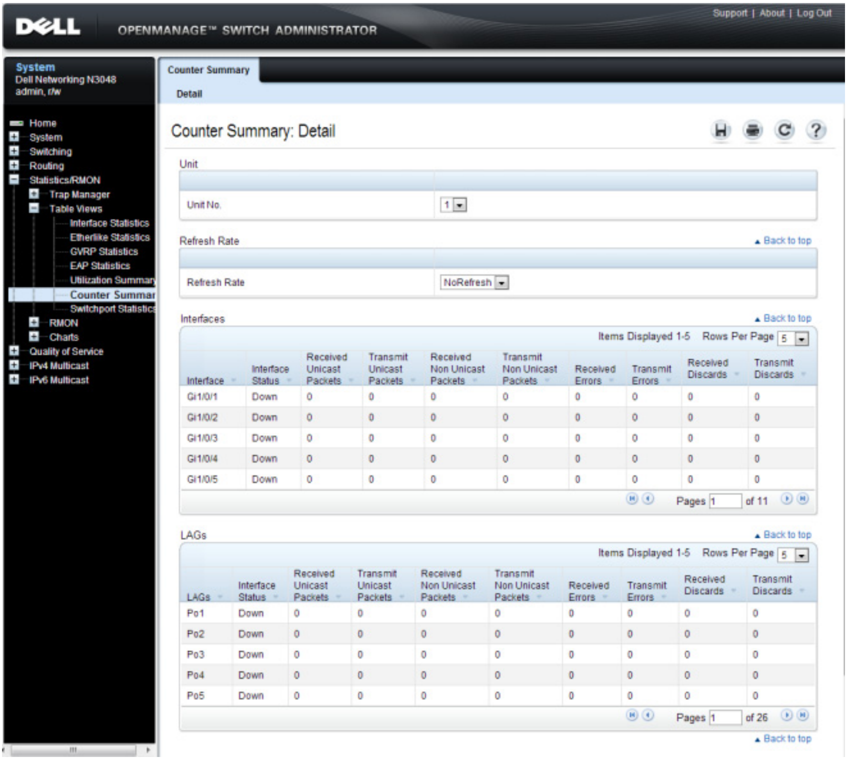


Counter Summary

Use the **Counter Summary** page to display interface utilization statistics in numeric sums as opposed to percentages.

To display the page, click **Statistics/RMON** → **Table Views** → **Counter Summary** in the navigation panel.

Figure 16-11. Counter Summary

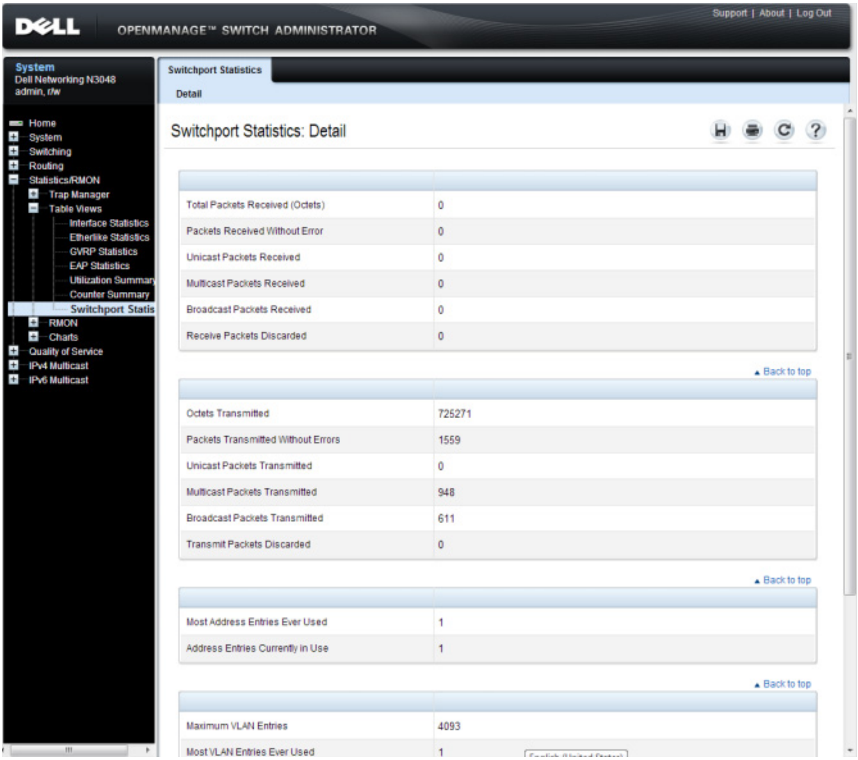


Switchport Statistics

Use the **Switchport Statistics** page to display statistical summary information about switch traffic, address tables, and VLANs.

To display the page, click **Statistics/RMON** → **Table Views** → **Switchport Statistics** in the navigation panel.

Figure 16-12. Switchport Statistics

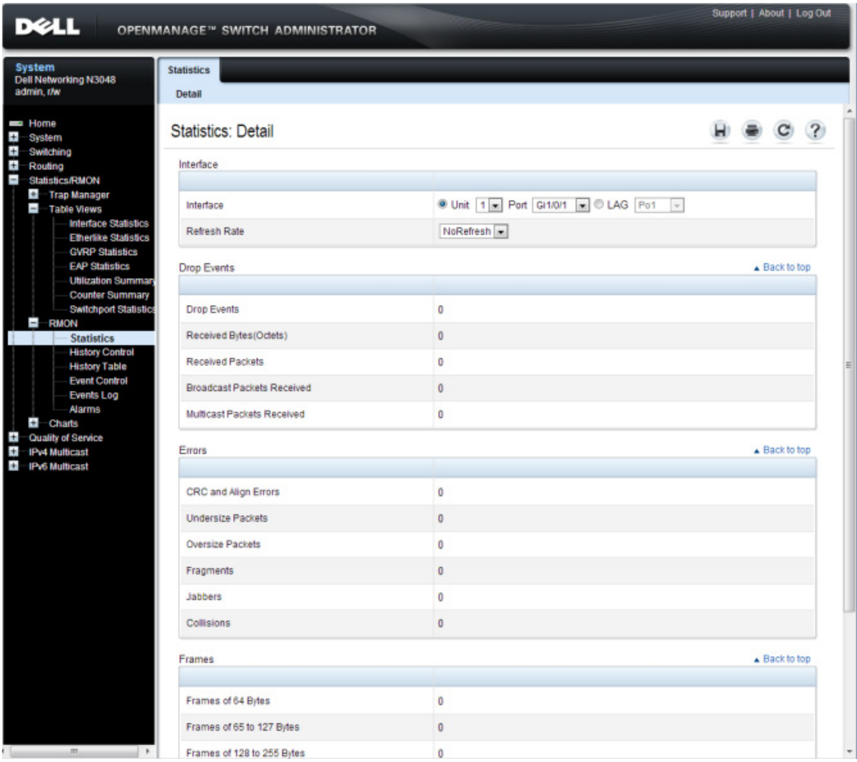


RMON Statistics

Use the **RMON Statistics** page to display details about switch use such as packet processing statistics and errors that have occurred on the switch.

To display the page, click **Statistics/RMON** → **RMON** → **Statistics** in the navigation panel.

Figure 16-13. RMON Statistics

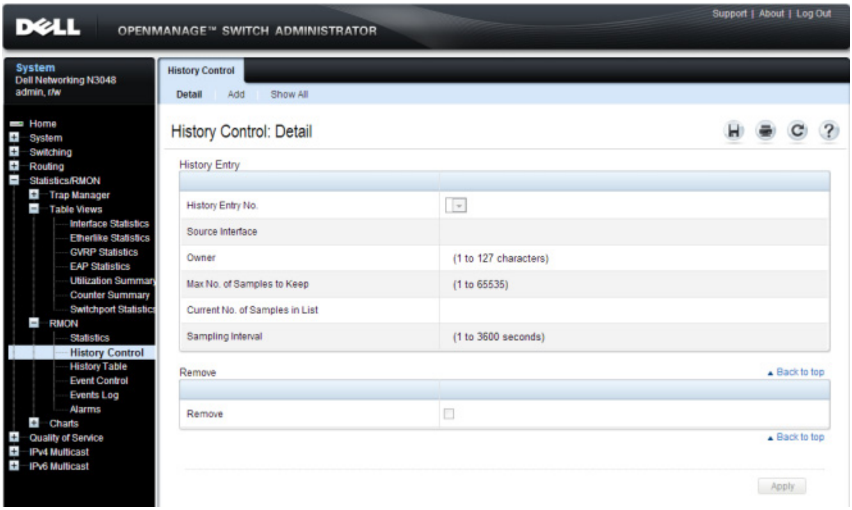


RMON History Control Statistics

Use the **RMON History Control** page to maintain a history of statistics on each port. For each interface (either a physical port or a port-channel), the number of buckets and the time interval between each bucket snapshot can be configured.

To display the page, click **Statistics/RMON** → **RMON** → **History Control** in the navigation panel.

Figure 16-14. RMON History Control



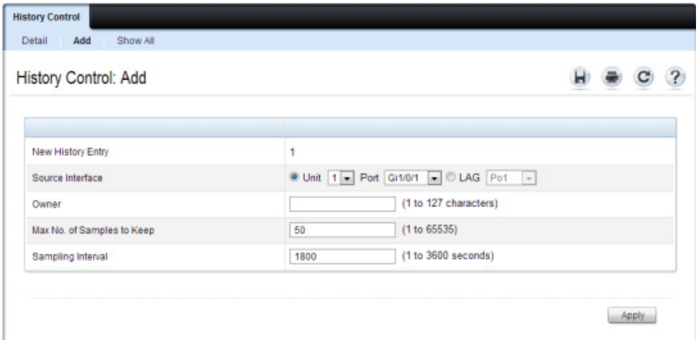
Adding a History Control Entry

To add an entry:

1. Open the **RMON History Control** page.
2. Click **Add**.

   The **Add History Entry** page displays.

Figure 16-15. Add History Entry



3 Select the port or LAG on which you want to maintain a history of statistics.

4 Specify an owner, the number of historical buckets to keep, and the sampling interval.

5 Click **Apply** to add the entry to the **RMON History Control Table**.

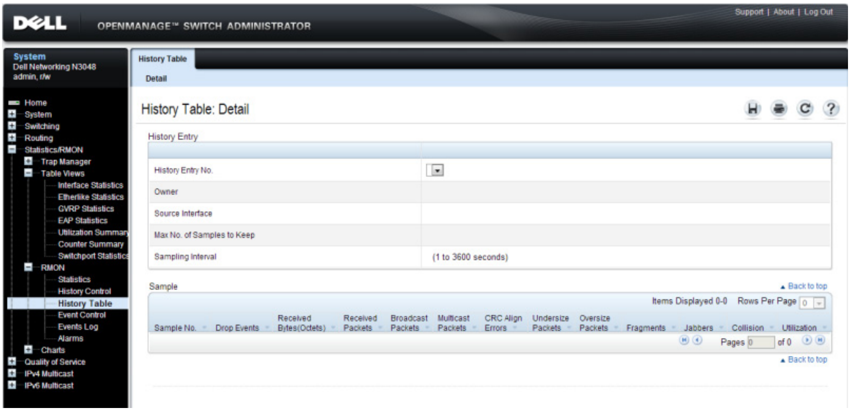
To view configured history entries, click the **Show All** tab. The **RMON History Control Table** displays. Configured history entries can be removed using this page.

RMON History Table

Use the RMON History Table page to display interface-specific statistical network samplings. Each table entry represents all counter values compiled during a single sample.

To display the **RMON History Table** page, click **Statistics/RMON** → **RMON** → **History Table** in the navigation panel.

Figure 16-16. RMON History Table

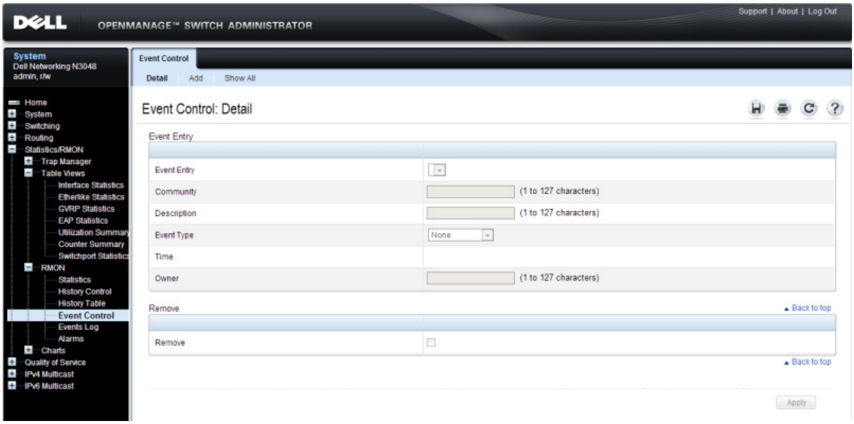


RMON Event Control

Use the **RMON Events Control** page to define RMON events. Events are used by RMON alarms to force some action when a threshold is crossed for a particular RMON counter. The event information can be stored in a log and/or sent as a trap to a trap receiver.

To display the page, click **Statistics/RMON** → **RMON** → **Event Control** in the navigation panel.

Figure 16-17. RMON Event Control



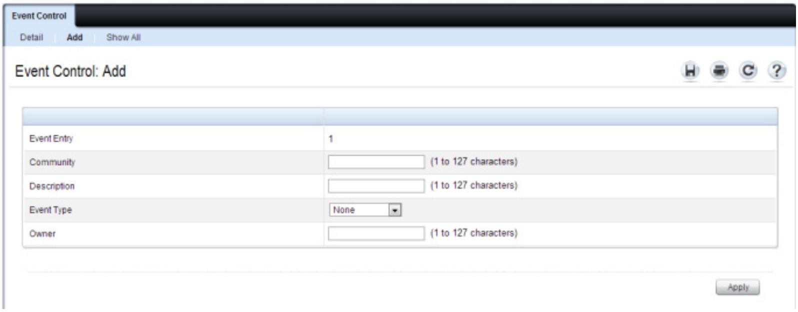
Adding an RMON Event

To add an event:

1. Open the **RMON Event Control** page.
2. Click **Add**.

   The **Add an Event Entry** page displays.

Figure 16-18. Add an Event Entry



3. If the event sends an SNMP trap, specify the SNMP community to receive the trap.
4. Optionally, provide a description of the event and the name of the event owner.
5. Select an event type.
6. Click **Apply**.

   The event is added to the **RMON Event Table**, and the device is updated.

Viewing, Modifying, or Removing an RMON Event

To manage an event:

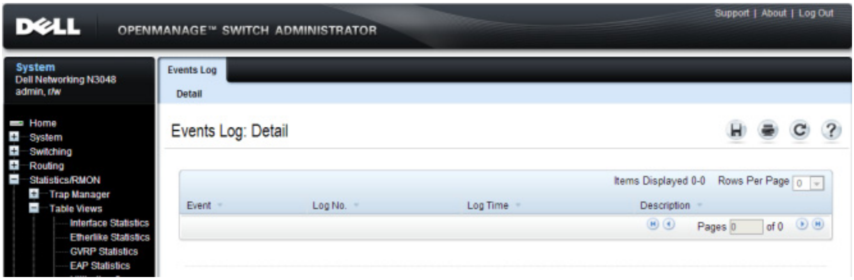
1. Open the **RMON Event Control** page.
2. Click **Show All** to display the **Event Control Table** page.
3. To edit an entry:
   a. Select the **Edit** check box in for the event entry to change.
   b. Modify the fields on the page as needed.
4. To remove an entry, select the **Remove** check box in for the event entry to remove.
5. Click **Apply**.

RMON Event Log

Use the **RMON Event Log** page to display a list of RMON events.

To display the page, click **Statistics/RMON** → **RMON** → **Events Log** in the navigation panel.

Figure 16-19. RMON Event Log

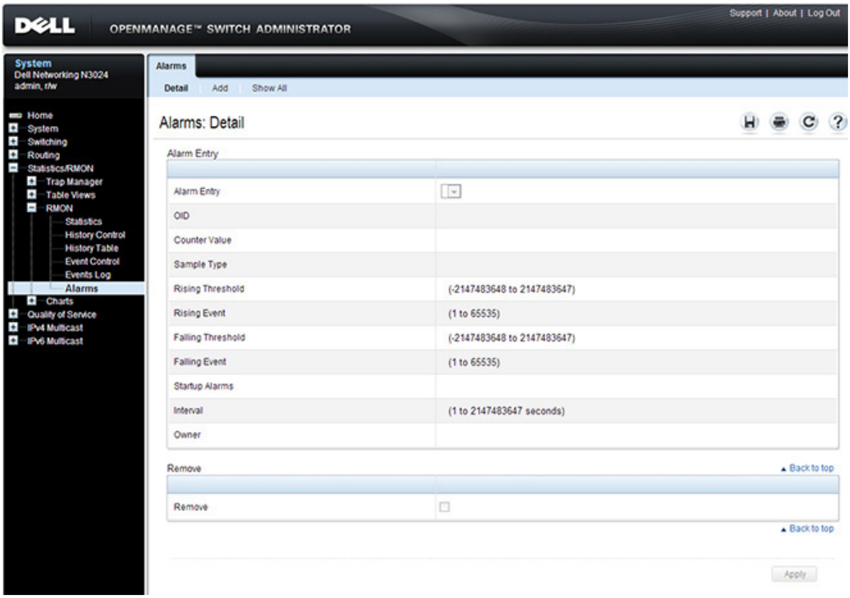


RMON Alarms

Use the **RMON Alarms** page to set network alarms. Alarms occur when certain thresholds are crossed for the configured RMON counters. The alarm triggers an event to occur. The events can be configured as part of the RMON Events group. For more information about events, see "RMON Event Log " on page 544.

To display the page, click **Statistics/RMON** → **RMON** → **Alarms** in the navigation panel.

Figure 16-20. RMON Alarms



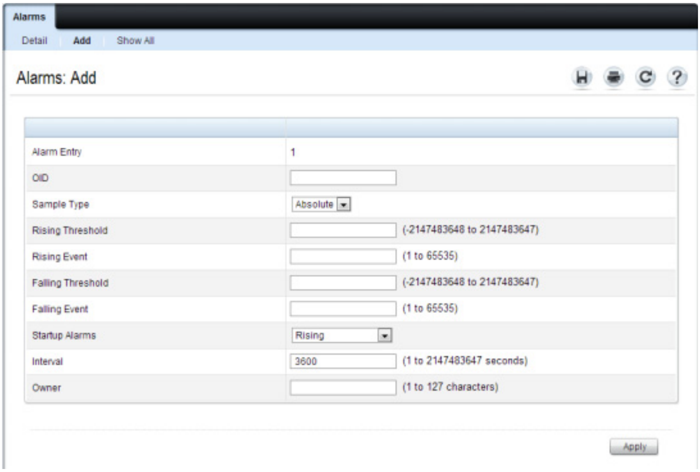
Adding an Alarm Table Entry

To add an alarm:

1. Open the **RMON Alarms** page.
2. Click **Add.**

   The **Add an Alarm Entry** page displays.

Figure 16-21. Add an Alarm Entry



3. Complete the fields on this page as needed. Use the help menu to learn more information about the data required for each field.
4. Click **Apply**.

   The RMON alarm is added, and the device is updated.

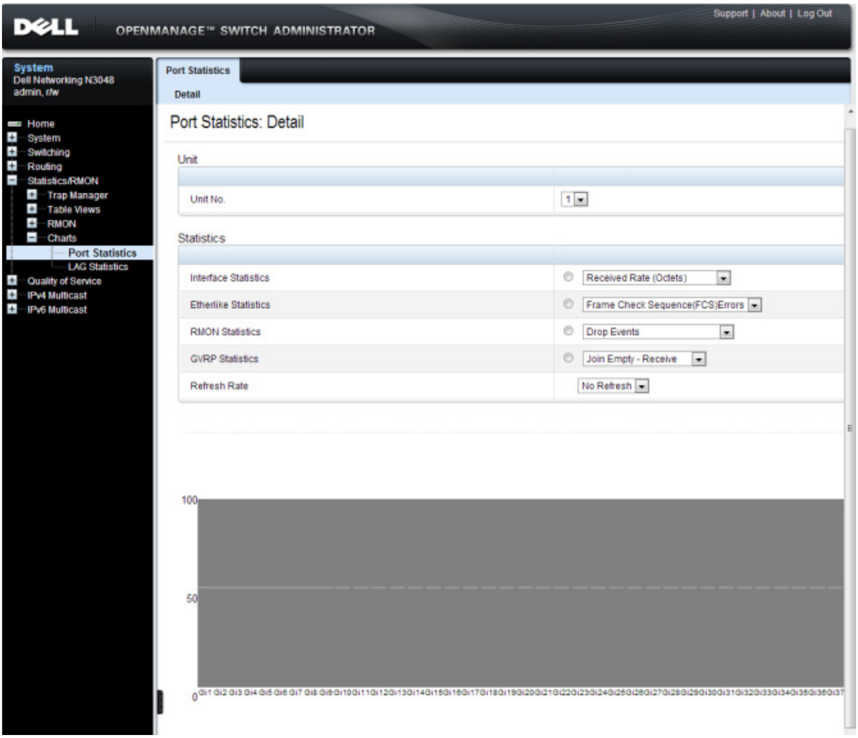
To view configured alarm entries, click the **Show All** tab. The **Alarms Table** displays. Configured alarms can be removed using this page.

Port Statistics

Use the **Port Statistics** page to chart port-related statistics on a graph.

To display the page, click **Statistics/RMON** → **Charts** → **Port Statistics** in the navigation panel.

Figure 16-22. Ports Statistics



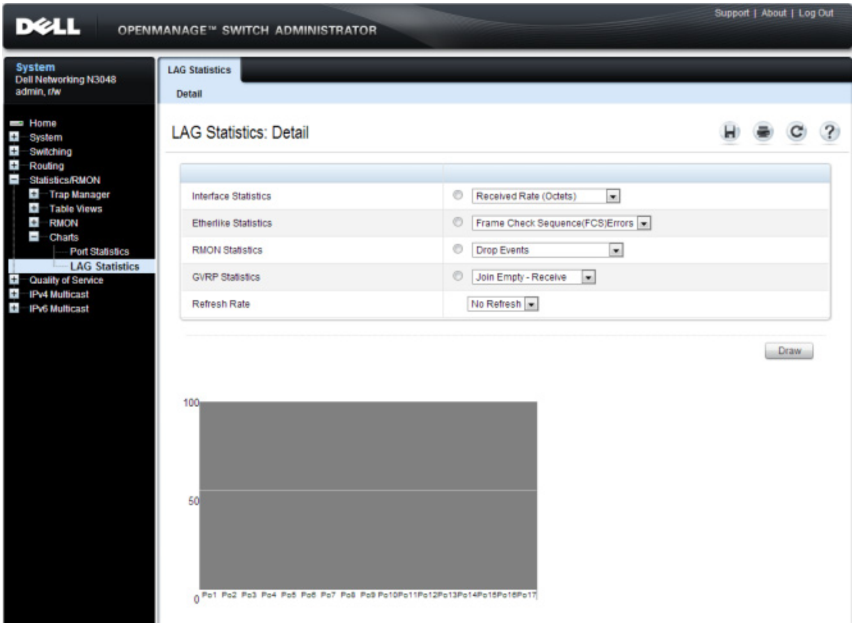
To chart port statistics, select the type of statistics to chart and (if desired) the refresh rate, then click **Draw**.

LAG Statistics

Use the **LAG Statistics** page to chart LAG-related statistics on a graph.

To display the page, click **Statistics/RMON** → **Charts** → **LAG Statistics** in the navigation panel.

Figure 16-23. LAG Statistics



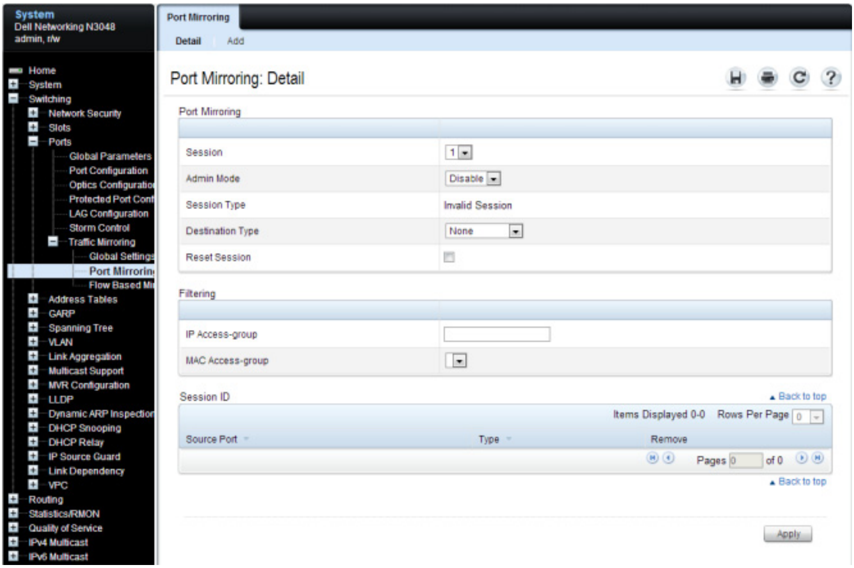
To chart LAG statistics, select the type of statistics to chart and (if desired) the refresh rate, then click **Draw**.

Port Mirroring

Use the **Port Mirroring** page to create a mirroring session in which all traffic that is sent or received (or both) on one or more source ports is mirrored to a destination port.

To display the **Port Mirroring** page, click **Switching** → **Ports** → **Traffic Mirroring** → **Port Mirroring** in the navigation panel.

Figure 16-24. Port Mirroring



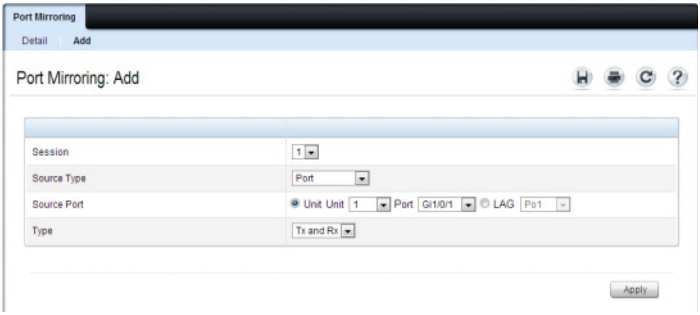
Configuring a Port Mirror Session

To configure port mirroring:

1. Open the **Port Mirroring** page.
2. Click **Add**.

   The **Add Source Port** page displays.
3. Select the port to be mirrored.
4. Select the traffic to be mirrored.

Figure 16-25. Add Source Port



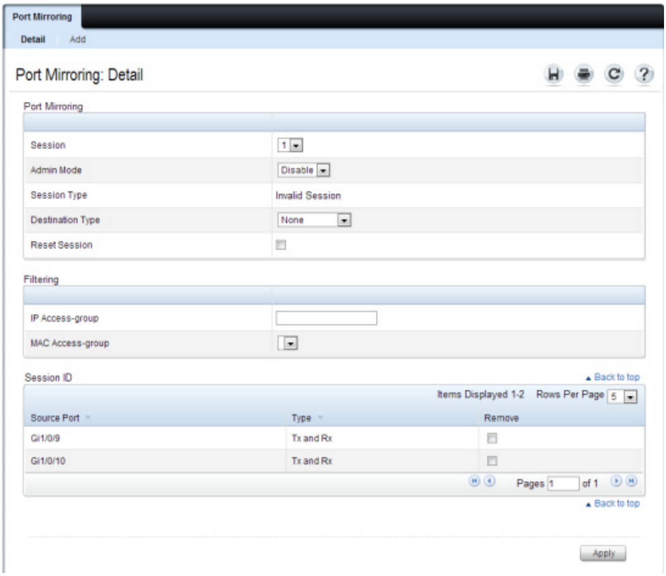
5 Click **Apply**.

6 Repeat the previous steps to add additional source ports.

7 Click **Port Mirroring** to return to the **Port Mirroring** page.

8 Enable the administrative mode and specify the destination port.

Figure 16-26. Configure Additional Port Mirroring Settings



9 Click **Apply**.

Monitoring Switch Traffic (CLI)

This section provides information about the commands you use to manage traffic monitoring features on the switch and to view information about switch traffic. For more information about these commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring sFlow

Beginning in Privileged EXEC mode, use the following commands to configure the sFlow receiver and to configure the sampling and polling on switch interfaces.

Command	Purpose
configure	Enter Global Configuration mode
sflow *rcvr_index* **destination** *ip-address* [*port*]	Configure the address of the sFlow receiver and (optionally) the destination UDP port for sFlow datagrams. • *rcvr_index*—The index of this sFlow receiver (Range: 1–8). • *ip-address*—The sFlow receiver IP address. • *port* —The destination layer-3 UDP port for sFlow datagrams. (Range: 1–65535).
sflow *rcvr_index* **destination owner** *owner_string* **timeout** *timeout*	Specify the identity string of the receiver and set the receiver timeout value. *timeout*—The number of seconds the configuration will be valid before it is automatically cleared. A value of 0 essentiality means the receiver is not configured.
sflow *rcvr_index* **maxdatagram** *size*	Specify the maximum number of data bytes that can be sent in a single sample datagram. The receiver should also be set to this value to avoid fragmentation of the sFlow datagrams. (Range: 200–9116 bytes).

Command	Purpose
sflow *rcvr-index* **polling** *if_type if_number poll-interval*	Enable a new sFlow poller instance on an interface range. • *rcvr-index* — The sFlow Receiver associated with the poller (Range: 1–8). • *if_type if_number* — The list of interfaces to poll. The interface type can be Gigabitethernet (gi) or Tengigabitethernet (te), for example te1/0/3-5 enables polling on ports 3, 4, and 5. • *poll-interval* — The sFlow instance polling interval. A value of *n* means once in *n* seconds a counter sample is generated. (Range: 0–86400).
sflow *rcvr-index* **sampling** *if_type if_number sampling-rate* [*size*]	Enable a new sflow sampler instance for the specified interface range. • *rcvr-index* — The sFlow Receiver for this sFlow sampler to which flow samples are to be sent. • *if_type if_number* — The list of interfaces to sample. The interface type can be Gigabitethernet (gi) or Tengigabitethernet (te), for example te1/0/3-5 enables polling on ports 3, 4, and 5. • *sampling-rate* — The statistical sampling rate for packet sampling from this source. A sampling rate of 1 counts all packets. A value of *n* means that out of *n* incoming packets, 1 packet will be sampled. (Range: 1024 - 65536). • *size* — The maximum number of bytes that should be copied from the sampler packet (Range: 20 - 256 bytes).
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3** or **te1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
sflow *rcvr-index* **polling** *poll-interval*	Enable a new sFlow poller instance for the interface.

Command	Purpose
sflow *rcvr-index* **sampling** *sampling-rate* [*size*]	Enable a new sflow sampler instance for the interface.
CTRL + Z	Exit to Privileged Exec mode.
show sflow agent	View information about the switch sFlow agent.
show sflow *index* **destination**	View information about a configured sFlow receivers.
show sflow *index* **polling**	View information about the configured sFlow poller instances for the specified receiver.
show sflow *index* **sampling**	View information about the configured sFlow sampler instances for the specified receiver.

Configuring RMON

Beginning in Privileged EXEC mode, use the following commands to configure RMON alarms, collection history, and events. The table also lists the commands you use to view information collected by the RMON probe.

Command	Purpose
configure	Enter Global Configuration mode
rmon event *number* [log] [**trap** *community*] [**description** *string*] [**owner** *string*]	Configure an RMON event. • *number* — The event index. (Range: 1–65535) • **log** — Specify that an entry is made in the log table for each event. • **trap** *community* — If the event is an SNMP trap to be sent, it is sent to the SNMP community specified by this octet string. (Range: 0-127 characters) • **description** *string* — A comment describing this event. (Range 0-127 characters) • **owner** *string* — Enter a name that specifies who configured this event. If unspecified, the name is an empty string.

Command	Purpose
rmon alarm *number variable interval* {**absolute** |**delta**} **rising-threshold** *value* [*event-number*] **rising-threshold** *value* [*event-number*] [**startup** *direction*] [**owner** *string*]	Add an alarm entry • *number* — The alarm index. (Range: 1–65535) • *variable* — A fully qualified SNMP object identifier that resolves to a particular instance of an MIB object. • *interval* — The interval in seconds over which the data is sampled and compared with the rising and falling thresholds. (Range: 1–4294967295) • **rising-threshold** *value* — Rising threshold value. (Range: 0–4294967295) • **rising-threshold** *value* — Falling threshold value. (Range: 0–4294967295) • event-number — The index of the event that is used when a rising or falling threshold is crossed. (Range: 1-65535) • **delta** — The sampling method for the selected variable and calculating the value to be compared against the thresholds. If the method is delta, the selected variable value at the last sample is subtracted from the current value, and the difference compared with the thresholds. • **absolute** — The sampling method for the selected variable and calculating the value to be compared against the thresholds. If the method is absolute, the value of the selected variable is compared directly with the thresholds at the end of the sampling interval. • **startup** *direction* — The type of startup alarm, which can be **rising**, **falling**, or **rising-falling**. • **owner** *string* — Enter a name that specifies who configured this alarm.
interface *interface*	Enter Interface Configuration mode for the specified port or LAG.

Command	Purpose
rmon collection history *index* [**owner** *ownername*] [**buckets** *bucket-number*] [**interval** *seconds*]	Enable an RMON MIB history statistics group on the interface. **NOTE:** You must configure RMON alarms and events before RMON collection history is able to display. • *index* — The requested statistics index group. (Range: 1–65535) • *ownername* — Records the RMON statistics group owner name. If unspecified, the name is an empty string. • *bucket-number* — A value associated with the number of buckets specified for the RMON collection history group of statistics. If unspecified, defaults to 50. (Range: 1 - 65535) • *seconds* — The number of seconds in each polling cycle. If unspecified, defaults to 1800. (Range: 1–3600)
CTRL + Z	Exit to Privileged EXEC mode.
show rmon {alarms |collection history | events | history | log | statistics}	View information collected by the RMON probe.

Viewing Statistics

Use the following commands in Privileged EXEC mode to view statistics about the traffic handled by the switch.

Command	Purpose
show interfaces counters [**errors**] [{*interface* | *port-channel*}]	Display the error counters or number of octets and packets handled by all interfaces or the specified interface.
show statistics {**switchport** | *interface*}	Display detailed statistics for a specific port or LAG, or for the entire switch. The *interface* variable includes the interface type and number.
show interfaces utilization [*interface-id*]	Display the TX and RX link utilization (frame rate and bits per second).

Command	Purpose
show interfaces traffic [*interface-id*]	Display the current TX and RX queue congestion and congestion discards.

Configuring Port Mirroring

Use the following commands in Privileged EXEC mode to configure a port mirroring session.

Command	Purpose
configure	Enter Global Configuration mode
monitor session *session_number* **source interface** {*interface-id*} [**rx** | **tx** | **both**]	Configure a source (monitored) port or CPU interface for a monitor session. • *session_number* —The monitoring session ID, which ranges from 1 to 4. The Dell Networking N1500 supports a single session. • *interface-id*—The interface to be monitored. • rx | tx — Monitor ingress (rx) or egress (tx) traffic. If no parameter is given, both ingress and egress traffic are monitored.
monitor session *session_number* **destination interface** *interface-id*	Configure a destination (probe) port for a monitor session. • *session_number* —The monitoring session ID, which ranges from 1 to 4. The Dell Networking N1500 supports a single session. • *interface*—The Ethernet interface to which the monitored source traffic is copied.
monitor session *session_number* **mode**	Enable the administrative mode for the configured port mirroring session to start sending the traffic from the source port to the destination (probe) port.
exit	Exit to Privileged EXEC mode.
show monitor session *session_number*	View information about the configured port mirroring session.

Configuring RSPAN

RSPAN is an extension of port mirroring that operates across multiple switches. Use the following commands in Privileged EXEC mode to configure RSPAN. Remember to assign VLANs to physical interfaces (steps not shown).

Configuring RSPAN (Source Switch)

Command	Purpose
configure	Enter Global Configuration mode.
vlan *vlan-id*	Configure an RSPAN VLAN.
remote-span	Configure the VLAN as a spanning VLAN.
exit	Exit to Global Configuration mode.
interface te1/0/1	Enter interface configuration mode.
switchport mode trunk	Set the egress span interface to trunk mode.
switch trunk allowed vlan *vlan-id*	Restrict the trunk to the spanning VLAN (optional).
exit	Exit to Global Configuration mode.
monitor session *session-number* **source** {**interface** *interface–id* | **vlan** *vlan-id* | **remote vlan** *rspan-vlan-id*} [**rx** | **tx**]	Configure a source (monitored) port for a monitor session. • *session_number* —The monitoring session ID, which ranges from 1 to 4. The Dell Networking N1500 supports a single session. • *interface–id* —The interface to be monitored. The internal CPU port may not be configured as an RSPAN source. • **rx | tx** — Monitor ingress (rx) or egress (tx) traffic. If no parameter is given, both ingress and egress traffic are monitored.
monitor session *session-number* **destination** {**interface** *interface–id* | **remote vlan** *rspan-vlan-id* **reflector-port** *interface–id*}	Configure a local RSPAN reflector port on the source switch. The reflector port should be configured as a trunk port.

Command	Purpose
monitor session *session_number* **mode**	Enable the administrative mode for the configured port mirroring session to start sending the traffic from the source port to the destination (probe) port.
exit	Exit to Privileged EXEC mode.

Configuring RSPAN (Transit Switch)

Command	Purpose
configure	Enter Global Configuration mode.
vlan *vlan-id*	Create an RSPAN VLAN.
remote-span	Configure the VLAN as a spanning VLAN.
exit	Exit to Global Configuration mode.
interface range te1/0/1-2	Configure the span interfaces.
switchport mode trunk	Configure the interface to be in trunking mode.
switch trunk allowed vlan *vlan-id*	Restrict the trunk to the spanning VLAN (optional).
exit	Exit to Global Configuration mode.

Configuring RSPAN (Destination Switch)

Command	Purpose
configure	Enter Global Configuration mode.
vlan *vlan-id*	Create an RSPAN VLAN.
remote-span	Configure the VLAN as a spanning VLAN.
exit	Exit to Global Configuration mode
monitor session *session_id* **source remote vlan** *vlan_id*	Configure a source RSPAN VLAN on the destination switch.
monitor session *session_id* **destination interface** *interface*	Configure the destination port on the RSPAN destination switch.

Command	Purpose
monitor session *session_id* **mode**	Enable the monitor session.

Traffic Monitoring Examples

This section contains the following examples:

- Showing Interface Traffic
- Configuring sFlow
- Configuring RMON
- Configuring Remote Capture
- Configuring RSPAN

Showing Interface Traffic

Use the **show interfaces utilization** and **show interfaces traffic** commands to display information about interface traffic and internal packet buffer usage. The following are examples of the output of these commands. Refer the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* for more details about the output.

This example shows Gi1/0/1 is suffering from congestion (Tx Queue high) and is dropping packets, either due to WRED drops or due to exceeding the internal buffer limits.

```
console#show interfaces traffic
Intf   Congestion  Tx Queue Rx Queue   Color Drops (Pkts)  Tx Queue
Name   Drops (Pkts) (Cells)  (Cells)   Yellow   Red        (Pkts)
------ ----------- -------- --------   -------- --------- --------
Gi1/0/1  18981     132      0          0        0         13
Gi1/0/2  0         0        0          0        0         0
Gi1/0/3  0         0        0          0        0         0
```

The following example shows a classical incast situation on interface Gi1/0/2 where the port is fully utilized or nearly fully utilized, buffering many frames (with increased latency) and beginning to drop frames as the internal thresholds for buffering on the port are reached. A conscientious network operator might want to examine why the devices attached to Gi1/0/5 and Gi1/0/6 are sending so much traffic to Gi1/0/2 attached devices and redistribute the devices, rate-limit traffic egressing the devices attached to Gi1/0/5 and Gi1/0/6, or increase the number of links available for the device attached to Gi1/0/2.

```
console#show interfaces utilization

Port     Load     Oper.  Rx    Tx     Rx     Tx     Buffer  Drop
         Interval Speed  Util  Util   PPS    PPS    Size    Count
------- -------- ------ ----- ----- ------ ------ ------- ---------
Gi1/0/1  300      10M    1     0      296    0      0       0
Gi1/0/2  300      1G     0     99     0      674500 938098  1102
Gi1/0/3  300      1G     0     15     0      112428 7       0
Gi1/0/4  300      0      0     0      0      1      0       0
Gi1/0/5  300      1G     37    0      249565 1      0       1
Gi1/0/6  300      1G     88    1      593560 3      0       0
Gi1/0/7  300      0      0     0      0      0      0       0
Gi1/0/8  300      0      0     0      0      1      0       0
```

Configuring sFlow

This example shows how to configure the switch so that ports 10-15 and port 23 send sFlow datagrams to an sFlow receiver at the IP address 192.168.20.34. The receiver owner is receiver1, and the timeout is 100000 seconds. A counter sample is generated on the ports every 60 seconds (polling interval), and 1 out of every 8192 packets is sampled. Note that sFlow monitoring is not enabled until a receiver owner string is configured.

To configure the switch:

1. Configure information about the sFlow receiver.

```
console#configure
console(config)#sflow 1 destination 192.168.30.34
console(config)#sflow 1 destination owner receiver1 timeout
100000
```

2. Configure the polling and sampling information for Tengigabit Ethernet ports 10-20.

```
console(config)#sflow 1 polling te1/0/10-15 60
console(config)#sflow 1 sampling te1/0/10-15 8192
```

3. Configure the polling and sampling information for Tengigabit Ethernet port 23.

```
console(config)#interface te1/0/23
console(config-if-Te1/0/23)#sflow 1 polling 60
console(config-if-Te1/0/23)#sflow 1 sampling 8192
```

4. Verify the configured information.

```
console#show sflow 1 destination
```

```
Receiver Index.................... 1
Owner String...................... receiver1
Time out.......................... 99994
IP Address:....................... 192.168.30.34
Address Type...................... 1
Port.............................. 6343
Datagram Version.................. 5
Maximum Datagram Size............. 1400
```

console#**show sflow 1 polling**

```
 Poller        Receiver  Poller
 Data Source   Index     Interval
 -----------   -------   -------
Te1/0/10        1         60
Te1/0/11        1         60
Te1/0/12        1         60
Te1/0/13        1         60
Te1/0/14        1         60
Te1/0/15        1         60
Te1/0/23        1         60
```

console#**show sflow 1 sampling**

```
 Sampler      Receiver  Packet          Max Header
 Data Source  Index     Sampling Rate   Size
 -----------  -------   -------------   ----------
Te1/0/10       1         8192           128
Te1/0/11       1         8192           128
Te1/0/12       1         8192           128
Te1/0/13       1         8192           128
Te1/0/14       1         8192           128
Te1/0/15       1         8192           128
Te1/0/23       1         8192           128
```

Configuring RMON

This example generates a trap and creates a log entry when the number of inbound packets are undeliverable due to errors increases by 20 or more.

First, an RMON event is created. Then, the alarm is created. The event (event 1) generates a trap and creates a log entry. The alarm is configured for the MIB object ifInErrors (OID: 1.3.6.1.2.1.2.2.1.14.1). The OID is the variable. The alarm checks the variable every 30 seconds to compare the MIB counter to the configured rising and falling thresholds. If the rise is equal to or greater than 20, event 1 goes into effect.

To configure the switch:

1 Create the event. The trap is sent to the private SNMP community.

```
console#configure
console(config)#rmon event 1 description "emergency event" log
trap private
```

2 Create the alarm.

```
console(config)#rmon alarm 1 1.3.6.1.2.1.2.2.1.14.1 30 delta
rising-threshold 20 1 falling-threshold 1
```

3 Verify the configuration.

```
console#show rmon events

Index  Description   Type     Community  Owner Last time sent
-------------------------------------------------------------
1      emergency    log-trap  private          0 days 0h:0m:0s
       event

console#show rmon alarms

Index    OID                      Owner
----------------------------------------------
1      1.3.6.1.2.1.2.2.1.14.1
```

Configuring Remote Capture

This example configures the switch to mirror packets transmitted and received by the switch CPU to a Wireshark client. This is useful to diagnose switch behavior and to determine if an attached device is sending properly formatted packets with correct information to the switch, or just to monitor traffic sent to the switch CPU. The capture feature can also be configured to capture to a local file in pcap format or to capture to an in-memory buffer (text format).

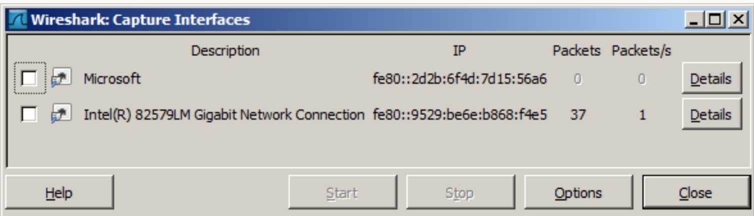
1 Configure capture for Wireshark remote access on port 2002.

```
console(config)#monitor capture mode remote
console(config)#exit
```

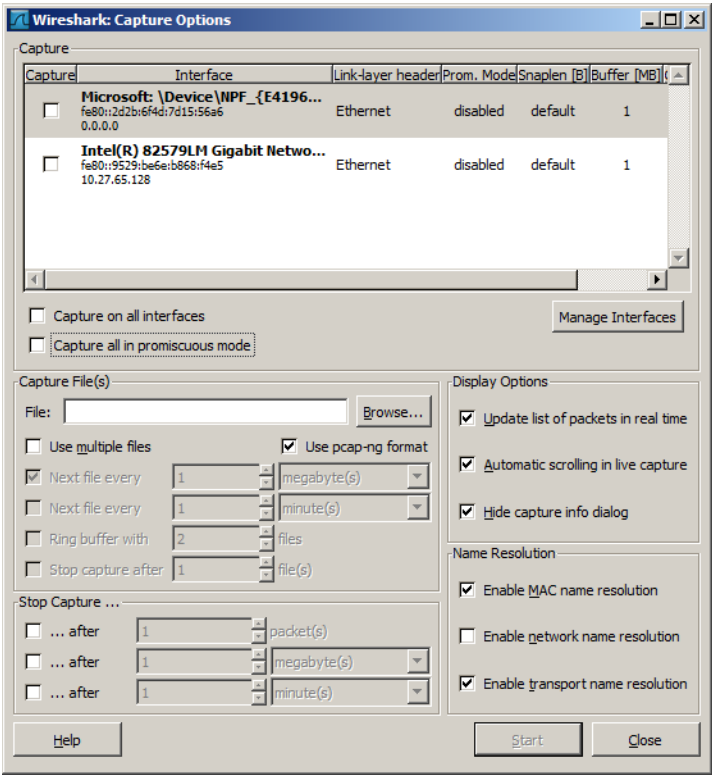
2 Start the capture enabling capture of both transmitted and received packets.

```
console#monitor capture start all
```

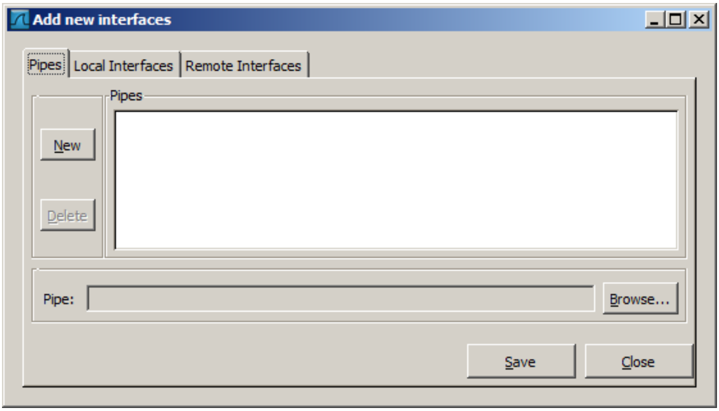
3 Configure Wireshark for remote capture by selecting **Capture > Interfaces** from the top tab. (The screens shown in this example are from Wireshark 1.10.1.)



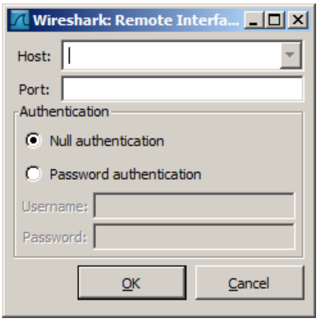
4 On the Capture Interfaces dialog, click **Options**.



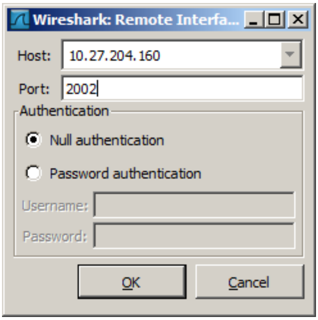
5 On the Capture Options dialog, click **Manage Interfaces**.



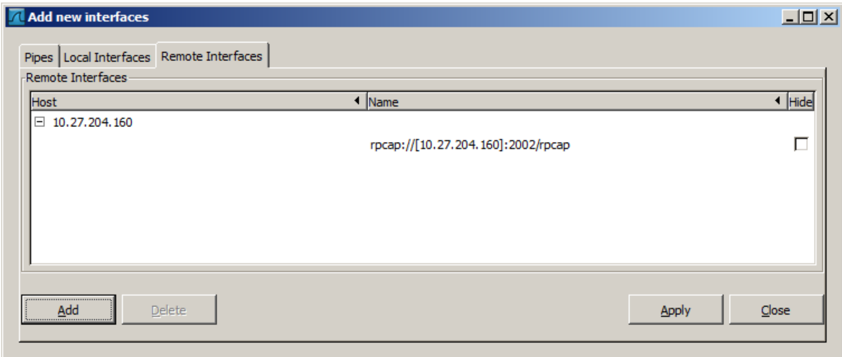
6 Add a new interface by giving the switch IP address and the default remote port (2002). First, select the **Remote Interfaces** tab and click **Add**.



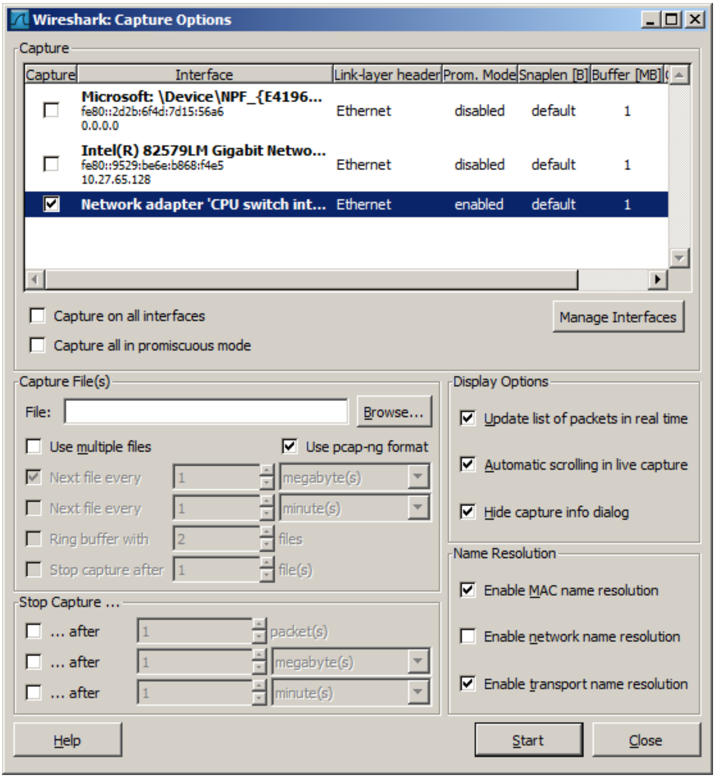
7 Enter the switch IP address and port (2002). Choose **Null authentication** (default).



8 Click **OK** to accept the entry.



9 On the Add new interfaces dialog, click **Apply** and then click **Close**.



10 From the Wireshark:Capture Options dialog, select the remote switch and click **Start**.

Remote Capture Caveats

Remote capture over an in-band port captures the capture packets transmitted to the Wireshark client. Therefore, when using remote capture over an in-band port, it is best to configure remote capture to capture only received packets, to configure remote capture to operate over the out-of-band port, or to configure local capture to capture to the in-memory buffer or a local pcap file.

Configuring RSPAN

RSPAN supports the transport of mirrored packets across the network to a remote switch. Ports may be configured as source ports, intermediate ports, or destination ports.

RSPAN Source Switch

This example mirrors interface gi1/0/3 to VLAN 723. VLAN 723 is the selected transit VLAN. Administrators should reserve a VLAN as the RSPAN VLAN when designing their network. The source switch requires a reflector port to carry packets to the transit switch. The reflector port must be configured as trunk port. Untagged packets on the source port are transmitted on the RSPAN VLAN tagged with the RSPAN VLAN. Tagged packets on the source port are transmitted over the RPSAN VLAN double-tagged with the outer tag containing the RSPAN VLAN.

The last line in this configuration enables the monitor session. It is recommended that configuration proceed with the destination switch first, followed by the intermediate switches, and then by the source switch.

1 Configure RSPAN on VLAN 723:

```
console#configure
console(config)#vlan 723
console(config-vlan723)#remote-span
console(config-vlan723)#exit
```

2 Configure interface te1/0/1 as the reflector port in trunk mode:

```
console(config)#interface te1/0/1
console(config-if-Te1/0/1)#switchport mode trunk
console(config-if-Te1/0/1)#switchport trunk allowed vlan 723
console(config-if-Te1/0/1)#exit
```

3 Configure a mirroring session with a source port gi1/0/3, the destination VLAN 723, and reflector port te1/0/1:

```
console(config)#monitor session 1 source interface gi1/0/3
console(config)#monitor session 1 destination remote vlan 723
reflector-port te1/0/1
```

4 Enable the monitor session:

```
console(config)#monitor session 1 mode
```

RSPAN cannot use the CPU as a mirror source. Instead, configure remote capture to view packets sent to or from the switch CPU.

RSPAN Transit Switch

The following is an example of an RSPAN transit switch configuration. The RSPAN VLAN should be configured as a remote-span in order to disable MAC learning on the VLAN. In this case, the transit switch ports are configured as trunk ports (members of all VLANs) and may be used by other traffic. Packets on the transit switch (in this example) are received and transmitted tagged.

1 Configure remote span on a VLAN:

```
console#configure
console(config)#vlan 723
console(config-vlan723)#remote-span
console(config-vlan723)#exit
```

2 Configure the transit switch ports in trunk mode:

```
console(config)#interface te1/0/1
console(config-if-Te1/0/1)#switchport mode trunk
console(config-if-Te1/0/1)#interface te1/0/2
console(config-if-Te1/0/2)#switchport mode trunk
```

RSPAN Destination Switch

The following example shows the configuration of the RSPAN destination switch. The RSPAN mirrored packets are transmitted over the destination port untagged.

1 Configure remote span on VLAN 723:

```
console#configure
console(config)#vlan 723
console(config-vlan723)#remote-span
console(config-vlan723)#exit
```

2 Configure interface te1/0/1 as the destination port.

```
console(config)#interface te1/0/1
console(config-if-Te1/0/1)#switchport mode trunk
console(config-if-Te1/0/1)#switchport trunk allowed vlan 723
console(config-if-Te1/0/1)#exit
```

3 Configure a mirroring session with the remote VLAN 723 as the source and interface gi1/0/1 as the destination port:

```
console(config)#monitor session 1 source remote vlan 723
console(config)#monitor session 1 destination interface
gi1/0/1
```

4 Enable the mirroring session:

```
console(config)#monitor session 1 mode
```

17

iSCSI Optimization

Dell Networking N2000, N3000, and N4000 Series Switches

NOTE: This feature is only available on the Dell Networking N2000, N3000, and N4000 Series switches.

This chapter describes how to configure Internet Small Computer System Interface (iSCSI) optimization, which enables special quality of service (QoS) treatment for iSCSI traffic.

The topics covered in this chapter include:

- iSCSI Optimization Overview
- Default iSCSI Optimization Values
- Configuring iSCSI Optimization (Web)
- Configuring iSCSI Optimization (CLI)
- iSCSI Optimization Configuration Examples

iSCSI Optimization Overview

iSCSI optimization provides a means of monitoring iSCSI sessions and iSCSI traffic on the switch. This is accomplished by monitoring, or “snooping,” traffic to detect packets used by iSCSI stations to establish iSCSI sessions and connections. Data from these exchanges may optionally be used to create classification rules to assign traffic between the stations to a configured traffic class. The traffic classification affects how the packets in the flow are queued and scheduled for egress on the destination port.

What Does iSCSI Optimization Do?

In networks containing iSCSI initiators and targets, iSCSI Optimization helps to monitor iSCSI sessions or give iSCSI traffic preferential QoS treatment. Dynamically-generated classifier rules generated by snooping iSCSI traffic are used to direct iSCSI data traffic to queues that can be given the desired preference characteristics over other data traveling through the switch. This may help to avoid session interruptions during times of congestion that would otherwise cause iSCSI packets to be dropped. However, in systems where a large proportion of traffic is iSCSI, it may also interfere with other network control-plane traffic, such as ARP or LACP.

The preferential treatment of iSCSI traffic needs to be balanced against the needs of other critical data in the network.

What Occurs When iSCSI Optimization Is Enabled or Disabled?

When iSCSI is enabled on the switch, the following actions occur:

- Flow control is globally enabled, if it is not already enabled.
- iSCSI session snooping is enabled
- iSCSI LLDP monitoring starts to automatically detect Dell EqualLogic arrays.

If the iSCSI feature is disabled on the switch, iSCSI resources are released and the detection of Dell EqualLogic arrays by using LLDP is disabled. Disabling iSCSI does not remove the MTU, flow control, portfast or storm control configuration applied as a result of enabling iSCSI. iSCSI Optimization is enabled by default.

How Does the Switch Detect iSCSI Traffic Flows?

The switch snoops iSCSI session establishment (target login) and termination (target logout) packets by installing classifier rules that trap iSCSI protocol packets to the CPU for examination. Devices that initiate iSCSI sessions generally use well-known TCP ports 3260 or 860 to contact targets. When iSCSI optimization is enabled, by default the switch identifies IP packets to or from these ports as iSCSI session traffic. In addition, the switch separately tracks connections associated with a login session (ISID) (dynamically allocated source/destination TCP port numbers). The switch can be configured to monitor traffic for additional port numbers or port number-target IP address combinations, and the well-known port numbers can be removed from monitoring. A target name can also be associated with a configured target TCP port entry. The maximum number of iSCSI sessions is 1024.

How Is Quality of Service Applied to iSCSI Traffic Flows?

The iSCSI CoS mode is configurable and controls whether CoS queue assignment and/or packet marking is performed on iSCSI traffic. When the iSCSI CoS mode is enabled, the CoS policy is applied to packets in detected iSCSI sessions. When the iSCSI CoS mode is disabled, iSCSI sessions and connections are detected and shown in the status tables, but no CoS policy is applied to packets.

When iSCSI CoS mode is enabled, iSCSI login sessions up to the switch limits are tracked, and data packets for those sessions are given the configured CoS treatment. iSCSI sessions in excess of the switch limits are not given the configured CoS treatment; therefore, it is not advisable to exceed the iSCSI session limit. Multiple connections within a session are counted against the session limit, even though they show in the session table as a single session.

In the switch, iSCSI connections are aged out using the session aging timer. If the connection has no detected data packets during the timeout period, the connection is deleted from the switch internal session table. When all connections associated with a session age out or disconnect, the session is deleted.

Whether the iSCSI optimization feature uses the VLAN priority or IP DSCP mapping to determine the traffic class queue is configurable. By default, iSCSI flows are assigned to the highest VLAN priority tag or DSCP value mapped to the highest queue not used for stack management or voice VLAN. Use the **classofservice dot1p-mapping** command or the **Quality of Service → Class of Service → Mapping Table Configuration** page to configure the relevant Class of Service parameters for the queue in order to complete the setting.

Whether iSCSI frames are remarked to contain the configured VLAN priority tag or the IP DSCP value when forwarded through the switch is configurable.

How Does iSCSI Optimization Use ACLs?

iSCSI Optimization borrows ACL lists from the global system pool. ACL lists allocated by iSCSI Optimization reduce the total number of ACLs available for use by the network operator. Enabling iSCSI Optimization uses one ACL list to monitor for iSCSI sessions. Each monitored iSCSI session utilizes two rules from additional ACL lists up to a maximum of two ACL lists. This means that the maximum number of ACL lists allocated by iSCSI is three.

What Information Does the Switch Track in iSCSI Traffic Flows?

Packets are examined to find the following data, which is used in tracking the session and creating the classifier entries that enable QoS treatment:

- Initiator's IP Address
- Target's IP Address
- ISID (Initiator defined session identifier)
- Initiator's IQN (iSCSI Qualified Name)
- Target's IQN
- Initiator's TCP Port
- Target's TCP Port

If no iSCSI traffic is detected for a session for a configurable aging period, the session data is cleared.

How Does iSCSI Optimization Interact With Dell EqualLogic Arrays?

The iSCSI feature includes auto-provisioning support with the ability to detect directly connected Dell EqualLogic (EQL) SAN storage arrays and automatically reconfigure the switch to enhance storage traffic flows.

The Dell Networking N-Series switches use LLDP, a vendor-neutral protocol, to discover Dell EQL devices on the network. LLDP is enabled by default. For more information about LLDP, see "Discovering Network Devices " on page 825.

When the switch detects a Dell EQL array, the following actions occur:

- An MTU of 9216 is enabled on the system, if it is not already enabled.
- Spanning tree portfast is enabled on the EQL-connected interface identified by LLDP.
- Unicast storm control is disabled on the EQL-connected interface identified by LLDP.

It is advisable to enable spanning tree portfast and disable unicast storm control on ports connected to the initiators as well.

If the iSCSI CoS policy feature is enabled on the switch and an EQL array is detected, the switch applies additional iSCSI CoS policies to the EQL inter-array traffic on TCP ports 9876 and 25555. If the iSCSI CoS policy is disabled and EQL arrays are present, the additional CoS policy is removed globally.

How Does iSCSI Optimization Interact with Dell Compellent Arrays?

Dell Networking N-Series switches support a macro that may be used to configure a port connected to a Dell Compellent storage array. The name of the macro is *profile-compellent-nas*. The macro takes a single argument: the interface identifier to which the Dell Compellent array is connected. The macro disables unicast storm control and sets the spanning tree configuration on the port to portfast. For an example of how to execute the macro, see "Configuring iSCSI Optimization Between Servers and a Disk Array " on page 587.

How Does iSCSI Optimization Interact with DCBx?

NOTE: The DCBx feature is available on the Dell Networking N4000 Series switches only.

The Data Center Bridging Exchange (DCBx) component supports the reception, decoding, and transmission of the Application Priority TLV. In general, if the Application Priority TLV has been received from the configuration source, it will be transmitted to the other auto configuration ports. The DCBx component contains a control to generate the Application Priority TLV for iSCSI if it is not already present in the DCBX information. DCBx generates an Application Priority TLV whenever the following conditions are met:

- An EqualLogic array has been detected on the port
- iSCSI CoS is enabled using a VPT value

The generated Application Priority TLV will contain the following values (in addition to any other information contained in the TLV):

- AE Selector=14
- AE Protocol=3260
- AE Priority=priority configured for iSCSI PFC by the **iscsi cos vpt** command (default priority is 4)

The existing application priority entries being transmitted, if any, will not be disturbed.

If DCBX is enabled (on Dell Networking N4000 Series switches only), the iSCSI application priority TLV is generated by the switch.

On Dell Networking N4000 Series switches, when the iSCSI CoS mode is disabled, the DCBX iSCSI Application Priority TLV is not generated by the switch. In either case, if DCBX is enabled and ports are configured as auto-up or auto-down, the Application Priority TLVs received from the configuration source are proxied to the other ports and, on the Dell Networking N4000 Series switches, the CoS policy for iSCSI received via DCBX is applied to iSCSI packets.

NOTE: If it is desired to utilize DCBX to configure lossless transport of iSCSI using PFC, the operator MUST configure a non-default VLAN end-to-end in order to transport the VLAN priority tag and ensure proper CoS treatment on every network enabled device, including CNAs and the EQL arrays.

iSCSI CoS and Priority Flow Control/Enhanced Transmission Selection Interactions

NOTE: The ETS feature is available on the Dell Networking N4000 Series switches only.

When manually or automatically enabling the classification of iSCSI flows on Dell Networking N4000 Series switches, enabling iSCSI CoS is not recommended unless required as follows.

When using manual configuration of the switch or auto-configuration via DCBX, the iSCSI packets are classified based on the user priority present in the VLAN tag and, in this case, enabling iSCSI CoS classification via the iSCSI command set provides no benefit. The only case for enabling iSCSI CoS prioritization is when using Dell Networking N4000 Series switches to originate iSCSI configuration information via DCBX. In this case, enabling iSCSI CoS classification configures the Dell Networking N4000 Series switch to generate the iSCSI TLV via DCBX in support of configuring directly connected storage and initiator devices. Dell Networking N4000 Series switches support both ETS and DCBx and support DCB configuration in conjunction with EQL devices.

Default iSCSI Optimization Values

Table 17-1 shows the default values for the iSCSI optimization feature.

Table 17-1. iSCSI Optimization Defaults

Parameter	Default Value
iSCSI optimization global status	Enabled
iSCSI CoS mode	Disabled
Jumbo frames	Disabled
Spanning tree portfast	Disabled
Unicast storm control	Disabled
Classification	iSCSI packets are classified by VLAN instead of by DSCP values.
VLAN priority tag	iSCSI flows are assigned by default the highest 802.1p VLAN priority tag mapped to the highest queue not used for stack management or the voice VLAN.
DSCP	When DSCP is selected as the classification, iSCSI flows are assigned by default the highest DSCP tag mapped to the highest queue not used for stack management or the voice VLAN.
Remark	Not enabled
iSCSI session aging time	10 minutes
iSCSI optimization target ports	iSCSI well-known ports 3260 and 860 are configured as default (with no IP address or name) but can be removed as any other configured target.

Configuring iSCSI Optimization (Web)

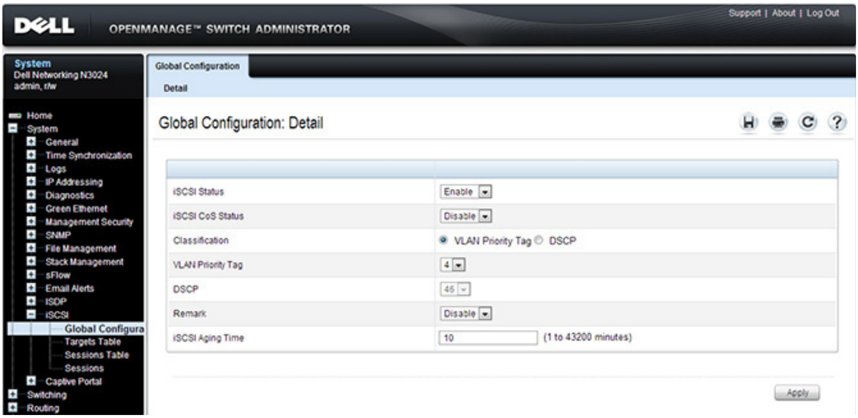
This section provides information about the OpenManage Switch Administrator pages to use to the iSCSI features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click  at the top of the page.

iSCSI Global Configuration

Use the **Global Configuration** page to allow the switch to snoop for iSCSI sessions/connections and to configure QoS treatment for packets where the iSCSI protocol is detected.

To access the iSCSI **Global Configuration** page, click **System** → **iSCSI Global Configuration** in the navigation panel.

Figure 17-1. iSCSI Global Configuration

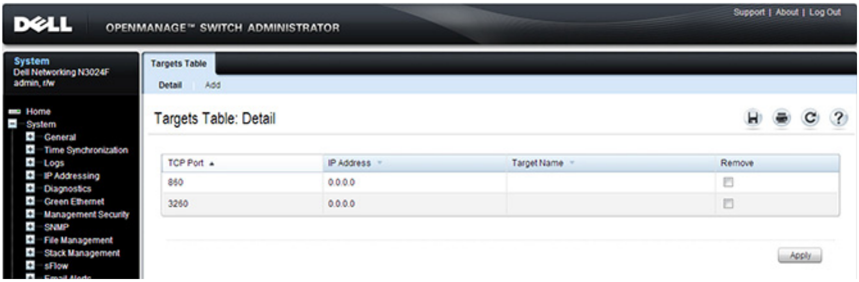


iSCSI Targets Table

Use the **Targets Table** page to view and configure iSCSI targets on the switch.

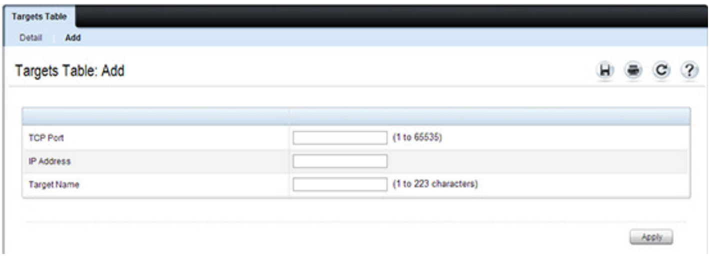
To access the **Targets Table** page, click **System** → **iSCSI** → **Targets** in the navigation panel.

Figure 17-2. iSCSI Targets Table



To add an iSCSI Target, click **Add** at the top of the page and configure the relevant information about the iSCSI target.

Figure 17-3. Add iSCSI Targets

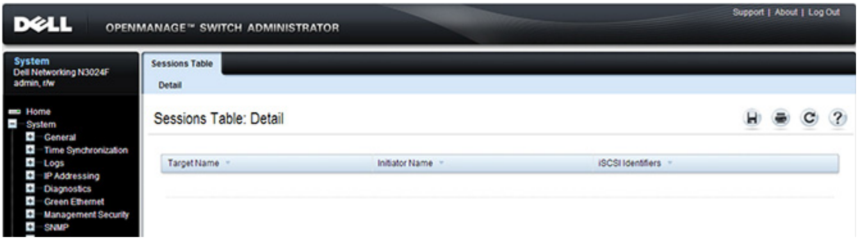


iSCSI Sessions Table

Use the **Sessions Table** page to view summary information about the iSCSI sessions that the switch has discovered. An iSCSI session occurs when an iSCSI initiator and iSCSI target communicate over one or more TCP connections. The maximum number of iSCSI sessions is 192. Redundant (MPIO paths) may not be accounted for in the iSCSI sessions table if a separate iSCSI login is not issued during establishment of the session.

To access the **Sessions Table** page, click **System** → **iSCSI** → **Sessions Table** in the navigation panel.

Figure 17-4. iSCSI Sessions Table

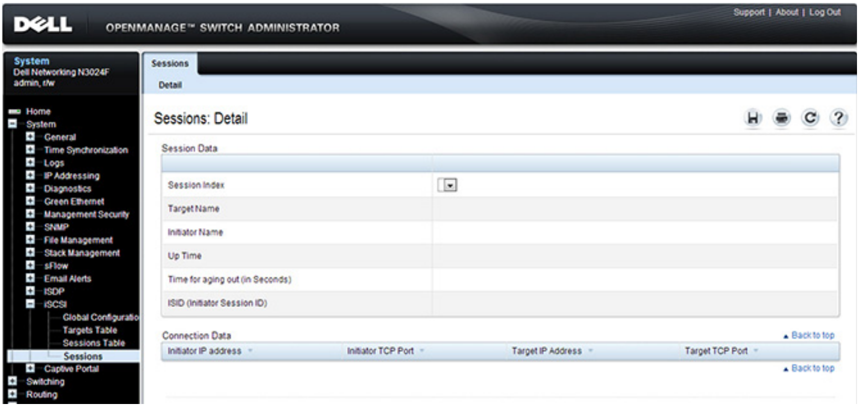


iSCSI Sessions Detailed

Use the **Sessions Detailed** page to view detailed information about an iSCSI sessions that the switch has discovered.

To access the **Sessions Detailed** page, click **System** → **iSCSI** → **Sessions Detailed** in the navigation panel.

Figure 17-5. iSCSI Sessions Detail



Configuring iSCSI Optimization (CLI)

This section provides information about the commands used for configuring iSCSI settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Command	Purpose
configure	Enter Global Configuration mode. iSCSI optimization is enabled by default.
iscsi target port *tcp-port-1* [*tcp-port-2...tcp-port-16*] [**address** *ip-address*] [**name** *targetname*]	Configure an iSCSI target port and, optionally, address and name. • *tcp-port-n*—TCP port number or list of TCP port numbers on which the iSCSI target listens to requests. Up to 16 TCP ports can be defined in the system in one command or by using multiple commands. • *ip-address*—IP address of the iSCSI target. When the no form of this command is used, and the tcp port to be deleted is one bound to a specific IP address, the address field must be present. • *targetname*—iSCSI name of the iSCSI target. The name can be statically configured; however, it can be obtained from iSNS or from sendTargets response. The initiator must present both its iSCSI Initiator Name and the iSCSI Target Name to which it wishes to connect in the first login request of a new session or connection.

Command	Purpose
iscsi cos {**enable** | **disable** | **vtp** *vtp* | **dscp** *dscp* [**remark**]	Optionally set the quality of service profile that will be applied to iSCSI flows. • **enable**—Enables application of preferential QoS treatment to iSCSI frames. On switches that support DCBX, this also enables the generation of the Application Priority TLV for iSCSI. • **disable**—Disables application of preferential QoS treatment to iSCSI frames. • **vpt/dscp**—The VLAN Priority Tag or DSCP value to assign received iSCSI session packets. • **remark**—Mark the iSCSI frames with the configured DSCP value when egressing the switch.
iscsi aging time *time*	Optionally set aging time (range: 1–43,200 seconds) for iSCSI connections. When all connections associated with a session are aged out, the session is deleted.
exit	Exit to Privilege Exec mode.
show iscsi	Display iSCSI settings.
show iscsi sessions	Display iSCSI session information. Redundant (MPIO paths) may not be accounted for in the iSCSI sessions table if a separate iSCSI login is not issued during establishment of the session.

iSCSI Optimization Configuration Examples

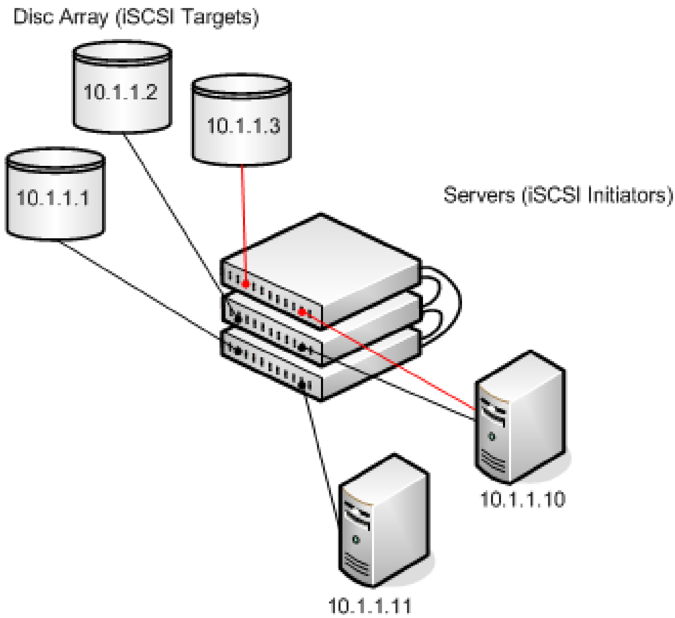
iSCSI optimization is enabled by default. The following procedure illustrates the configuration steps required if configuring iSCSI manually.

Configuring iSCSI Optimization Between Servers and a Disk Array

Figure 17-6 illustrates a stack of three Dell Networking N-Series switches connecting two servers (iSCSI initiators) to a disk array (iSCSI targets).

An iSCSI application running on the management unit (the top unit in the diagram) has installed priority filters to ensure that iSCSI traffic that is part of these two sessions receives priority treatment when forwarded in hardware.

Figure 17-6. iSCSI Optimization



The following commands show how to configure the iSCSI example depicted in Figure 17-6. Remember that iSCSI optimization is enabled by default.

1 Set the system MTU to 9216 to enable the use of jumbo frames.

```
console#config
console(config)#system jumbo mtu 9216
```

2 Optionally configure the switch to associate CoS queue 5 with detected iSCSI session traffic.

```
console(config)#iscsi cos enable
console(config)#exit
```

The default target port and IP address criteria is used to determine which packets are snooped for iSCSI session data (ports 860 and 3260; any IP address).

3 If the array is a Compellent storage array, execute the Compellent macro on the ports attached to the array:

```
console#config
console(config)#macro global apply profile-compellent-nas
$interface_name te1/0/21
console(config)#macro global apply profile-compellent-nas
$interface_name te1/0/22
console(config)#macro global apply profile-compellent-nas
$interface_name te1/0/23
```

To configure a Dell Networking N4000 Series switch in a lossless DCBX environment where another switch connected to storage arrays supplies the DCBX configuration, perform the following steps starting with a default configuration:

1 Enter global configuration mode and configure the system MTU on the switch.

```
console#configure
console(config)#system jumbo mtu 9216
```

2 Create VLAN 100. This command also enters the VLAN configuration mode for VLAN 100.

```
console(config)#vlan 100
console(config-vlan100)#exit
```

3 Enable VLAN tagging to allow the CNA ports 1-4 to carry 802.1p priority values through the network.

```
console(config)#interface range te1/0/1-4
console(config-if)#switchport mode trunk
```

4 Configure the DCBx port role as auto-downstream. This step automatically enables PFC and ETS on the ports using the configuration received from the other switch.

```
console(config-if)#lldp dcbx port-role auto-down
console(config-if)#exit
```

5 Enter interface configuration mode for the switch-facing ports and configure the DCBx port role as auto-up. This step automatically enables PFC and ETS on the ports using the configuration received from the other switch.

```
console(config)#interface range te1/0/16-17
console(config-if)#lldp dcbx port-role auto-up
```

6 Add the ports to port-channel 1:

```
console(config-if)#channel-group 1 mode active
console(config-if)#exit
```

7 Configure the port-channel to be in trunk mode:

```
console(config)#interface po1
console(config-if)#switchport mode trunk
console(config-if)#exit
```

To configure a Dell Networking N4000 Series switch in a lossless DCBX environment where the switch is directly connected to storage arrays and the CNAs (no other switch is present), perform the following steps starting from a default configuration:

1 Enter global configuration mode and configure the system MTU on the switch.

```
console#configure
console(config)#system jumbo mtu 9216
```

2 Create VLAN 100. This command also enters the VLAN configuration mode for VLAN 100.

```
console(config)#vlan 100
console(config-vlan100)#exit
```

3 Enable iSCSI CoS. This enables generation of the iSCSI Application Priority TLV required by the CNAs.

```
console(config)#iscsi cos enable
```

4 Map VLAN priority 4 onto traffic class 4.

```
(config)#classofservice dot1p-mapping 4 4
```

5 Enter Interface Configuration mode for CNA connected ports 1-4 and array connected ports 16-17.

```
console(config)#interface range te1/0/1-4,te1/0/16-17
```

6 Enable VLAN tagging to allow the CNA connected ports to carry 802.1p priority values through the network.

```
console(config-if)#switchport mode trunk
```

7 Enter datacenter bridging mode to enable PFC on the ports.

```
console(config-if)#datacenter-bridging
```

8 Enable PFC and configure traffic marked with 802.1p priority 4 to be paused rather than dropped when congestion occurs.

```
console(config-if-dcb)#priority-flow-control mode on
console(config-if-dcb)#priority-flow-control priority 4
no-drop
console(config-if-dcb)#exit
```

9 Configure ETS by mapping the lossless traffic onto TC 1 and sharing bandwidth equally between the lossless and lossy traffic classes.

```
console(config-if)#classofservice traffic-class-group 0 0
console(config-if)#classofservice traffic-class-group 1 0
console(config-if)#classofservice traffic-class-group 2 0
console(config-if)#classofservice traffic-class-group 3 0
console(config-if)#classofservice traffic-class-group 4 1
console(config-if)#classofservice weight 50 50 0
```

10 Exit interface configuration mode for the range of interfaces.

```
console(config-if)#exit
```

18

Port Characteristics

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure physical switch port characteristics, including settings such as administrative status and maximum frame size. This chapter also describes the link dependency feature.

The topics covered in this chapter include:

- Port Overview
- Default Port Values
- Configuring Port Characteristics (Web)
- Configuring Port Characteristics (CLI)
- Port Configuration Examples

Port Overview

A port is a physical interface. Cables physically connect ports on devices such as PCs or servers to ports on the switch to provide access to the network. The number and type of physical ports available on your Dell Networking N1500, N2000, N3000, and N4000 Series switches depends on the model.

What Physical Port Characteristics Can Be Configured?

Table 18-1 provides a summary of the physical characteristics that can be configured on the switch ports.

Table 18-1. Port Characteristics

Feature	Description
Administrative status	Controls whether the port is administratively enabled or disabled.
Description	Provides a text-based description of the port.
Auto-negotiation	Enables a port to advertise its transmission rate, duplex mode and flow control abilities to its partner.

Table 18-1. Port Characteristics

Feature	Description
Speed	Specifies the transmission rate for frames.
Duplex mode	Specifies whether the interface supports transmission between the switch and the connected client in one direction at a time (half) or both directions simultaneously (both).
Maximum frame size	Indicates the maximum frame size that can be handled by the port.
Green Ethernet features	Green Ethernet features include: • Energy Detect mode • Energy Efficient Ethernet (EEE), which enables the low-power idle mode
Flow control	This is a global setting that affects all ports. For more information about this feature, see "Port-Based Traffic Control " on page 849.
Storm control	For more information about this feature, see "Port-Based Traffic Control " on page 849.
Port security	For more information about this feature, see "Port and System Security " on page 623.
Protected port	For more information about this feature, see "Port-Based Traffic Control " on page 849.

Auto-Negotiation

Dell Networking N-Series switches implement IEEE 802.3 auto-negotiation for 1000BASE-T, 1000BASE-X, and 10GBASE-T based copper interfaces. 1000BASE-X fiber interfaces also implement auto-negotiation. Auto-negotiation is required to be present and enabled for 1000BASE-T and 10GBASE-T copper interfaces in order for a clock master to be selected.

The administrator can configure the advertised capabilities, including the acceptable link speeds, or may disable auto-negotiation altogether. Auto-negotiation must be disabled and full-duplex must be enabled on certain fiber interfaces. Disabling auto-negotiation on copper interfaces is not recommended as it can lead to a duplex mismatch, where one or both interfaces may appear to come up but, in fact, they have not agreed on the speed, duplex, or clock master. This may occur when the devices are connected as follows:

- One end is set manually to half-duplex and the other is manually set to full-duplex
- One end is set to auto-negotiation and the other is manually set to full-duplex
- Both sides are manually set to full-duplex, with one side set to auto-negotiate with the link partner and the other side configured with auto-negotiation disabled.

Maximum Transmission Unit

Dell Networking N-Series switches allow the operator to configure the maximum transmission unit for the switch to a value larger than the IEEE 802.3 standard allows. This jumbo frames technology is employed in certain situations to reduce the task load on a server CPU and to transmit large amounts of data efficiently. The need for jumbo frames predominantly appears where certain applications would benefit from using a larger frame size (for example, Network File System (NFS). The larger frame size eliminates some of the need for fragmentation, leading to greater throughput. The increase in throughput is particularly valuable on data center servers, where the larger frame size increases the efficiency of the system and allows processing of more requests. The Dell Networking jumbo frames feature extends the standard Ethernet MTU (Max Frame Size) from 1518 bytes

(1522 bytes with a VLAN header) to 9216 bytes. Dell Networking N-Series switches assumes that all packets are in Ethernet format. Any device connecting to the same broadcast domain must support the same MTU.

Dell Networking N-Series switches do not fragment L2 or L3 forwarded traffic. Received frames larger than the system MTU are discarded. The switch will not transmit a frame larger than the system MTU.

Packets originated by the switch are fragmented based upon path MTU discovery. IPv4 packets forwarded in software are dropped if they exceed the IPv4 MTU of the outgoing interface, whether or not the Do Not Fragment bit is set in the IP header. An ICMP Fragmentation Needed message is returned to the sender. Dell Networking IPv4 software forwarding does not fragment packets.

An IPv4 packet originated on the switch is fragmented in the IP stack if it is larger than the IPv4 path MTU to the packet's destination. For each IPv4 route in the IP stack's routing table, the default IPv4 path MTU is the IPv4 MTU of the outgoing interface. The IP stack updates the path MTU for each route when it receives Fragmentation Needed ICMP messages.

An IPv6 packet originated on the switch is fragmented in the IP stack if it is larger than the IPv6 path MTU to the packet's destination. IPv6 does not allow forwarded packets to be fragmented. If a forwarded IPv6 packet is larger than the path MTU to its destination, the packet is dropped and an ICMPv6 Packet Too Big message is returned.

What is Link Dependency?

The link dependency feature provides the ability to enable or disable one or more ports based on the link state of one or more different ports. With link dependency enabled on a port, the link state of that port is dependent on the link state of another port. For example, if port A is dependent on port B and the switch detects a link loss on port B, the switch automatically brings down the link on port A. When the link is restored to port B, the switch automatically restores the link to port A.

A maximum of 72 dependency groups can be created. The ports participating in the Link Dependency can be across all the Stack Units (Manager/Member unit).

Link Action

The link action specifies the action that the group members will take when the dependent port is down. The group members can transition to the same state as the dependant port, or they can transition to the opposite state. In other words, if the link action is **down** and the dependent port goes down, the members ports will go down as well. Conversely, when the link action is **up** and the dependant link goes down, the group member ports are enabled (brought up).

Creating a link dependency group with the **up** link action essentially creates a backup link for the dependent link and alleviates the need to implement STP to handle the fail-over.

Link Dependency Scenarios

The Link Dependency feature supports the scenarios in the following list.

NOTE: Whether the member ports or LAGs are brought *up* or *down* depends on the link action.

- Port dependent on port — If a port loses the link, the switch brings up/down the link on another port.
- Port dependent on LAG — If all ports in a channel-group lose the link, the switch brings up/down the link on another port.
- LAG dependent on port — If a port loses the link, the switch brings up/down all links in a channel-group.
- Multiple port command — If a group of ports lose their link, the switch brings up/down the link on another group of ports.
- Overlapping ports — Overlapping ports on different groups will be brought up/down only if both dependent ports lose the link.

What Interface Types are Supported?

The physical ports on the switch include the out-of-band (OOB) interface (Dell Networking N3000 and N4000 Series only) and Ethernet switch ports. The OOB interface supports a limited set of features and is for switch management only. The Ethernet switch ports support many logical features that are often supported by logical interfaces. The switch supports the following types of logical interfaces:

- Port-based VLANs — For more information, see "VLANs " on page 701.
- VLAN routing interfaces — For more information, see "Routing Interfaces " on page 1141.
- Link Aggregation Groups (LAGs), which are also called port-channels) — For more information, see "Link Aggregation " on page 979.
- Tunnels — For more information, see "Routing Interfaces " on page 1141.
- Loopback interfaces — For more information, see "Routing Interfaces " on page 1141.

The Dell Networking N-Series switches includes the following Power over Ethernet (PoE) Plus models: the N1524P, N1548P, N2024P, N2048P, N3024P, N3048P. For information about configuring PoE plus features for the ports, see "Managing General System Settings " on page 389.

Dell Networking N3000 and N4000 Series switches have a single expansion slot and can support the following module types:

- 10GBaseT module
- SFP+ module
- QSFP+ module (Dell Networking N4000 Series only)

What is Interface Configuration Mode?

When you use the CLI to configure physical or logical characteristics for an interface, you must enter Interface Configuration Mode for that interface. To enter the mode, type the keyword **interface** followed by the interface type and additional information to identify the interface, such as the interface number.

To enter Interface Configuration mode for a physical switch port, the following information is required:

- Type — For physical switch ports, the type is Gigabit Ethernet (gigabitethernet or gi) for 10/100/1000 Mbps Ethernet ports or 10-Gibabit Ethernet (tengigabitethernet or te) for 10,000 Mbps Ethernet ports.
- Stack member number— The unit number within the stack. The range is 1–12. The default unit number for a switch that has not been in a stack is 1. To view the member number assigned to each switch in a stack, use the **show switch** command.
- Module (slot) number—For the Dell Networking N4000 Series, the slot number is always 0. The expansion module slot. The number is 1 for a module inserted in the left slot or 2 when it is in the right slot (when viewing the back panel of the switch). For front-panel ports, the slot number is 0.
- Port number—The number assigned to the port. For front-panel ports the port number is written above or below each port. Odd-numbered ports are on the top row, and even-numbered ports are on the bottom row. The port numbers increase from left to right. For ports on the optional modules, the left port is 1, and the right port is 2.

For example, to enter Interface Configuration mode for Gigabit Ethernet port 10 on a switch that is not part of a stack, use the following command:

```
console(config)#interface gigabitEthernet 1/0/10
```

For example, to enter Interface Configuration mode for 10-Gigabit Ethernet port 10, use the following command:

```
console(config)#interface tengigabitEthernet 1/0/10
```

NOTE: When you enter Interface Configuration mode, the command prompt changes and identifies the interface. In the previous example, the command prompt becomes `console(config-if-Te1/0/10)#`.

To enter Interface Configuration mode for Gigabit Ethernet port 6 on stack member 3, use the following command:

```
console(config)#interface gigabitEthernet 3/0/6
```

To enter Interface Configuration mode for port 1 on a 10-Gigabit Ethernet module in the left slot, use the following command:

```
console(config)#interface tengigabitEthernet 1/1/1
```

For many features, a range of interfaces can be specified. When you enter Interface Configuration mode for multiple interfaces, the commands you execute apply to all interfaces specified in the range.

To enter Interface Configuration mode for a range of interfaces, include the keyword **range** and specify the interfaces to configure. For example, to apply the same configuration to ports 1-10 on a standalone switch, use the following command:

```
console(config)#interface range tengigabitEthernet 1/0/1-10
```

To enter Interface Configuration mode for ports 3, 4, 5, 12, and 14 on a standalone switch, use the following command:

```
console(config)#interface range tengigabitEthernet 1/0/3-
5,1/0/12,1/0/14
```

NOTE: To switch to another interface or range of interfaces, enter the interface command while in Interface Configuration mode. It is not necessary to exit Interface Configuration mode to select a different interface.

What Are the Green Ethernet Features?

The Green Ethernet feature supports two per-port power-saving modes:

- Energy-detect Mode
- EEE

All integrated 1G and module-based 10G copper ports on Dell Networking N-Series switches are capable of utilizing the Energy Detect and EEE modes for reduced power consumption.

When the Energy Detect mode is enabled and the port link is down, the PHY automatically goes down for short period of time and then wakes up to check link pulses. This mode reduces power consumption on the port when no link partner is present.

EEE enables ports to enter a low-power mode to reduce power consumption during periods of low link utilization. EEE is defined by IEEE 802.3az. EEE enables both the send and receive sides of the link to disable some functionality for power savings when the link is lightly loaded.

NOTE: Cable diagnostics may give misleading results if green mode is enabled on the port. Disable green mode prior to running any cable diagnostics.

Switchport Modes

Each port on the Dell Networking N1500, N2000, N3000, and N4000 Series switches can be configured to be in one of the following modes:

- Access — Access ports are intended to connect end-stations to the system, especially when the end-stations are incapable of generating VLAN tags. Access ports support a single VLAN (the PVID). Packets received untagged are processed as if they are tagged with the access port PVID. Packets received that are tagged with the PVID are also processed. Packets received that are tagged with a VLAN other than the PVID are dropped. If the VLAN associated with an access port is deleted, the PVID of the access port is set to VLAN 1. VLAN 1 may not be deleted.
- Trunk — Trunk-mode ports are intended for switch-to-switch links. Trunk ports can receive both tagged and untagged packets. Tagged packets received on a trunk port are forwarded on the VLAN contained in the tag if the trunk port is a member of the VLAN. Untagged packets received on a trunk port are forwarded on the native VLAN. Packets received on another interface belonging to the native VLAN are transmitted untagged on a trunk port.
- General — General ports can act like access or trunk ports or a hybrid of both.

VLAN membership rules that apply to a port are based on the switchport mode configured for the port. Table 18-2 shows the behavior of the three switchport modes.

Table 18-2. Switchport Mode Behavior

Mode	VLAN Membership	Frames Accepted	Frames Sent	Ingress Filtering
Access	One VLAN	Untagged/ Tagged	Untagged	Always On
Trunk	All VLANs that exist in the system (default)	Untagged/ Tagged	Tagged and Untagged	Always On
General	As many as desired	Tagged or Untagged	Tagged or Untagged	On or Off

When a port is in General mode, all VLAN features are configurable. When ingress filtering is on, the frame is dropped if the port is not a member of the VLAN identified by the VLAN ID in the tag. If ingress filtering is off, all tagged frames are forwarded. The port decides whether to forward or drop the frame when the port receives the frame.

Default Port Values

Table 18-3 lists the default values for the port characteristics that this chapter describes.

Table 18-3. Default Port Values

Feature	Description
Administrative status	All ports are enabled
Description	None defined
Auto-negotiation	Enabled
Speed	Auto-negotiate
Duplex mode	Auto-negotiate (Full duplex only)
Flow control	Enabled
Maximum frame size	1518
Energy Detect mode	Disabled
EEE mode	Disabled
Link Dependency	None configured

Configuring Port Characteristics (Web)

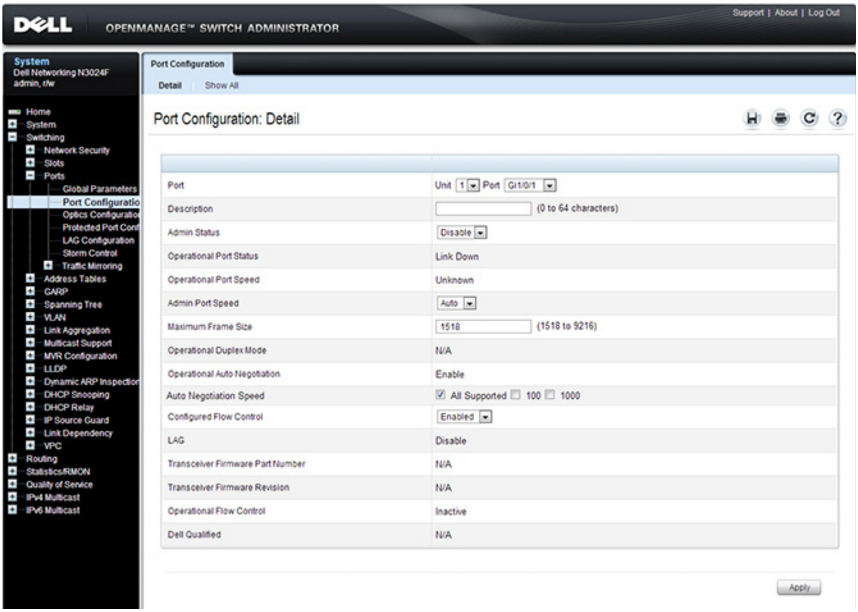
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring port characteristics on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

Port Configuration

Use the **Port Configuration** page to define port parameters.

To display the **Port Configuration** page, click **Switching** → **Ports** → **Port Configuration** in the navigation panel.

Figure 18-1. Port Configuration



Configuring Multiple Ports

To configure port settings on multiple ports:

1 Open the **Port Configuration** page.

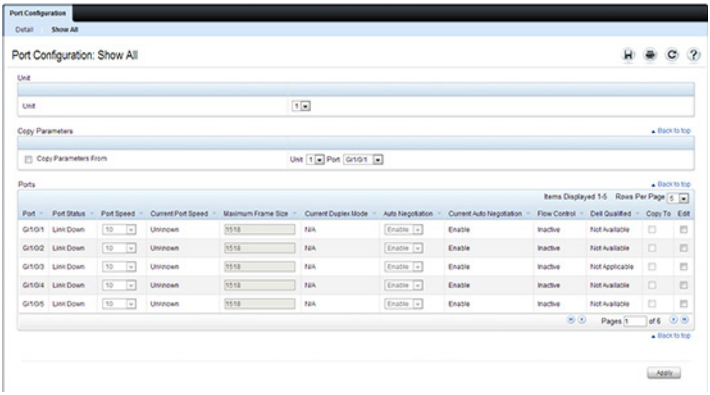
2 Click **Show All** to display the **Port Configuration Table** page.

3 In the **Ports** list, select the check box in the **Edit** column for the port to configure.

4 Select the desired settings.

5 Click **Apply**.

Figure 18-2. Configure Port Settings

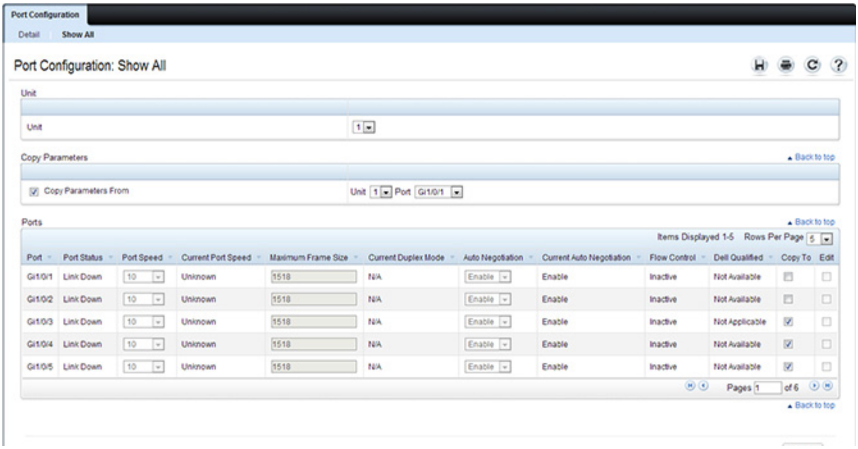


6 Select the **Copy Parameters From** check box, and select the port with the settings to apply to other ports.

7 In the **Ports** list, select the check box(es) in the **Copy To** column that will have the same settings as the port selected in the **Copy Parameters From** field.

In the following example, Ports 3, 4, and 5 will be updated with the settings that are applied to Port 1.

Figure 18-3. Copy Port Settings



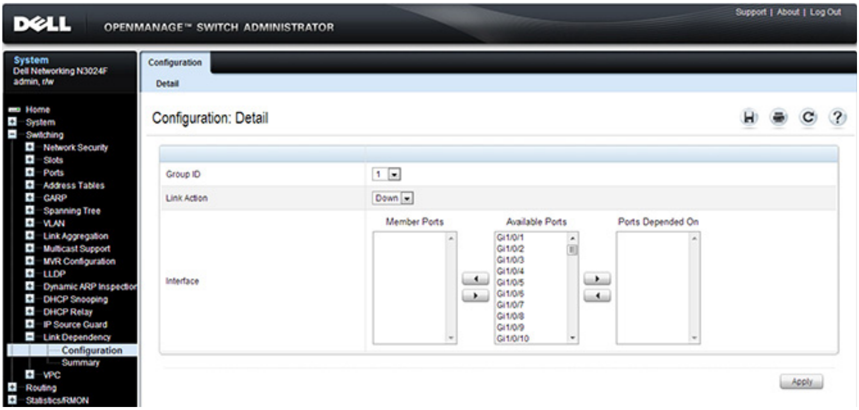
8 Click Apply.

Link Dependency Configuration

Use the **Link Dependency Configuration** page to create link dependency groups. A maximum of 16 dependency groups can be created. The page displays the groups whether they have been configured or not.

To display the **Link Dependency Configuration** page, click **Switching** → **Link Dependency** → **Configuration** in the navigation panel.

Figure 18-4. Link Dependency Configuration



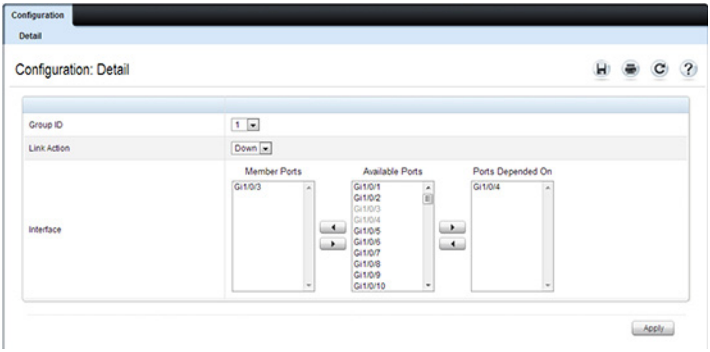
Creating a Link Dependency Group

To create link dependencies:

1. Open the **Link Dependency Configuration** page.
2. In the **Group ID** field, select the ID of the group to configure.
3. Specify the link action.
4. To add a port to the **Member Ports** column, click the port in the **Available Ports** column, and then click the < button to the left of the **Available Ports** column. Ctrl + click to select multiple ports.
5. To add a port to the **Ports Depended On** column, click the port in the **Available Ports** column, and then click the > button to the right of the **Available Ports** column.

In the following example, Group 1 is configured so that Port 3 is dependent on Port 4.

Figure 18-5. Link Dependency Group Configuration



6 Click **Apply**.

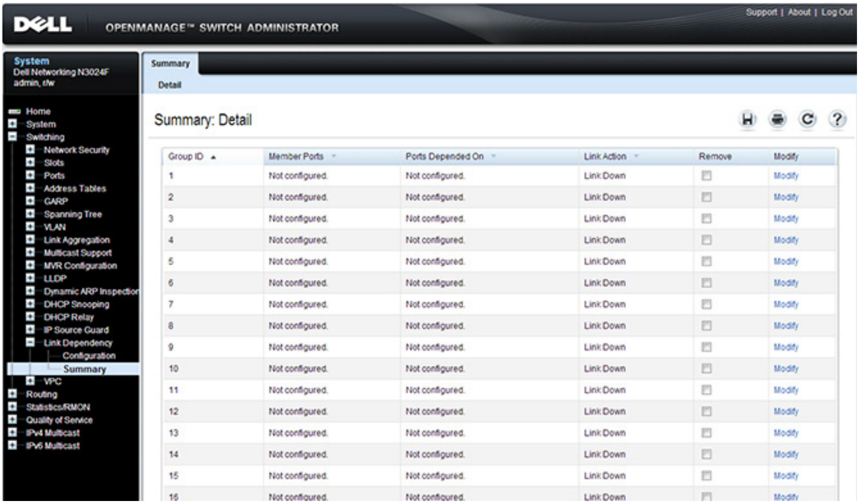
The Link Dependency settings for the group are modified, and the device is updated.

Link Dependency Summary

Use the **Link Dependency Summary** page to view all link dependencies on the system and to access the **Link Dependency Configuration** page. A maximum of 16 dependency groups can be created. The page displays the groups whether they have been configured or not.

To display the **Link Dependency Summary** page, click **Switching → Link Dependency → Link Dependency Summary** in the navigation panel.

Figure 18-6. Link Dependency Summary



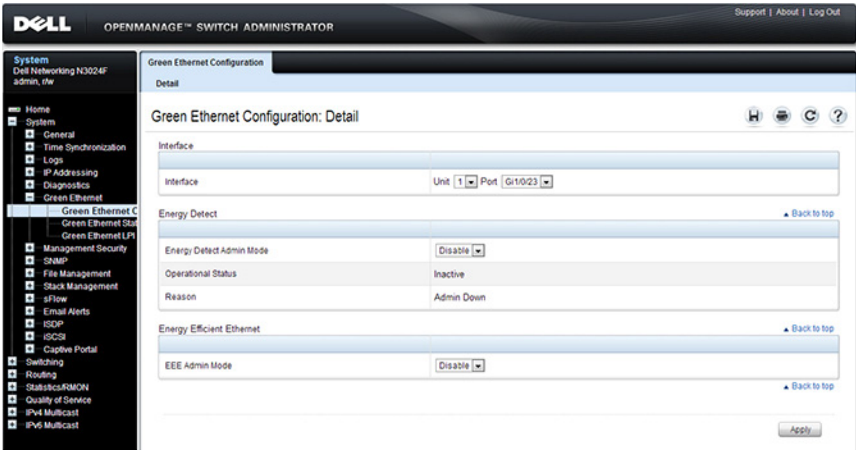
To configure a group, click the **Modify** link associated with the ID of the group to configure. Clicking the **Modify** link takes you to the **Link Dependency Configuration** page. The Group ID is automatically selected based on the link that was clicked.

Port Green Ethernet Configuration

Use the **Green Ethernet Configuration** page to enable or disable energy-saving modes on each port.

To display the **Green Ethernet Configuration** page, click **System** → **Green Ethernet** → **Green Ethernet Configuration** in the navigation panel.

Figure 18-7. Green Ethernet Configuration

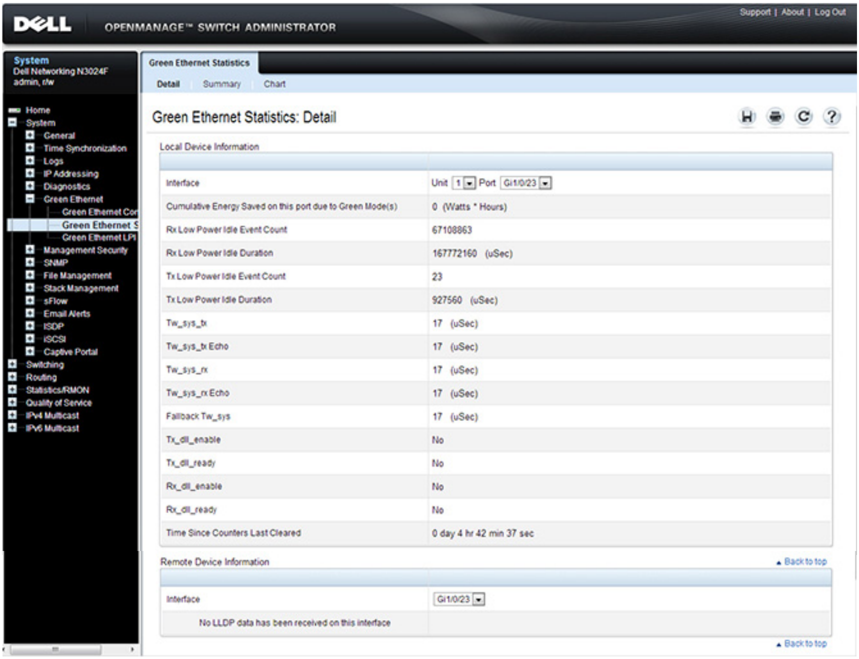


Port Green Ethernet Statistics

Use the **Green Ethernet Statistics** page to view information about per-port energy savings.

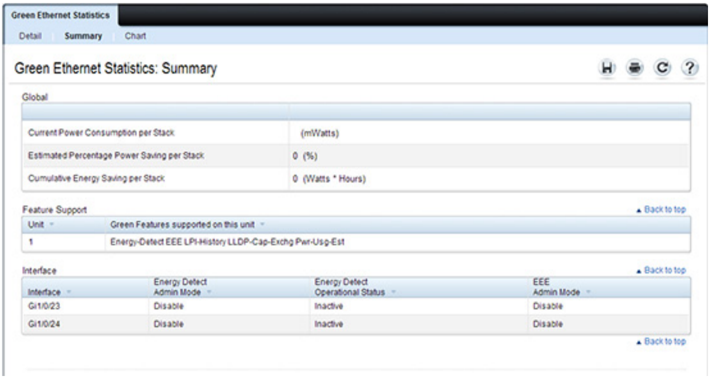
To display the **Green Ethernet Statistics** page, click **System** → **Green Ethernet** → **Green Ethernet Statistics** in the navigation panel.

Figure 18-8. Green Ethernet Statistics



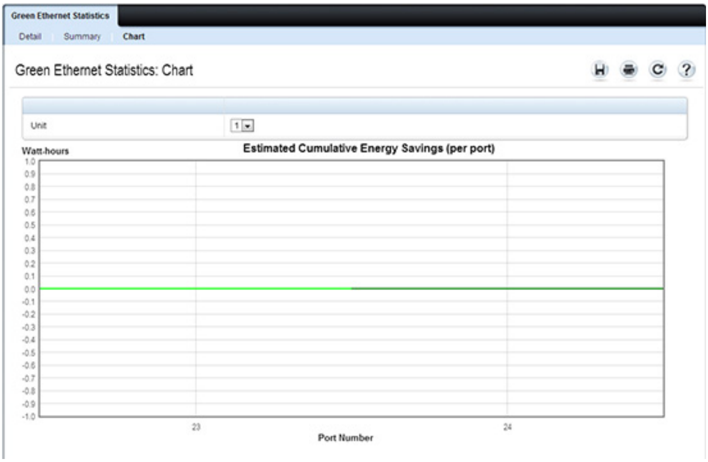
To view a summary of energy savings for the switch and all ports, click **Summary**.

Figure 18-9. Green Ethernet Statistics Summary



To view a chart that shows the estimated per-port energy savings, click **Chart**.

Figure 18-10. Green Ethernet Statistics Chart

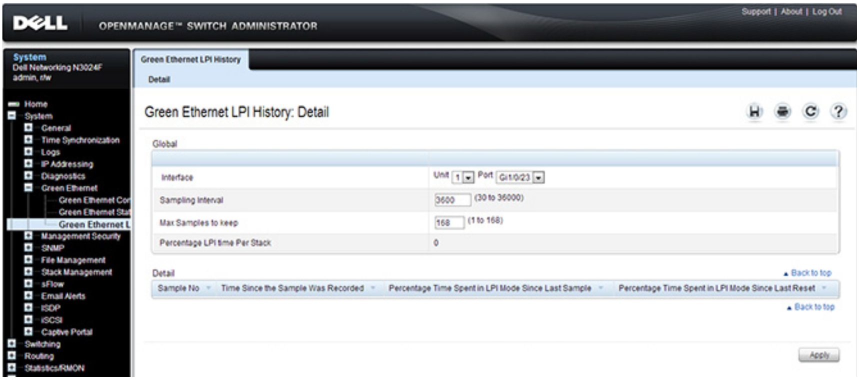


Port Green Ethernet LPI History

Use the **Green Ethernet LPI History** page to view data about the amount of time the switch has spent in low-power idle (LPI) mode.

To display the **Green Ethernet LPI History** page, click **System** → **Green Ethernet** → **Green Ethernet LPI History** in the navigation panel.

Figure 18-11. Green Ethernet LPI History



Configuring Port Characteristics (CLI)

This section provides information about the commands used for configuring port characteristics. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Port Settings

Beginning in Privileged EXEC mode, use the following commands to configure various port settings.

Command	Purpose
configure	Enter Global Configuration mode.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
description *string*	Add a description to the port. The text string can be from1-64 characters.
shutdown	Administratively disable the interface.
speed {10 |100|1000|10000 | **auto** [100|1000|10000]}	Configure the speed of a given Ethernet interface or allow the interface to automatically detect the speed. If you use the **100**, **1000**, or **10000** keywords with the auto keyword, the port auto-negotiates only at the specified speeds. Setting the speed without the auto keyword forces the speed to the selected value and disables auto-negotiation. It is possible to configure a fiber port for a speed not supported by the transceiver. In this case, the port will not link up. On combo ports, it is possible to configure auto-negotiation even if only the fiber interface is active. The auto-negotiation settings will be utilized when the copper port is active or a 1000BASE-X transceiver is inserted. Fiber ports always operate in full-duplex mode.

Command	Purpose
system jumbo mtu *size*	Enable jumbo frames on the switch by adjusting the maximum size of a packet.
CTRL + Z	Exit to Privileged EXEC mode.
show interfaces status	Show summary information about all interfaces.
show interfaces configuration	View a summary of the configuration for all ports.
show interfaces advertise	View a summary of the speeds that are advertised on each port.
show interfaces description	View configured descriptions for all ports.
show interfaces detail *interface*	View detailed information about the specified port.

Configuring Link Dependencies

Beginning in Privileged EXEC mode, use the following commands to configure ports that are dependent on the state of other ports.

Command	Purpose
configure	Enter Global Configuration mode.
link-dependency group *group_id*	Enter the link-dependency mode to configure a link-dependency group.
add *interface*	Add member ports to the group. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. Port-channels (LAGs) can also be added as members by using the keyword **port-channel** followed by an ID. A range of interfaces can also be specified. For example, **interface tengigabitethernet 1/0/8-10,1/0/20** configures interfaces 8, 9, 10 and 20.
depends-on *interface*	Specify the port(s) upon which the member ports are dependent. For information about the *interface* variable, see the previous command description.
action {**down**|**up**}	Specifies the action the member ports take when the dependent link goes down. • **down**—When the dependent link is down, the group members are down (the members are up otherwise). • **up**—When the dependent link goes down, the group members are brought up (the members are down otherwise)
CTRL + Z	Exit to Privileged EXEC mode.
show link-dependency [**group** *group_id*]	View link dependency settings for all groups or for the specified group, along with the group state.

Configuring Green Features

Beginning in Privileged EXEC mode, use the following commands to configure and monitor energy-saving features for the ports and the switch.

Command	Purpose
configure	Enter Global Configuration mode.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **gigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range gigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
green-mode energy-detect	Enable energy-detect mode on the interface.
green-mode eee	Enable EEE low power idle mode on the interface.
exit	Exit to global configuration mode.
green-mode eee-lpi-history {**sampling-interval** *seconds* | **max-samples** *max*}	Configure the global EEE LPI history collection interval and buffer size.
exit	Exit to Privileged EXEC mode.
show green-mode *interface*	View green mode settings for the specified port.
show green-mode eee-lpi-history interface *interface*	View the EEE LPI history statistics for the specified port.

Port Configuration Examples

This section contains the following examples:

- Configuring Port Settings
- Configuring a Link Dependency Groups

Configuring Port Settings

The commands in this example specify the speed for port 1 (gigabitEthernet 1/0/1) and change the system MTU size.

To configure the switch:

1 Enter Interface Configuration mode for port 1.

```
console#configure
console(config)#interface gigabitEthernet 1/0/1
```

2 Change the speed settings for the port.

```
console(config-if-Gi1/0/1)#speed 100
console(config-if-Gi1/0/1)#exit
```

3 Enable jumbo frame support on the interfaces.

```
console(config)#system jumbo mtu 9216
console(config)#CTRL + Z
```

4 View summary information about the ports

```
console#show interfaces configuration

Port       Type                Duplex  Speed    Neg   Admin St.
---------  -------------      ------  -------  ----  -----
Gi1/0/1    Gigabit - Level     Full    100      Off   Up
Gi1/0/2    Gigabit - Level     N/A     Unknown  Auto  Up
Gi1/0/3    Gigabit - Level     N/A     Unknown  Auto  Up
Gi1/0/4    Gigabit - Level     N/A     Unknown  Auto  Up
Gi1/0/5    Gigabit - Level     N/A     Unknown  Auto  Up

--More-- or (q)uit
```

Configuring a Link Dependency Groups

The commands in this example create two link dependency groups. Group 1 has port 3 as a member port that is dependent on port 4. The group uses the default link action, which is down. This means that if port 4 goes down, port 3 goes down. When port 4 returns to the up state, port 3 is brought back up. In Group 2, port 6 dependent on port-channel (LAG) 1, and the link action is up. If port-channel 1 goes down, port 6 is brought up. This also means that when port-channel 1 is up, port 6 is down.

To configure the switch:

1 Enter the configuration mode for Group 1.

```
console#configure
console(config)#link-dependency group 1
```

2 Configure the member and dependency information for the group.

```
console(config-linkDep-group-1)#add tengigabitethernet 1/0/3
console(config-linkDep-group-1)#depends-on tengigabitethernet
1/0/4
console(config-linkDep-group-1)#exit
```

3 Enter the configuration mode for Group 2

```
console(config)#link-dependency group 2
console(config-linkDep-group-2)#add tengigabitethernet 1/0/6
console(config-linkDep-group-2)#depends-on port-channel 1
console(config-linkDep-group-2)#action up
console(config-linkDep-group-2)#CTRL + Z
```

4 View the configured link dependency groups.

```
console#show link-dependency

GroupId  Member Ports     Ports Depended On    Link Action
-------  -------------   -----------------     ----------
1        Te1/0/3         te/0/4                Link Down
2        te/0/6          ch1                   Link Up
```

Configuring a Port in Access Mode

Beginning in Privileged EXEC mode, use the following commands to configure an access mode VLAN interface and, optionally, assign the interface to a VLAN. When a port is in access mode, it can only be a member of one VLAN and will accept tagged packets with the access VLAN ID or untagged packets. Untagged packets are treated as belonging to the access VLAN. Packets received with a VLAN ID other than the access VLAN ID are discarded. When you configure an interface as an access mode port, the interface is automatically made a member of VLAN 1 and removed from all other VLAN memberships. Each interface can be configured separately, or a range of interfaces can be configured with the same settings.

Command	Purpose
configure	Enter global configuration mode.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
switchport mode access	Configure the interface as an access mode VLAN interface. Access mode VLANs accept tagged or untagged packets for the access VLAN only.
switchport access vlan *vlan-id*	Configure the interface as a member of the specified VLAN. By default, access mode ports are members of VLAN 1. *vlan-id* — A valid VLAN ID of the VLAN to which the port is configured. (Range: 1–4093)
CTRL + Z	Exit to Privileged EXEC mode.
show interfaces switchport *interface*	Display information about the VLAN settings configured for the specified interface.

Configuring a Port in Trunk Mode

Beginning in Privileged EXEC mode, use the following commands to configure an interface as a layer-2 trunking interface, which connects two switches. Trunk mode ports support traffic tagged with different VLAN IDs. Untagged received traffic is switched in the native VLAN. A trunk port is automatically configured as a member of all VLANs. Trunk ports can be removed from membership in specific VLANs. By default, the native VLAN for a trunk port is VLAN 1.

Command	Purpose
configure	Enter global configuration mode.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command; For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
switchport mode trunk	Configure the interface as a tagged layer-2 VLAN interface.

<table>
<tr><th>Command</th><th>Purpose</th></tr>
<tr><td>switchport trunk {allowed vlan vlan-list|native vlan vlan-id}</td><td>Set the list of allowed VLANs that can receive and send traffic on this interface in tagged format when in trunking mode.

• allowed vlan-list — Set the list of allowed VLANs that can receive and send traffic on this interface in tagged format when in trunking mode. Separate non-consecutive VLAN IDs with a comma and no spaces. Use a hyphen to designate a range of IDs.

The vlan-list format is all | [add | remove | except] vlan-atom [vlan-atom...] where:

• all—Specifies all VLANs from 1 to 4093. This keyword is not allowed on commands that do not permit all VLANs in the list to be set at the same time.

• add—Adds the list of VLANs to the allowed set.

• remove—Removes the list of VLANs from the allowed set. Removing the native VLAN from a trunk port forces the port to allow tagged packets only.

• except—Allows all VLANs other than those in the list.

• vlan-atom —Either a single VLAN number from 1 to 4093 or a continuous range of VLANs described by two VLAN numbers, the lesser one first, separated by a hyphen.

• native vlan-id— The untagged VLAN. Untagged packets received on this interface are switched in the native VLAN. Transmitted packets in this VLAN are sent untagged.</td></tr>
<tr><td>CTRL + Z</td><td>Exit to Privileged EXEC mode.</td></tr>
<tr><td>show interfaces switchport interface</td><td>Display information about the VLAN settings configured for the specified interface. The interface variable includes the interface type and number.</td></tr>
</table>

Configuring a Port in General Mode

Beginning in Privileged EXEC mode, use the following commands to configure an interface with full 802.1q support and configure the VLAN membership information for the interface. Except when noted as required (for example, when configuring MAB, Voice VLAN, or 802.1x), it is recommended that operators use either trunk or access mode.

Command	Purpose
configure	Enter global configuration mode.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
switchport mode general	Configure the interface as a tagged and untagged layer-2 VLAN interface.
switchport general allowed vlan [**add**|**remove**] *vlan-list* {**tagged**|**untagged**}	Configure the VLAN membership for the port. This command can also be used to change the egress tagging for packets without changing the VLAN assignment. • **add** *vlan-list* — List of VLAN IDs to add. Separate nonconsecutive VLAN IDs with a comma and no spaces. Use a hyphen to designate a range of IDs. (Range: 1–4093) • **remove** *vlan-list* — List of VLAN IDs to remove. Separate nonconsecutive VLAN IDs with a comma and no spaces. Use a hyphen to designate a range of IDs. • **tagged** — Sets the port to transmit tagged packets for the VLANs. If the port is added to a VLAN without specifying tagged or untagged, the default is untagged. • **untagged** — Sets the port to transmit untagged packets for the VLANs.

Command	Purpose
switchport general pvid *vlan-id*	(Optional) Set the port VLAN ID. Untagged traffic that enters the switch through this port is tagged with the PVID. *vlan-id* — PVID. The selected PVID assignment must be to an existing VLAN. (Range: 1–4093). Entering a PVID value does not remove the previous PVID value from the list of allowed VLANs.
switchport general acceptable-frame-type tagged-only	(Optional) Specifies that the port will only accept tagged frames. Untagged frames are dropped at ingress.
switchport general ingress-filtering disable	(Optional) Turn off ingress filtering so that all received tagged frames are forwarded whether or not the port is a member of the VLAN in the tag.
CTRL + Z	Exit to Privileged EXEC mode.
show interfaces switchport *interface*	Display information about the VLAN settings configured for the specified interface. The *interface* variable includes the interface type and number.

19

Port and System Security

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure port-based and system security features, which control access to the network through the switch ports, and the denial of service (DoS) feature.

The topics covered in this chapter include:

- Port-based Security—Port MAC Locking
- Denial of Service

Port-based Security—Port MAC Locking

Port MAC locking is used to enable security on a per-port basis. When a port is locked, only packets with allowable source MAC addresses can be forwarded. All other packets are discarded. Port-MAC locking allows a configurable limit to the number of source MAC addresses that can be learned on a port.

NOTE: Port-based security can also be accomplished by using Access Control Lists (ACLs). For information about configuring ACLs, see "Access Control Lists " on page 629.

The Port Security feature allows you to limit the number of source MAC addresses that can be learned on a port. If a port reaches the configured limit, any other addresses beyond that limit are not learned and the frames are discarded. Frames with a source MAC address that has already been learned will be forwarded.

The purpose of this feature, which is also known as port-MAC locking, is to help secure the network by preventing unknown devices from forwarding packets into the network. For example, to ensure that only a single device can be active on a port, set the number of allowable dynamic addresses to one. After the MAC address of the first device is learned, no other devices will be allowed to forward frames into the network.

When link goes down on a port, all of the dynamically locked addresses are cleared from the source MAC address table the feature maintains. When the link is restored, that port can once again learn addresses up to the specified limit.

The port can learn MAC addresses dynamically, and a list of static MAC addresses can be specified for a port. The number of static addresses that may be configured is limited by the port security limit, regardless of whether port security is enabled on the interface or not.

Default 802.1X Values

Table 19-1 lists the default values for the Port Security feature.

Table 19-1. Default Port Security Values

Feature	Description
Port security	Unlocked
Port security traps	Disabled
Maximum learned MAC addresses	100 (when locked)
Maximum static MAC addresses	100
Monitor mode	Disabled

Configuring Port Security Configuration (Web)

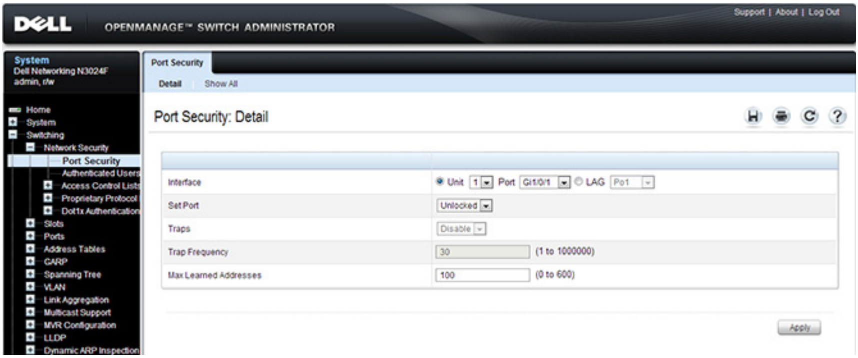
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring the IEEE 802.1X features and Port Security on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

Port Security

Use the **Port Security** page to enable MAC locking on a per-port basis. When a port is locked, a limit can be specified for the number of source MAC addresses that are allowed to transmit traffic on the port.

To display the **Port Security** page, click **Switching → Network Security → Port Security** in the navigation panel.

Figure 19-1. Network Security Port Security

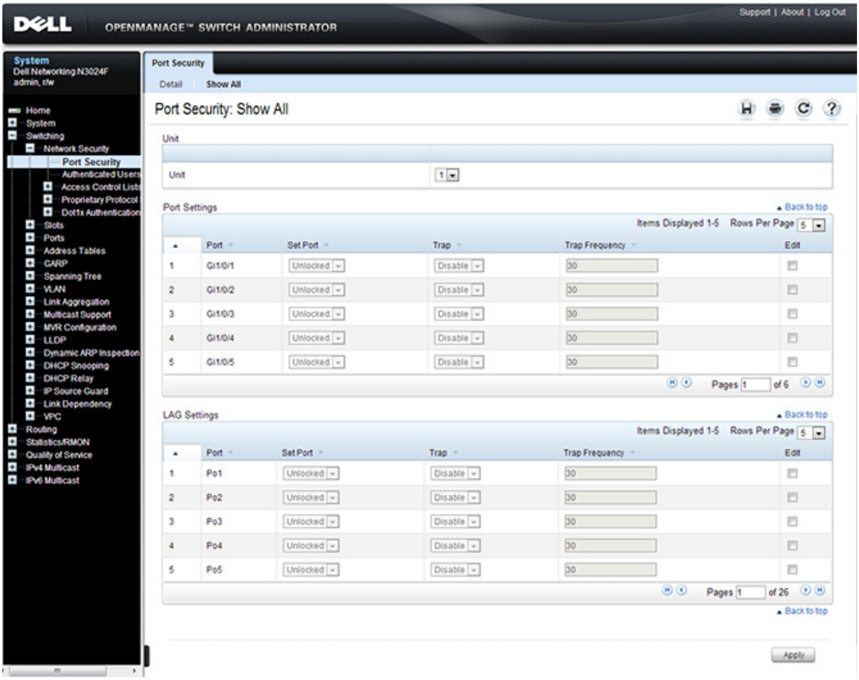


Configuring Port Security Settings on Multiple Ports

To configure port security on multiple ports:

1. Open the **Port Security** page.
2. Click **Show All** to display the **Port Security Table** page.
3. In the Ports list, select the check box in the **Edit** column for the port to configure.
4. Select the desired settings for all ports that are selected for editing.

Figure 19-2. Configure Port Security Settings



5 Click **Apply**.

Configuring Port Security (CLI)

Beginning in Privileged EXEC mode, use the following commands to enable port security on an interface to limit the number of source MAC addresses that can be learned.

Command	Purpose
configure	Enter Global Configuration mode.
switchport port-security	Enable port-security administrative mode. Port security must be enabled globally in order to operate on any interfaces.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
switchport port-security [**mac-address** {*mac-address* **vlan** {*vlan-id*}}] | **dynamic** *value* | **maximum** *value*]	Enable port security on the port. This prevents the switch from learning new addresses on this port after the maximum number of addresses has been learned. • **mac-address** — configure a static MAC address on the interface and VLAN. This command performs the same function as the **mac address-table static** command. • **dynamic** — set the maximum number of dynamic MAC addresses that may be learned on the interface. • **maximum** — set the maximum number of static MAC addresses that may be configured on the interface. This limit applies regardless of the port security administrative setting.
CTRL + Z	Exit to Privileged EXEC mode.
show port-security [*interface-id* | **all** | **dynamic** *interface-id* | **static** *interface-id* | **violation** *interface-id*]	View port security settings on all interfaces or the specified interface. Use the **dynamic** keyword to display learned MAC addresses and the **static** keyword to display configured MAC addresses.

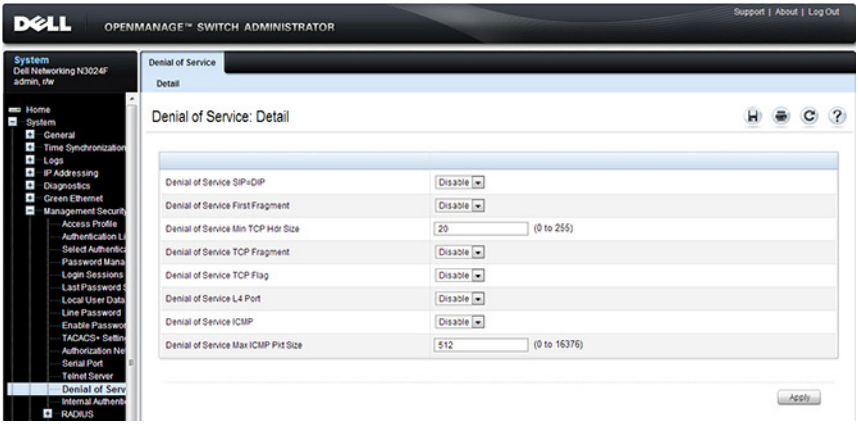
Denial of Service

Denial of Service (DoS) refers to the exploitation of a variety of vulnerabilities which would interrupt the service of a host or make a network unstable. Use the **Denial of Service** page to configure settings to help prevent DoS attacks.

DoS protection is disabled by default.

To display the **Denial of Service** page, click **System** → **Management Security** → **Denial of Service** in the navigation panel.

Figure 19-3. Denial of Service



20

Access Control Lists

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure Access Control Lists (ACLs), including IPv4, IPv6, and MAC ACLs. This chapter also describes how to configure time ranges that can be applied to any of the ACL types.

The topics covered in this chapter include:

ACL Overview

Access Control Lists (ACLs) are a collection of rules that provide security by blocking selected packets from ingressing the switch. ACLs are implemented in hardware and processed at line rate for the front-panel ports. A reduced functionality set of ACLs is implemented in firmware for the OOB port.

ACLs can also provide ingress traffic rate limiting and decide which types of traffic are forwarded or blocked. Egress ACLs support traffic shaping. ACLs support deployment as a firewall router, a router connecting two internal networks, or a layer-3 router implementing routing policies.

To harden the switch against external threats, it is possible to create an ACL that limits access to the management interfaces based on the connection method (for example, Telnet or HTTP) and/or the source IP address.

The Dell Networking N-Series switches support ACL configuration in both the ingress and egress direction. Egress ACLs provide the capability to implement security rules on the egress flows (traffic leaving a port) rather than the ingress flows (traffic entering a port). Ingress and egress ACLs can be applied to any physical port, port-channel (LAG), or VLAN routing port.

Depending on whether an ingress or egress ACL is applied to a port, when the traffic enters (ingress) or leaves (egress) a port, the ACL compares the criteria configured in its rules, in list order, to the fields in a packet or frame to check for matching conditions. The ACL processes the traffic based on the actions contained in the rules.

ACLs are organized into access groups. Access groups are numbered in priority (lowest number has highest priority). Multiple access groups can be configured on an interface, in which the lowest numbered access group is processed first, followed by the next lowest numbered access group, etc.

NOTE: Conceptually, ACL processing proceeds by attempting to match each of the ACLs listed in the first match term or clause in the first access group in order. If an ACL does not match, processing moves to the next ACL in order until an ACL matches or the ACL group is exhausted. If there are more access groups configured, processing proceeds with the next access group.

In reality, all interface ACL matches are attempted in parallel at once, and the priority of the ACL is used to determine the action. Then, all VLAN ACL matches are attempted in parallel at once, and the priority of the ACL is used to determine the action. This implies that a packet that matches both a physical interface ACL and a VLAN ACL will always take the physical interface action.

Within an access group, ACL rules are processed in sequence, from the first (lowest numbered) rule to the last (highest numbered) rule in the access group. If a matching rule is found, the rule action is taken and no subsequent rules are processed for that packet. Frequently matched rules should be placed near or at the front of the list. At least one access list within the access groups configured on an interface must contain at least one permit rule or all traffic is denied (dropped). ACL entries may be numbered by the administrator when configured or automatically numbered by the system. Additionally, remarks may be entered for an ACL entry.

Packets generated by the switch are sent regardless of any egress ACL deny rules.

NOTE: The last access group configured on an interface is terminated by an implicit **deny all** rule, which drops any packet not matching a preceding **permit** rule. The implicit **deny all** rule is not configured if Policy-Based Routing is configured on the interface.

ACLs may be used to control traffic at layer 2, layer 3, or layer 4. MAC ACLs contain packet match criteria based on layer-2 fields in Ethernet frames. IP ACLs contain packet match criteria based on layer-3 and layer-4 fields in the packet. Dell Networking N-Series switches support both IPv4 and IPv6 ACLs and supports ACLs applied to up to 24 VLAN interfaces.

ACL Counters

Matching rules in an ACL are counted. The counts may be displayed using the **show ip access-list** or **show mac access-list** commands. For ACL counters, if an ACL rule is configured without a rate-limit, the counter value is the count of the permitted or denied packets. (Example: If a burst of 100 matching packets is received, the counter value is 100.)

If an ACL rule is configured with a rate limit, the counter value will be the matched packet count. If the received traffic rate exceeds the configured limit, the counters still display matched packet count despite the packets which exceed the configured limit since match criteria is met. For example, if the rate limit is set to 10 Kbps and 'matching' traffic is received at 100 Kbps, the counters reflect the 100 Kbps value. If the received traffic rate is less than the configured limit, the counters display only the matched packet count. ACL counters do not interact with Diffserv policies.

What Are MAC ACLs?

MAC ACLs are layer-2 ACLs. MAC ACLs can filter on the following fields of a packet:

- Source MAC address
- Source MAC mask
- Destination MAC address
- Destination MAC mask
- VLAN ID
- Class of Service (CoS) (802.1p)
- EtherType

MAC access list actions include CoS queue assignment, logging, mirroring, redirection to another port, and logging, as well as the usual permit and deny actions. It is possible to configure MAC access groups in conjunction with IP access groups on the same interface. MAC ACLs can be configured on a VLAN interface as well as a physical interface or port channel.

What Are IP ACLs?

IP ACLs contain filters for layers 3 and 4 on IPv4 or IPv6 traffic.

Each IP ACL is a set of up to 100 rules applied to inbound or outbound traffic. IP ACLs support logging, redirect, mirroring, and drop. The following fields may be specified in the permit or deny rules.

- Destination IP with wildcard mask
- Destination layer-4 port
- Every protocol or a specific protocol
- IP DSCP
- IP precedence
- IP TOS
- TCP flags
- Source IP with wildcard mask
- Source layer-4 port, with eq, ne, gt, and lt operators and ranges (IP/TCP/UDP packets only)
- Destination layer-4 port, with eq, ne, gt, and lt operators and ranges (TCP/UDP packets only)

IP access lists may be configured on physical interfaces and port channels as well as VLANs.

ACL Actions

The following actions are available for ingress ACLs. Not all actions are available for all types of ACLs. Refer to "ACL Limitations " on page 635 for more details.

- CoS queue assignment—assign the matching packet to the specific CoS queue. This action does not rewrite any fields in the packet.

- Log—perform the logging action on the matching packet as described below.
- Mirror—forward a copy of the matching packet to the designated interface. The original packet continues to be forwarded to its original destination.
- Redirect—forward the matching packet to the designated interface. The original destination of the packet is ignored.
- Rate limit—forward matching packets that do not exceed the rate limit. Drop packets exceeding the rate limit. Refer to the Diffserv section for more sophisticated ingress rate limiting.

The following actions are available for egress ACLs. Not all actions are available for all types of ACLs. Refer to "ACL Limitations " on page 635 for more details.

- CoS queue assignment—rewrite the matching packet CoS value. This action does not affect processing of the packet within the switch.
- Log—perform the logging action on the matching packet as described below.
- Mirror—forward a copy of the matching packet to the designated interface. The original packet continues to be forwarded to its original destination.
- Redirect—forward the matching packet to the designated interface. The original destination of the packet is ignored.
- Rate limit—forward matching packets that do not exceed the rate limit. Drop packets exceeding the rate limit. Refer to "Differentiated Services " on page 1445 for more sophisticated ingress rate limiting.

What Is the ACL Redirect Function?

The redirect function allows traffic that matches a permit rule to be redirected to a specific physical port or LAG instead of processed on the original port. A packet that is redirected does not go through the normal forwarding process. It is sent to the redirect target port. The redirect function and mirror function are mutually exclusive. In other words, a given ACL rule cannot be configured with both mirror and redirect attributes.

What Is the ACL Mirror Function?

ACL mirroring provides the ability to send a copy of traffic that matches a permit rule to a specific physical port or LAG. Using ACLs to mirror traffic is called flow-based mirroring, since the traffic flow is defined by the ACL classification rules. This is in contrast to port mirroring, where all traffic encountered on a specific interface is replicated out of another interface.

Mirroring is similar to the redirect function, except that in flow-based mirroring a copy of the permitted traffic is delivered to the mirror interface while the packet itself is forwarded normally through the device. A given ACL rule cannot be configured with both mirror and redirect attributes.

What Is ACL Logging

ACL Logging provides a means for counting the number of “hits” against an ACL rule. To configure ACL Logging, augment the ACL deny rule specification with a “log” parameter that enables hardware hit count collection and reporting. The switch uses a fixed five minute logging interval, at which time trap log entries are written for each ACL logging rule that accumulated a non-zero hit count during that interval. The hit count is the number of times the rule has been invoked since the expiry of the last logging interval. It is not possible to configure the logging interval.

What Are Time-Based ACLs?

The time-based ACL feature allows the switch to dynamically apply an explicit ACL rule within an ACL for a predefined time interval by specifying a time range on a per-rule basis within an ACL, so that the time restrictions are imposed on the ACL rule.

With a time-based ACL, one can define when and for how long an individual rule of an ACL is in effect. To apply a time to an ACL, first define a specific time interval and then apply it to an individual ACL rule so that it is operational only during the specified time range, for example, during a specified time period or on specified days of the week.

A time range can be absolute (specific time) or periodic (recurring). If an absolute and periodic time range entry are defined within the same time range, the periodic timer is active only when the absolute timer is active.

NOTE: Adding a conflicting periodic time range to an absolute time range will cause the time range to become inactive. For example, consider an absolute time range from 8:00 AM Tuesday March 1st 2011 to 10 PM Tuesday March 1st 2011. Adding a periodic entry using the 'weekend' keyword will cause the time-range to become inactive because Tuesdays are not on the weekend.

A named time range can contain up to 10 configured time ranges. Only one absolute time range can be configured per time range. During the ACL configuration, a configured time range can be associated with the ACL to provide additional control over permitting or denying a user access to network resources.

Benefits of using time-based ACLs include:

- Providing more control over permitting or denying a user access to resources, such as an application (identified by an IP address/mask pair and a port number).
- Providing control of logging messages. Individual ACL rules defined within an ACL can be set to log traffic only at certain times of the day so access can simply be denied without the need to analyze many logs generated during peak hours.

ACL Limitations

There are two hardware matching engines visible to the Dell switch administrator: the ingress processor and the egress processor. Each of these processors has different limits and actions. The ingress matching engine processes packets on ingress to the switch and can apply actions such as applying CoS processing, diverting to a different port, etc. The egress matching engine processes packets after they are switched and queued for egress and supports policies such as rewriting the DSCP or CoS values, as well as the normal permit (forward) and deny (drop) actions.

ACLs operate by matching on specific fields within packets. Various match conditions (operators) are supported (e.g., equal, less than, not equal, etc.), along with masks that support selection of all or a portion of a field. Each field to be matched is assigned to a matching engine (a slice). A slice is defined by an offset into the packet that is compared against a set of matching values and masks along with an associated action (ACEs). Each Dell Networking N-Series switch supports a fixed number of slices and each slice supports a fixed number of matching criteria (values and masks). Slices operate in parallel to perform the configured matching operations. An ACL with a different offset requires the use of a new hardware slice but multiple matching values can be specified for a single slice (e.g., an IPv4 destination address with a 32-bit mask is 192.168.21.1 or 192.168.12.3). Slices can also be joined together to match widths larger than 32 bits or they can be concatenated to provide a larger number of matching values with a single offset. In general, ACLs that match on less than 32 bits will be expanded internally to match on 32 bits with a variable mask. This allows other ACLs using the same offset to utilize the same slice with potentially different masks and match values.

The user interface limits for ACLs are 1023 rules per access list and 100 access lists. The switch automatically combines slices to operate in parallel over greater field widths (e.g., IPv6 source address) or combines slices to supply more match conditions (IPv4 destination address equal to multiple ranges of addresses). In the case of a match condition specifying a match wider than 32 bits (e.g., a 128-bit IPv6 address), additional slices are assigned to operate in parallel on the additional match fields. This reduces the overall number of slices available to match on other key fields. The switch attempts to assign slices to match conditions in an optimal manner; however, combinations of match conditions can reduce the maximum number of ACLs that can be configured to fewer than the published limits. As an example, the smallest IPv6 QoS match will utilize six slices in the switch hardware.

If encountering a situation where the hardware limit is exceeded when configuring an ingress ACL, consider disabling features that use ACLs internally, such as iSCSI or CFM.

The Dell Networking N4000 Series switches support the following hardware limits:

- 2047 ingress rules and 1023 egress rules, for a total of 3072 rules.
- The hardware has 10 ingress slices and 4 egress slices, with 4 ingress slices having a depth of 128 rules, and 6 ingress slices having a depth of 256 rules. The egress slices have a depth of 256 rules.

The Dell Networking N3000 Series switches support the following hardware limits:

- 3072 ingress rules and 1024 egress rules, for a total of 4096 rules.

- The hardware has 14 ingress slices and 4 egress slices, with the 14 ingress slices having a depth of 256 rules. The egress slices have a depth of 256 rules.

The Dell Networking N2000 Series switches support the following hardware limits:

- 1024 ingress rules and 512 egress rules, for a total of 1536 rules.
- The hardware has 14 ingress slices and 4 egress slices, with the 14 ingress slices having a depth of 256 rules. The egress slices have a depth of 256 rules.

The Dell Networking N1500 Series switches support the following hardware limits:

- 1024 ingress rules and 512 egress rules, for a total of 1536 rules.
- The hardware has 6 ingress slices and 4 egress slices, with the ingress slices having a depth of 256 rules. The egress slices have a depth of 128 rules.

The software limits are shown in Table 20-1:

Table 20-1. ACL Software Limits

Limitation	Dell Networking N1500 Series	Dell Networking N2000 Series	Dell Networking N3000 Series	Dell Networking N4000 Series
Maximum number of ACLs (any type)	100	100	100	100
Maximum number of configurable rules per list.	1023	1023	1023	1023
Maximum ACL Rules per Interface and Direction (IPv4/L2)	1023 ing., 1023 egr.	1023 ing., 1023 egr.	2048 ing., 1023 egr.	1023 ing., 1023 egr.
Maximum ACL Rules per Interface and Direction (IPv6)	255 ing., 125 egr.	1023 ing., 509 egr.	1659 ing., 509 egr.	635 ing., 509 egr.
Maximum ACL Rules (system-wide)	2030	3914	3914	3060
Maximum VLAN interfaces with ACLs applied	24	24	24	24
Maximum ACL Logging Rules (system-wide)	128	128	128	128

Please note the following additional limitations on ingress and egress ACLs:

- Port ranges are not supported for egress ACLs for either IPv4 or IPv6 ACLs.
- It is possible to configure mirror or redirect attributes for a given ACL rule, but not both.
- The Dell Networking N-Series switches support a limited number of counter resources, so it may not be possible to log every ACL rule. It is possible to define an ACL with any number of logging rules, but the rules that are actually logged cannot be determined until the ACL is configured in the interface hardware. Furthermore, hardware counters that become available after an ACL is applied are not retroactively assigned to rules that were unable to be logged (the ACL must be disassociated from the interface and then re-associated). Rules that are unable to be logged are still active in the ACL for purposes of permitting or denying a matching packet. If console logging is enabled and the severity is set to a numerically equal or lower severity than the console severity setting, a log entry may appear on the screen.
- The order of the rules is important: when a packet matches multiple rules, the first rule takes precedence. Once a packet has matched a rule, the corresponding action is taken and no further attempts to match the packet are made. Also, once an access group is configured on an interface, all traffic not specifically permitted by an ACL is dropped by the implicit deny all the system supplies at the end of the last configured access group.
- Egress (out) ACLs only affect switched/routed traffic. They have no effect on packets generated locally by the switch, e.g., LACPDUs or spanning tree BPDUs.
- Ingress ACLs filter packets before they are processed by the switching fabric. Egress ACLs filter packets after they have been processed by the switching fabric.
- User-defined ingress ACLs are prioritized before system ACLs. User-defined ingress ACLs that match control plane packets such as BPDUs may interfere with switch operation.
- The **fragments** and **routing** keywords are not supported for egress IPv6 ACLs. The **fragments** keyword is not supported on IPv4 egress ACLs.
- On the Dell Networking N4000 Series switches, the IPv6 ACL **routing** keyword is not supported when any IPv6 address is specified. The **routing** keyword is not support for IPv4 ACLs.

- On the Dell Networking N4000 Series switches, the IPv6 ACL **fragment** keyword matches only on the first two IPv6 extension headers for the fragment header (next header code 44). If the fragment header appears in the third or subsequent header, it is not matched
- On the Dell Networking N2000 and N3000 Series switches, the IPv6 ACL **fragment** keyword matches only on the first IPv6 extension header (next header code 44). If the fragment header appears in the second or subsequent header, it is not matched.
- The IPv6 ACL **routing** keyword matches only on the first IPv6 extension header (next header code 43). If the fragment header appears in the second or subsequent header, it is not matched.

NOTE: The actual number of ACLs and rules supported depends on the resources consumed by other processes and configured features running on the switch. If the switch does not allow a rule to be configured, consider disabling features that consume user ACL space such as iSCSI, CFM, or IPv6 RA Guard.

ACL Configuration Details

How Are ACLs Configured?

To configure ACLs, follow these steps:

1. Create a IP or MAC ACL by specifying a name.
2. Add new rules to the ACL.
3. Configure the match criteria for the rules.
4. Apply the ACL to one or more interfaces.

Editing Access Lists

When editing access lists, entries are added in the order specified by the rule sequence number. It is recommended that rule sequence number indices be separated by a fixed offset (e.g., 10). The ACL sequence number can range from 1 to 2147483647.

If no sequence number is specified, new entries are added to the end of the list. There is an implicit deny all statement at the end of the last access-group that is not shown and is not editable. To insert a rule in the middle of an

ACL, enter a sequence number less than the following rule and greater than the preceding rule. Use the **no** [***sequence-number***] command in ACL Configuration mode to remove rules from an ACL.

NOTE: When configuring access lists, complete checks are made only when the access list is applied to an active interface. It is recommended that you configure and test an access list on an active (up) interface prior to deploying it on links in the production network. If an ACL is configured on an interface that is not up, error messages regarding ACL resource allocation may be logged when the interface is brought up.

Preventing False ACL Matches

Be sure to specify ACL access-list, permit, and deny rule criteria as fully as possible to avoid false matches. This is especially important in networks with protocols that have different frame or EtherType values. For example, layer-3 ACL rules that specify a TCP or UDP port value should also specify the TCP or UDP protocol. MAC ACL rules that specify an EtherType value for the frame should also specify a source or destination MAC address wherever possible. Likewise, MAC ACLs that specify a source MAC address should specify an Ethertype to avoid interfering with control-plane traffic.

In general, any rule that specifies matching on an upper-layer protocol field should also include matching constraints for as many of the lower-layer as where possible. For example, a rule to match packets directed to the well-known UDP port number 22 (SSH) should also include matching constraints on the IP protocol field (protocol=0x11 or UDP) and the source or destination IP address. Table 20-2 lists commonly-used EtherTypes numbers:

Table 20-2. Common EtherType Numbers

EtherType	Protocol
0x0800	Internet Protocol version 4 (IPv4)
0x0806	Address Resolution Protocol (ARP)
0x0842	Wake-on LAN Packet
0x8035	Reverse Address Resolution Protocol (RARP)
0x8100	VLAN tagged frame (IEEE 802.1Q)
0x86DD	Internet Protocol version 6 (IPv6)

Table 20-2. Common EtherType Numbers (Continued)

EtherType	Protocol
0x8808	MAC Control
0x8809	Slow Protocols (IEEE 802.3)
0x8870	Jumbo frames
0x888E	EAP over LAN (EAPOL – 802.1x)
0x88CC	Link Layer Discovery Protocol
0x8906	Fibre Channel over Ethernet
0x9100	Q in Q

Figure 20-3 lists commonly-used IP protocol numbers:

Table 20-3. Common IP Protocol Numbers

IP Protocol Number	Protocol
0x00	IPv6 Hop-by-hop option
0x01	ICMP
0x02	IGMP
0x06	TCP
0x08	EGP
0x09	IGP
0x11	UDP

Using IP and MAC Address Masks

Masks are used with IP and MAC addresses to specify what should be considered in the address for a match. Masks are expanded internally into a bit mask and are applied bit-wise in the hardware even though they are entered in decimal or hexadecimal format. Masks need not have contiguous 0 or 1 bits. A 0 bit value in the mask indicates that the address field in the packet being compared must match the address bit exactly. A 1 value in the mask indicates a wildcard or don't care value, i.e. the access bits are not compared and match any possible value. For example, an IP address of 3.3.3.3 with a mask of 0.0.0.0 indicates that the ACL matches on all four bytes of the

IP address. Likewise, a MAC address of 68:94:23:AD:F3:18 with a mask of 00:00:00:00:00:ff indicates that the first five bytes must match (e.g., 68:94:23:AD:F3) and the last byte may take on any value from 0x00 to 0xff (0–255) and still be considered a match.

The following ACL equivalents are noted:

Address	Mask	Equivalent Address
0.0.0.0	255.255.255.255	any
x.x.x.x	host	x.x.x.x
0:0:0:0:0:0	ff:ff:ff:ff:ff:ff	any

Policy-Based Routing

In contemporary inter-networks, network administrators often need to implement packet forwarding/routing according to specific organizational policies. Policy-Based Routing (PBR) exactly fits this purpose. Policy-Based Routing provides a flexible mechanism to implement solutions where organizational constraints dictate that traffic be routed through specific network paths. PBR does not affect route redistribution that occurs via routing protocols.

PBR is a true routing policy solution. The packet TTL is decremented in PBR-routed packets. The destination MAC is rewritten in PBR-routed packets. ARP lookups are sent when required for unresolved next-hop addresses.

Configuring PBR consists of installing a route-map with **match** and **set** commands and then applying the corresponding route-map to the routing VLAN interface. IP routing must be enabled on the interfaces by assigning IP addresses to the VLAN interfaces, assigning the VLANs to physical interfaces, and enabling IP routing globally.

Packet Classification

Route-maps may specify multiple packet attributes in match statements. These attributes can be matched through a “match” clause based on length of the packet or a “match” clause linked with up to 16 ACLs.

The match attributes listed below for each ACL type indicate the criteria used to classify layer-3 routed traffic for PBR. At least one of the listed attributes must be present in the ACL of the given type:

- VLAN tag (implicitly added)
- MAC access list (**match mac-list**)
  - Source MAC address
  - 802.1p priority
- IP access list (**match ip address**)
  - Source or destination IP address
- Protocol ID field in the IP header
- L3 packet length in the IP header (**match length**)

Additional match criteria may be configured by the administrator if desired. Since a route-map is configured in the context of a routing VLAN, a VLAN tag is automatically added to the match criteria without the need for the administrator to specify the VLAN ID.

Route-Map Processing

An incoming packet is matched against the criteria in the 'match' terms specified in each route-map in the policy. The 'match' terms (clauses) must refer to one or more MAC or IPv4 addresses or a packet length. Multiple MAC or IPv4 match terms are allowed in a route-map, each consisting of a list of ACLs.

Conceptually, ACL processing proceeds by attempting to match each of the ACLs listed in the first match clause, in order. If an ACL does not match, processing moves to the next ACL, in order, until an ACL matches or the ACL list is exhausted. If there are more match terms in the route-map, processing proceeds with the next match term, in order. In reality, all ACLs matches are attempted in parallel at once, and the priority of the ACL is used to implement the conceptual match process.

An ACL that is used in a 'match' term itself has one or more permit and/or deny rules. The incoming packet is matched sequentially against the permit rules in each ACL in the match list, in order, and a permit/deny decision is reached. If a permit rule in an ACL in the list matches, the ACL match term criteria is met and no further match processing takes place in the route-map. If none of the permit rules in an ACL matches, the packet match is attempted against the next ACL in the route-map match list. Deny ACLs are optimized out of both permit and deny route-maps and are not processed.

Once a match has occurred:

- For a permit route-map, if the decision reached in the above step is permit, then PBR executes the action specified in the **set** term(s) of the route-map statement. The counter for the route-map is incremented for each matching packet.
- For a permit route-map, if the decision reached in the above step is deny, then PBR does not apply any action that is specified in **set** term(s) in the route-map statement. In this situation, the counter for this match statement is not incremented. The processing logic terminates, and the packet goes through the standard destination-based routing logic.

- For a deny route-map, if the decision reached in the above step is permit, then PBR processing logic terminates and the packet goes through standard destination-based routing logic. The counter is incremented for each matching packet.
- For a deny route-map, if the decision reached in the above step is deny, the counter for this match statement is not incremented. The processing logic terminates, and the packet goes through the standard destination-based routing logic.

PBR counters increment when a packet matches the corresponding ACL. They do not indicate the outcome of the processing logic; i.e., PBR counters do not count packets that are policy-routed vs. not policy-routed. ACL packet matching occurs in parallel across all ACLs. If a policy ACL matches a packet, and an interface or VLAN ACL also matches the packet, the PBR counter may be incremented even though the interface or VLAN ACL caused the packet to be dropped.

If no match occurs, then the packet goes through the standard destination-based routing logic.

Route-Map Actions

Policy-Based Routing overrides the normal routing decisions taken by the router and attempts to route the packet using the criteria in the set clause:

- List of next-hop IP addresses—The **set ip next-hop command** checks for the next-hop address in the routing table and, if the next-hop address is present and active in the routing table, then the policy routes the ACL matching packets to the next hop. If the next hop is not present in the routing table, the command uses the normal routing table to route the packet. Non-matching packets are routed using the normal routing table. The IP address must specify an adjacent next-hop router in the path toward the destination to which the packets should be routed. The first available IP address associated with a currently active routing entry is used to route the packets. This type of rule takes priority over all entries in the routing table.
- List of default next-hop IP addresses—The **set ip default next-hop** command checks the list of destination IP addresses in the routing table and, if there is no explicit route for the packet's destination address in the routing table, the next-hop destinations are evaluated, and packets are routed to the first-available next hop. Packets that do not match are routed

using the routing table. A default route in the routing table is not considered an explicit route for an unknown destination address. This type of rule takes priority over default entries in the routing table.

- IP precedence—Packets matching the ACL criteria have their IP precedence rewritten. The IP precedence value is the 4 ToS bits in the IP packet header.

The following table summarizes the actions taken by the routing policy depending on the combination of ACL and route-map permit and deny rules:

ACL	Match	Route-Map	Outcome	Action	Optimized
Permit	Yes	Permit	Permit	Set	No
Permit	No	Permit	Deny	Next	No
Permit	Yes	Deny	Permit	Route	No
Permit	No	Deny	Deny	Next	No
Deny	Yes	Permit	Deny	Next	Yes
Deny	No	Permit	Deny	Next	Yes
Deny	Yes	Deny	Deny	Next	Yes
Deny	No	Deny	Deny	Next	Yes

In the table, the Action column indicates the disposition of the packet:

- *Next* means fall through to next route-map, and if there are no further route-maps to be processed, route the packet using the default routing table.
- *Set* means route the packet per the action in the set clause.
- *Route* means route the packet with the default routing table.

In the last column of the table (Optimized), a Yes entry means the rule is never processed in hardware because the action, if any, is to fall through to the next match criteria. The system optimizes out deny ACL match clauses and never processes them in the system hardware. Counters for these match clauses will always show 0.

Interface ACLs and PBR Interaction

PBR can be configured only on VLAN routing interfaces. However, user-defined ACLs can be configured on all types of interfaces, including physical interfaces, port-channels, and VLANs. When processing packets on which both PBR and user-defined ACLs are configured, routing policy is performed only after the application of all user-defined VLAN and interface ACLs.

Only packets that do not match any user-defined ingress deny ACLs rules configured on an incoming interface are eligible for processing by PBR. User-defined interface ACLs have a higher precedence than user-defined VLAN ACLs or PBR ACLs. In the case of conflicting actions, the user-defined interface ACL takes precedence. Specifically, if a user-defined interface ACL drops a packet (deny), routing policy is not applied to the packet. Likewise, if a user defined VLAN interface ACL drops a packet, routing policy is not applied to the packet.

In many cases, the switch is capable of taking multiple actions on a packet, irrespective of whether the action is configured on an ACL used in a route-map or on an ACL configured on a port. For example, the system can both rate limit packets on ingress with an interface ACL and set the ip precedence on packets that do not exceed the rate limit with a PBR ACL.

The following table describes the action resolution mechanism when a packet matches both the PBR rules configured on a VLAN routing interface and a permit ACL rule configured on a physical interface (the deny ACL action is included for emphasis):

PBR Action (VLAN)	ACL Action (Interface)	Result
set ip precedence	deny	deny
	mirror	both
	redirect	both (see Note 1)
	rate limit	both
set interface null0	deny	deny (see Note 2)
	mirror	mirror
	redirect	redirect
	rate limit	deny
set ip next-hop (default)	deny	deny

PBR Action (VLAN)	ACL Action (Interface)	Result
	mirror	both
	redirect	both (see Note 1)
	rate limit	both

1. In the case of redirect ACL action, both the redirect and PBR actions are honored, if possible. This implies the PBR routed packet is redirected to the configured physical port and the redirected port is participating in the egress VLAN to which the packet is being routed. In other words, the system will select the interface specified by the ACL which is a member of the egress VLAN. If the physical interface is not a member of the egress VLAN, the behavior is undefined.

2. In case of the PBR **set interface Null0** action, the PBR routed packet is dropped only if no conflicting port ACL is configured. Configuring ACL deny statements that also match packets with a PBR **set interface Null0** action is redundant and wastes system resources.

PBR and Implicit Deny-all

Configuring a routing policy on an interface overrides the implicit “deny all” ACL at the end of the last interface access group. Administrators should ensure all appropriate ACL deny rules are configured on an interface on which PBR is configured in order to ensure system security.

Limitations

Internally Generated Packets

Packets that are generated internally by the router are never policy routed.

Set Clause Required

Route-map deny/permit statements without “set” clauses are ignored except in the case where a deny route-map refers to an ACL with a permit statement.

No Implicit “deny all” Rule

When an access-group is configured on an interface, an implicit rule of “deny all” is applied to the last access-group on the interface. Since PBR processing occurs after normal ACL processing, when a “permit” route-map associated ACL is applied to an interface, the implicit “deny all” rule is not applied. When match rules in an ACL associated with a route-map are successful, packets are considered as candidates for routing according to rules specified in route-map. If none of the match rules are successful, then packet is routed by the standard L3 routing process. The implicit "deny all" rule is not applicable to interfaces on which a routing policy is configured. Configuring an explicit deny all ACL that not associated with a route-map will drop packets prior to them being processed by PBR.

Black Holes Possible

If the next hop specified by a policy-based rule is not reachable, packets matching the ACL are routed using the routing table. If the routing table does not supply a route to the destination, then the packets are lost. If a set interface null0 statement is present in the policy map, the packets are dropped. The set interface null0 statement can also be used to drop undesirable or unwanted traffic, i.e. create a black hole route.

Counter Support for Route-map ACL

A counter is associated with each ACL rule associated with a route-map in order to indicate how many packets have been policy routed. There is no provision to non-destructively clear these counters from the UI. Counters associated with route-map statement are cleared when the route-map is removed from the VLAN. The hardware does not support both a counter and a rate-limit. Therefore, the system does not support configuring ACLs with a rate-limit being used for PBR. In this case, a separate interface or VLAN ACL with a rate-limit can be used at the cost of consuming additional resources.

Packets matching PBR-associated ACLs that contain deny statements are not counted. Deny ACLs in PBR rules are optimized out of the system as they always fall through to the next PBR statement.

PBR Associated ACLs Processed After User-defined ACLs

Each ACL in an access-group is associated with a sequence number indicating the order in which the ACL is processed by the hardware. Likewise, a route-map may have multiple statements with different sequence numbers associated with each ACL entry. These statements are processed in sequential order beginning with the lowest numbered rule, but only after all user configured ACLs that are not associated with any route-map.

ACL Resource Usage

When a route-map defines a “match” rule associated with an ACL, except for the implicit routing behavior mentioned above, the resource consumption is the same as if a normal ACL is applied on an interface. Rules consumed by an ACL corresponding to route-map “match” clause share hardware resources with the ACL component. Some resources cannot be shared. For example, it is not permitted to utilize the rate-limit clause in a PBR ACL, as the hardware cannot support both a counter (allocated by every PBR route-map) and a rate limit.

ACLs associated with a route-map and general ACLs share the same hardware resources. If PBR consumes the maximum number of hardware resources on an interface/system wide, general purpose ACLs can't be configured later and vice versa. Hardware allocation is performed on a first-come first-serve basis when the interface becomes active.

ACL Resource Sharing

An ACL rule contains match and action attributes. For example, an ACL rule may have a match clause on source IP address and action attributes independent of PBR such as queue assignment as shown below:

```
console#config
console(config)#ip access-list example-1
console(config-ip-acl)#permit ip 1.1.1.1 0.0.0.255 any assign-queue
2
console(config-ip-acl)#permit every
console(config-ip-acl)#exit
```

Actions specified in the “set” clauses of a route-map utilize the hardware entries of the corresponding ACL. This sharing does not consume additional hardware resources as Dell Networking supports multiple actions in an ACL rule. However, if conflicting actions are specified, an error is thrown when the switch attempts to configure the conflicting actions in the hardware.

No IPv6 support

PBR does not support IPv6 match ACLs.

Locally Generated Packets

Policy-Based Routing does not affect locally generated packets, i.e. packets generated by protocols running on the switch.

Configuring ACLs (Web)

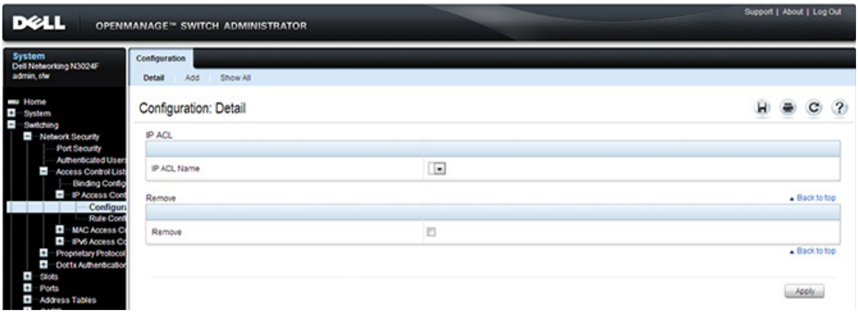
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring ACLs on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

IP ACL Configuration

Use the **IP ACL Configuration** page to add or remove IP-based ACLs.

To display the **IP ACL Configuration** page, click **Switching** → **Network Security** → **Access Control Lists** → **IP Access Control Lists** → **Configuration** in the navigation panel.

Figure 20-1. IP ACL Configuration



Adding an IPv4 ACL

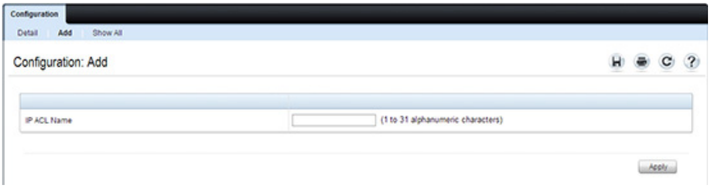
To add an IPv4 ACL:

1 Open the **IP ACL Configuration** page.

2 Click **Add** to display the **Add IP ACL** page.

3 Specify an ACL name.

Figure 20-2. Add IP ACL



4 Click **Apply**.

Removing IPv4 ACLs

To delete an IPv4 ACL:

1 From the **IP ACL Name** menu on the **IP ACL Configuration** page, select the ACL to remove.

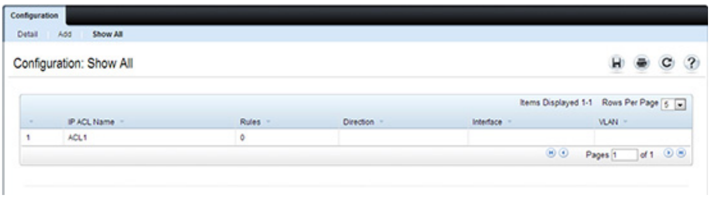
2 Select the **Remove** checkbox.

3 Click **Apply**.

Viewing IPv4 ACLs

To view configured ACLs, click **Show All** from the **IP ACL Configuration** page.

Figure 20-3. View IPv4 ACLs



IP ACL Rule Configuration

Use the **IP ACL Rule Configuration** page to define rules for IP-based ACLs. The access list definition includes rules that specify whether traffic matching the criteria is forwarded normally or discarded. Additionally, rules can be used to assign traffic to a particular queue, filter on some traffic, change a VLAN tag, shut down a port, and/or redirect the traffic to a particular port.

NOTE: There is an implicit **deny all** rule at the end of an ACL list. This means that if an ACL is applied to a packet and if none of the explicit rules match, then the final implicit "deny all" rule applies and the packet is dropped.

To display the **IP ACL Rule Configuration** page, click **Switching → Network Security → Access Control Lists → IP Access Control Lists → Rule Configuration** in the navigation panel.

Figure 20-4. IP ACL - Rule Configuration

Removing an IP ACL Rule

To delete an IP ACL rule:

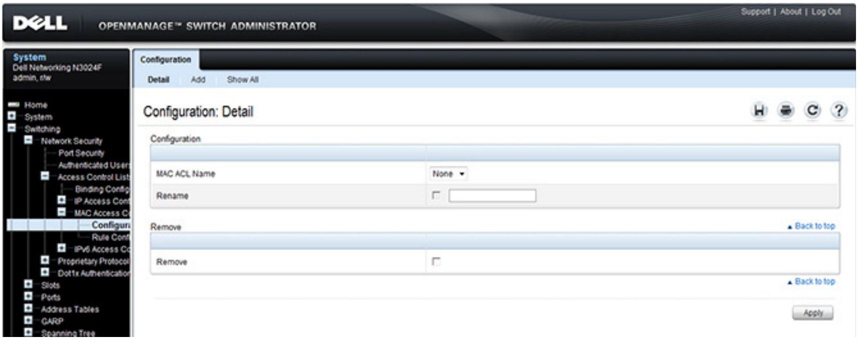
1. From the **Rule ID** menu, select the ID of the rule to delete.
2. Select the **Remove** option near the bottom of the page.
3. Click **Apply** to remove the selected rule.

MAC ACL Configuration

Use the **MAC ACL Configuration** page to define a MAC-based ACL.

To display the **MAC ACL Configuration** page, click **Switching** → **Network Security** → **Access Control Lists** → **MAC Access Control Lists** → **Configuration** in the navigation panel.

Figure 20-5. MAC ACL Configuration



Adding a MAC ACL

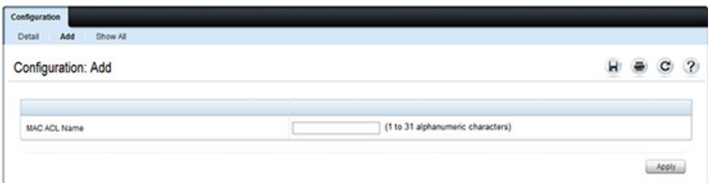
To add a MAC ACL:

1 Open the **MAC ACL Configuration** page.

2 Click **Add** to display the **Add MAC ACL** page.

3 Specify an ACL name.

Figure 20-6. Add MAC ACL



4 Click **Apply**.

Renaming or Removing MAC ACLs

To rename or delete a MAC ACL:

1. From the **MAC ACL Name** menu on the **MAC ACL Configuration** page, select the ACL to rename or remove.
2. To rename the ACL, select the **Rename** checkbox and enter a new name in the associated field.
3. To remove the ACL, select the **Remove** checkbox.
4. Click **Apply**.

Viewing MAC ACLs

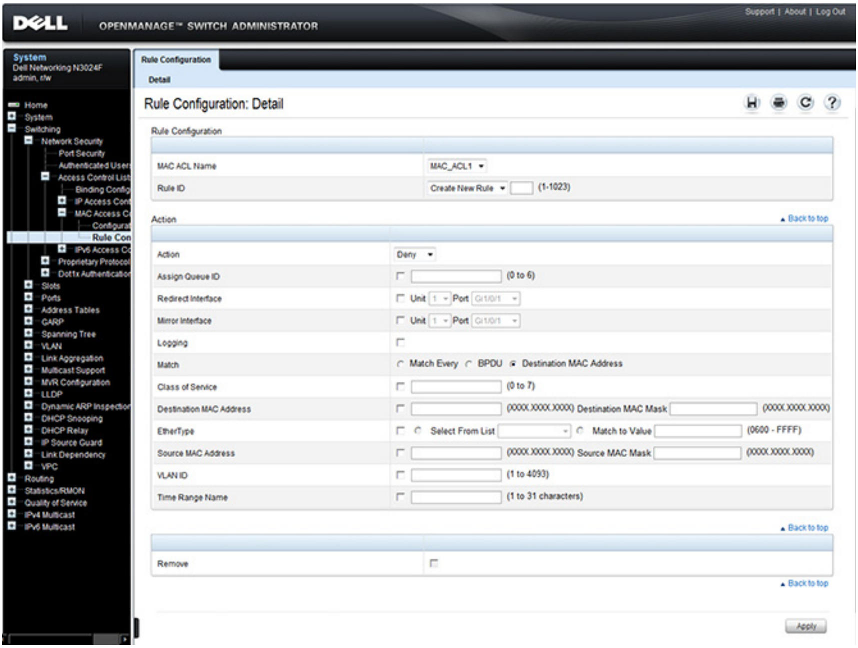
To view configured ACLs, click **Show All** from the **MAC ACL Configuration** page.

MAC ACL Rule Configuration

Use the **MAC ACL Rule Configuration** page to define rules for MAC-based ACLs. The access list definition includes rules that specify whether traffic matching the criteria is forwarded normally or discarded. A default **deny all** rule is the last rule of every list.

To display the **MAC ACL Rule Configuration** page, click **Switching** → **Network Security** → **Access Control Lists** → **MAC Access Control Lists** → **Rule Configuration** in the navigation panel.

Figure 20-7. MAC ACL Rule Configuration



Removing a MAC ACL Rule

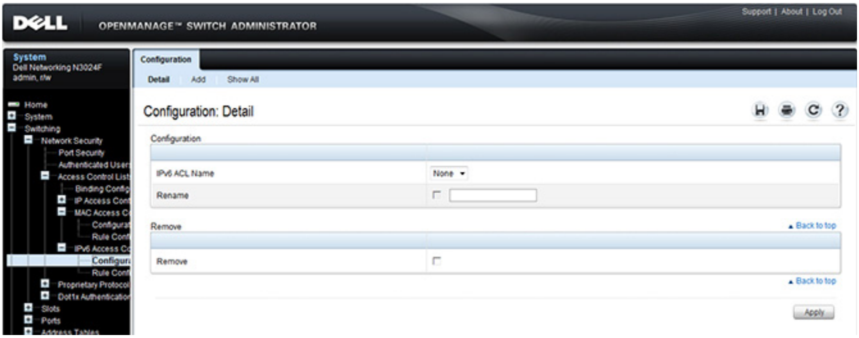
To delete a MAC ACL rule:

1. From the **Rule ID** menu, select the ID of the rule to delete.
2. Select the **Remove** option near the bottom of the page.
3. Click **Apply** to remove the selected rule.

IPv6 ACL Configuration

Use the **IPv6 ACL Configuration** page to add or remove IP-based ACLs. To display the IP ACL Configuration page, click **Switching** → **Network Security** → **Access Control Lists** → **IPv6 Access Control Lists** → **IPv6 ACL Configuration** in the navigation panel.

Figure 20-8. IPv6 ACL Configuration

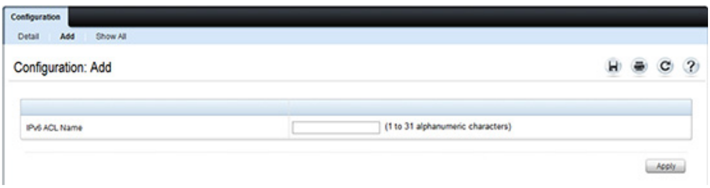


Adding an IPv6 ACL

To add an IPv6 ACL:

1. Open the **IPv6 ACL Configuration** page.
2. Click **Add** to display the **Add IPv6 ACL** page.
3. Specify an ACL name.

Figure 20-9. Add IPv6 ACL



4. Click **Apply**.

Removing IPv6 ACLs

To delete an IPv6 ACL:

1 From the **IPv6 ACL Name** menu on the **IPv6 ACL Configuration** page, select the ACL to remove.
2 Select the **Remove** checkbox.
3 Click **Apply**.

Viewing IPv6 ACLs

To view configured ACLs, click **Show All** from the **IPv6 ACL Configuration** page. The **IPv6 ACL Table** page displays.

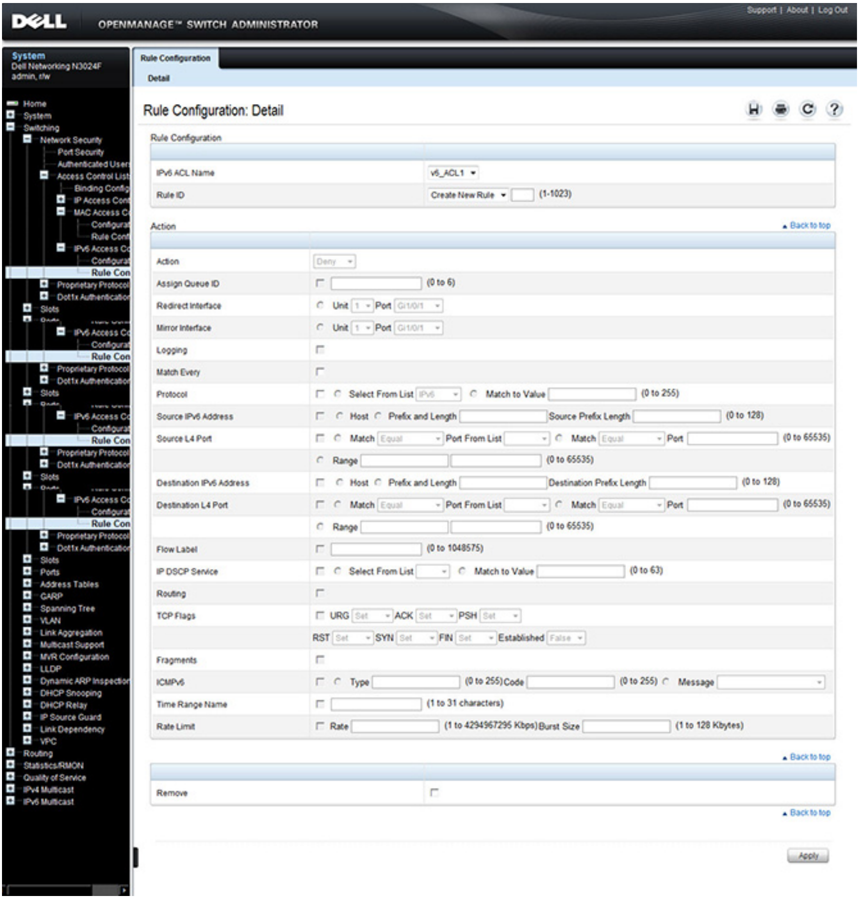
IPv6 ACL Rule Configuration

Use the IPv6 ACL Rule Configuration page to define rules for IPv6-based ACLs. The access list definition includes rules that specify whether traffic matching the criteria is forwarded normally or discarded. Additionally, rules can specify to assign traffic to a particular queue, filter on some traffic, change VLAN tag, shut down a port, and/or redirect the traffic to a particular port. By default, no specific value is in effect for any of the IPv6 ACL rules.

There is an implicit **deny all** rule at the end of an ACL list. This means that if an ACL is applied to a packet and if none of the explicit rules match, then the final implicit **deny all** rule applies and the packet is dropped.

To display the IPv6 ACL Rule Configuration page, click **Switching → Network Security → Access Control Lists → IPv6 Access Control Lists → Rule Configuration** in the navigation menu.

Figure 20-10. IPv6 ACL - Rule Configuration



Removing an IPv6 ACL Rule

To delete an IPv6 ACL rule:

1. From the **Rule ID** menu, select the ID of the rule to delete.
2. Select the **Remove** option near the bottom of the page.
3. Click **Apply** to remove the selected rule.

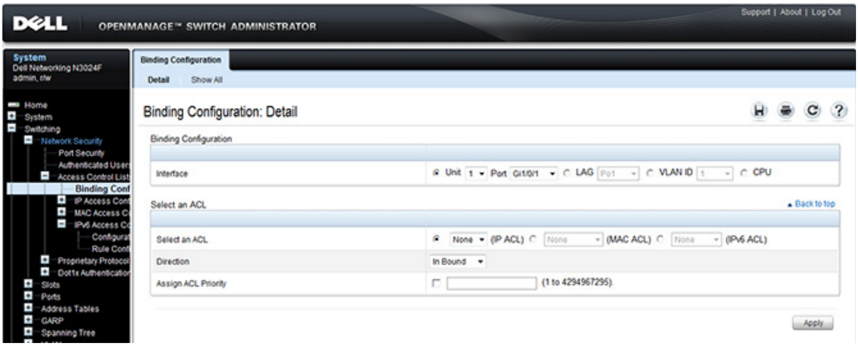
ACL Binding Configuration

When an ACL is bound to an interface, all the rules that have been defined are applied to the selected interface. Use the **ACL Binding Configuration** page to assign ACL lists to ACL Priorities and Interfaces.

From the web interface, the ACLs rules can be configured in the ingress or egress direction so that they implement security rules for packets entering or exiting the port. ACLs can be applied to any physical (including 10 Gb) interface, LAG, or routing port.

To display the **ACL Binding Configuration** page, click **Switching** → **Network Security** → **Access Control Lists** → **Binding Configuration** in the navigation panel.

Figure 20-11. ACL Binding Configuration

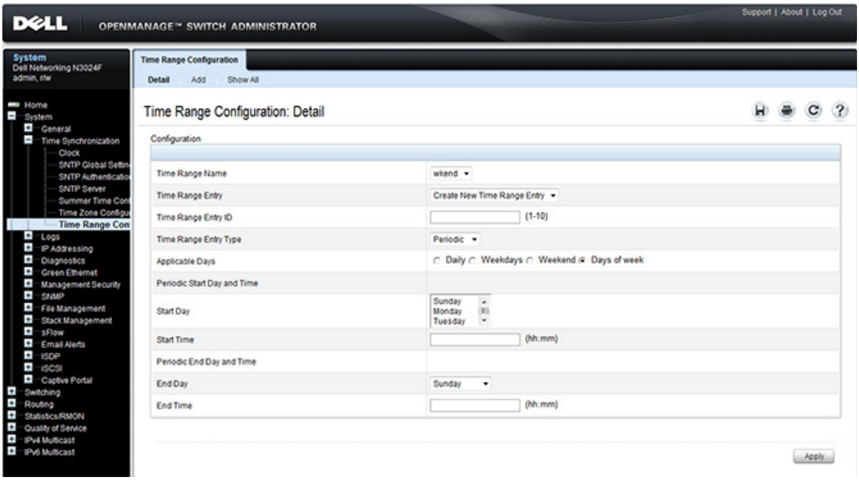


Time Range Entry Configuration

Use the **Time Range Entry Configuration** page to define time ranges to associate with ACL rules.

To display the **Time Range Entry Configuration** page, click **System** → **Time Synchronization** → **Time Range Configuration** in the navigation panel. The following image shows the page after at least one time range has been added. Otherwise, the page indicates that no time ranges are configured, and the time range configuration fields are not displayed.

Figure 20-12. Time Range Configuration



NOTE: A time-range parameter in an ACL that is referred to by a route-map statement is active only during the time range specified. When the ACL is not active (outside the time range), the route-map simply treats the ACL as a “no match”.

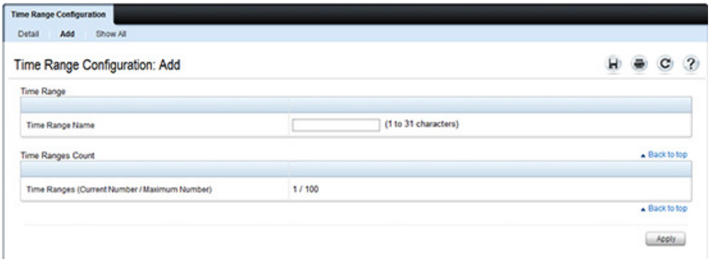
Adding a Time Range

To configure a time range:

1 From the **Time Range Entry Configuration** page, click **Add**.

2 Specify a name to identify the time range.

Figure 20-13. Add a Time Range



3 Click **Apply**.

4 Click **Configuration** to return to the **Time Range Entry Configuration** page.

5 In the **Time Range Name** field, select the name of the time range to configure.

6 Specify an ID for the time range. Up to 10 different time range entries can be configured to include in the named range. However, only one absolute time entry is allowed per time range.

7 Configure the values for the time range entry.

8 Click **Apply**.

9 To add additional entries to the named time range, repeat step 5 through step 8.

Configuring ACLs (CLI)

This section provides information about the commands you use to create and configure ACLs. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring an IPv4 ACL

Beginning in Privileged EXEC mode, use the following commands to create an IPv4 ACL, configure rules for the ACL, and bind the ACL to an interface.

NOTE: The **ip access-group** command can be issued in Global Configuration mode or Interface configuration mode. If it is applied in Global Configuration mode, the ACL binding is applied to all interfaces. If it is applied in Interface Configuration mode, it is applied only to the specified interfaces within the mode.

Command	Purpose
configure	Enter global configuration mode.
ip access-list *name*	Create an extended ACL and enter IPv4 access-list configuration mode.

<table>
<tr><th>Command</th><th>Purpose</th></tr>
<tr><td>[sequence-number] {deny | permit} {every | {{ipv4-protocol | 0-255 | every} {srcip srcmask | any | host srcip} [{range {portkey | startport} {portkey | endport} | {eq | neq | lt | gt} {portkey | 0-65535}] {dstip dstmask | any | host dstip} [{range {portkey | startport} {portkey | endport} | {eq | neq | lt | gt} {portkey | 0-65535}] [flag [+fin | -fin] [+syn | -syn] [+rst | -rst] [+psh | -psh] [+ack | -ack] [+urg | -urg] [established]] [icmp-type icmp-type [icmp-code icmp-code] | icmp-message icmp-message] [igmp-type igmp-type] [fragments] [precedence precedence | tos tos [tosmask] | dscp dscp]}} [time-range time-range-name] [log] [assign-queue queue-id] [{mirror | redirect} unit/slot/port] [rate-limit rate burst-size]</td>
<td>Enter the permit and deny conditions for the extended ACL.
• sequence-number — Identifies the order of application of the permit/deny statement. If no sequence number is assigned, permit/deny statements are assigned a sequence number beginning at 1000 and incrementing by 10. Statements are applied in hardware beginning with the lowest sequence number. Sequence numbers apply only within an access group; i.e., the ordering applies within the access-group scope. The range for sequence numbers is 1–2147483647.
• {deny | permit} — Specifies whether the IP ACL rule permits or denies the matching traffic.
• {ipv4-protocol | number | every} — Specifies the protocol to match for the IP ACL rule.
– IPv4 protocols: eigrp, gre, icmp, igmp, ip, ipinip, ospf, tcp, udp, pim
– every: Match any protocol (don’t care)
• srcip srcmask | any | host srcip — Specifies a source IP address and netmask to match for the IP ACL rule.
– Specifying “any” implies specifying srcip as “0.0.0.0” and srcmask as “255.255.255.255” for IPv4.
– Specifying “host A.B.C.D” implies srcip as “A.B.C.D” and srcmask as “0.0.0.0”.
• [{{eq | neq | lt | gt} {portkey | number} | range startport endport}] — Specifies the layer-4 destination port match condition for the IP ACL rule. A destination port number, which ranges from 0-65535, can be entered, or a portkey, which can be one of the following keywords: domain, echo, ftp, ftp-data, http, smtp, snmp, telnet, tftp, and www. Each of these keywords translates into its equivalent destination port number.</td></tr>
</table>

Command	Purpose
continued	– When **range** is specified, IP ACL rule matches only if the layer-4 port number falls within the specified port range. The *startport* and *endport* parameters identify the first and last ports that are part of the port range. They have values from 0 to 65535. The ending port must have a value equal or greater than the starting port. The starting port, ending port, and all ports in between will be part of the layer-4 port range. – When **eq** is specified, IP ACL rule matches only if the layer-4 port number is equal to the specified port number or portkey. – When **lt** is specified, IP ACL rule matches if the layer-4 destination port number is less than the specified port number or portkey. It is equivalent to specifying the range as 0 to <specified port number – 1>. – When **gt** is specified, IP ACL rule matches if the layer-4 destination port number is greater than the specified port number or portkey. It is equivalent to specifying the range as <specified port number + 1> to 65535. – When **neq** is specified, IP ACL rule matches only if the layer-4 destination port number is not equal to the specified port number or portkey. – IPv4 TCP port names: **bgp, domain, echo, ftp, ftp-data, http, smtp, telnet, www, pop2, pop3** – IPv4 UDP port names: **domain, echo, ntp, rip, snmp, tftp, time, who** • *dstip dstmask* | **any** | **host** *dstip*—Specifies a destination IP address and netmask for match condition of the IP ACL rule. – Specifying **any** implies specifying *dstip* as "0.0.0.0" and *dstmask* as "255.255.255.255". – Specifying **host** A.B.C.D implies *dstip* as "A.B.C.D" and *dstmask* as "0.0.0.0". • [**precedence** *precedence* | **tos** *tos* [*tosmask*] | **dscp** *dscp*]—Specifies the TOS for an IP/TCP/UDP ACL rule depending on a match of precedence or DSCP values using the parameters **dscp**, **precedence**, or **tos** *tosmask*.

Command	Purpose
continued	• **flag** [**+fin** | **-fin**] [**+syn** | **-syn**] [**+rst** | **-rst**] [**+psh** | **-psh**] [**+ack** | **-ack**] [**+urg** | **-urg**] [**established**]—Specifies that the IP/TCP/UDP ACL rule matches on the TCP flags. – **Ack** – Acknowledgement bit – **Fin** – Finished bit – **Psh** – push bit – **Rst** – reset bit – **Syn** – Synchronize bit – **Urg** – Urgent bit – When "**+**<tcpflagname> "is specified, a match occurs if specified <tcpflagname> flag is set in the TCP header. – When "**-**<tcpflagname> "is specified, a match occurs if specified <tcpflagname> flag is *NOT* set in the TCP header. – When **established** is specified, a match occurs if either the RST or ACK bits are set in the TCP header. – This option is visible only if protocol is **tcp**. • [**icmp-type** *icmp-type* [**icmp-code** *icmp-code*] | **icmp-message** *icmp-message*] —Specifies a match condition for ICMP packets. – When **icmp-type** is specified, IP ACL rule matches on the specified ICMP message type, a number from 0 to 255. – When **icmp-code** is specified, IP ACL rule matches on the specified ICMP message code, a number from 0 to 255. – Specifying **icmp-message** implies both **icmp-type** and **icmp-code** are specified. – **icmp-message** is decoded into corresponding ICMP type and ICMP code within that ICMP type. This option is visible only if the protocol is **icmp**. – IPv4 ICMP message types: echo, echo-reply, host-redirect, mobile-redirect, net-redirect, net-unreachable, redirect, packet-too-big, port-unreachable, source-quench, router-solicitation, router-advertisement, time-exceeded, ttl-exceeded, unreachable.

Command	Purpose
continued	• **igmp-type** *igmp-type*—When **igmp-type** is specified, the IP ACL rule matches on the specified IGMP message type (i.e., a number from 0 to 255). • **fragments**—Specifies the rule matches packets that are non-initial fragments (fragment bit asserted). Not valid for rules that match L4 information such as TCP port number since that information is carried in the initial packet. This keyword is visible only if the protocol is **ip**, **tcp**, or **udp**. • **log**—Specifies that this rule is to be logged. • **time-range** *time-range-name*—Allows imposing time limitation on the ACL rule as defined by the parameter time-range-name. If a time range with the specified name does not exist and the ACL containing this ACL rule is applied to an interface or bound to a VLAN, then the ACL rule is applied immediately. If a time range with specified name exists and the ACL containing this ACL rule is applied to an interface or bound to a VLAN, then the ACL rule is applied when the time-range with specified name becomes active. The ACL rule is removed when the time-range with specified name becomes inactive. • **assign-queue** *queue-id*—Specifies the assign-queue, which is the queue identifier to which packets matching this rule are assigned. • {**mirror** | **redirect**} *unit/slot/ port*—Specifies the mirror or redirect interface which is the unit/slot/port to which packets matching this rule are copied or forwarded, respectively. • **rate-limit** *rate burst-size*—Specifies the allowed rate of traffic as per the configured rate in kbps, and burst-size in kbytes. – **Rate** – the committed rate in kilobits per second – **Burst-size** – the committed burst size in Kilobytes.

Command	Purpose
interface *interface*	(Optional) Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
ip access-group *name direction seqnum*	Bind the specified ACL to an interface. **NOTE:** To apply this ACL to all interfaces, issue the command in Global Configuration mode. • *name* — Access list name. (Range: Valid IP access-list name up to 31 characters in length) • *direction* — Direction of the ACL. (Range: In or out. Default is *in*.) • *seqnum* — Precedence for this interface and direction. A lower sequence number has higher precedence. Range: 1 – 4294967295. Default is 1.
CTRL + Z	Exit to Privileged EXEC mode.
show ip access-lists [*name*]	Display all IPv4 access lists and all of the rules that are defined for the IPv4 ACL. Use the optional *name* parameter to identify a specific IPv4 ACL to display.

Configuring a MAC ACL

Beginning in Privileged EXEC mode, use the following commands to create an MAC ACL, configure rules for the ACL, and bind the ACL to an interface.

Command	Purpose
configure	Enter global configuration mode.
mac access-list extended *name*	Create a named MAC ACL. This command also enters MAC Access List Configuration mode. If a MAC ACL with this name already exists, this command enters the mode to update the existing ACL.

Command	Purpose
[*sequence-number*] **{deny | permit}** {*srcmac srcmacmask* | **any**} {*dstmac dstmacmask* | **any** | **bpdu**} [{*ethertypekey* | *0x0600-0xFFFF}* [**vlan eq** *0-4093*] [**cos** *0-7*] [**secondary-vlan eq** *0-4093*] [**secondary-cos** *0-7*] [**log**] [**time-range** *time-range-name*] [**assign-queue** *queue-id*] [{**mirror** |**redirect**} *interface*]	Specify the rules (match conditions) for the MAC access list. • *sequence-number* — Identifies the order of application of the permit/deny statement. If no sequence number is assigned, permit/deny statements are assigned a sequence number beginning at 1000 and incrementing by 10. Statements are applied in hardware beginning with the lowest sequence number. Sequence numbers are applicable only within an access group; i.e., the ordering applies within the access-group scope. The range for sequence numbers is 1–2147483647. • *srcmac* — Valid source MAC address in format xxxx.xxxx.xxxx. • *srcmacmask* — Valid MAC address bitmask for the source MAC address in format xxxx.xxxx.xxxx. • **any** — Packets sent to or received from any MAC address • *dstmac* — Valid destination MAC address in format xxxx.xxxx.xxxx. • *destmacmask* — Valid MAC address bitmask for the destination MAC address in format xxxx.xxxx.xxxx. • **bpdu** — Bridge protocol data unit • *ethertypekey* — Either a keyword or valid four-digit hexadecimal number. (Range: Supported values are **appletalk**, **arp**, **ibmsna**, **ipv4**, **ipv6**, **ipx**, **mplsmcast**, **mplsucast**, **Netbios**, **novell**, **pppoe**, **rarp**.) • *0x0600-0xFFFF* — Specify custom EtherType value (hexadecimal range 0x0600-0xFFFF) • **vlan eq** — VLAN number. (Range 0–4093) • **cos** — Class of service. (Range 0–7) • **secondary-vlan** — An outer VLAN tag, if present in the frame • **secondary-cos** — The CoS value contained in the outer VLAN tag, if present in the frame.

Command	Purpose
continued	• **log**—Specifies that this rule is to be logged. • **time-range** *time-range-name*—Allows imposing time limitation on the ACL rule as defined by the parameter time-range-name. If a time range with the specified name does not exist and the ACL containing this ACL rule is applied to an interface or bound to a VLAN, then the ACL rule is applied immediately. If a time range with specified name exists and the ACL containing this ACL rule is applied to an interface or bound to a VLAN, then the ACL rule is applied when the time-range with specified name becomes active. The ACL rule is removed when the time-range with specified name becomes inactive. • **assign-queue** *queue-id*—Specifies the assign-queue, which is the queue identifier to which packets matching this rule are assigned. • {**mirror** | **redirect**} *unit/slot/ port*—Specifies the mirror or redirect interface which is the unit/slot/port to which packets matching this rule are copied or forwarded, respectively.
interface *interface*	(Optional) Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.

Command	Purpose
mac access-group *name direction seqnum*	Bind the specified MAC ACL to an interface. **NOTE:** To apply this ACL to all interfaces, issue the command in Global Configuration mode. • *name* — Access list name. (Range: Valid MAC access-list name up to 31 characters in length) • *direction* — Direction of the ACL. (Range: In or out. Default is *in*.) • *seqnum* — Precedence for this interface and direction. A lower sequence number has higher precedence. Range: 1 – 4294967295. Default is 1.
CTRL + Z	Exit to Privileged EXEC mode.
show mac access-lists [*name*]	Display all MAC access lists and all of the rules that are defined for the MAC ACL. Use the optional *name* parameter to identify a specific MAC ACL to display.

Configuring an IPv6 ACL

Beginning in Privileged EXEC mode, use the following commands to create an IPv6 ACL, configure rules for the ACL, and bind the ACL to an interface.

Command	Purpose
configure	Enter global configuration mode.
ipv6 traffic-filter *name*	Create an extended IPv6 ACL. This command also enters IPv6 Access List Configuration mode. If an IPv6 ACL with this name already exists, this command enters the mode to update the existing ACL.

Command	Purpose
[*sequence-number*] {**deny** | **permit**} {**ipv6-protocol** | **number** | **every**} {*source-ipv6-prefix/prefix-length* | **any** | *host source-ipv6-address*} [{**range** {*portkey* | *startport*} {*portkey* | *endport*} | {**eq** | **neq** | **lt** | **gt**} {*portkey* | *0-65535*}] {*destination-ipv6-prefix/prefix-length* | **any** | *host destination-ipv6-address*} [{**range** {*portkey* | *startport*} {*portkey* | *endport*} | {**eq** | **neq** | **lt** | **gt**} {*portkey* | *0-65535*}] [**flag** [**+fin** | **-fin**] [**+syn** | **-syn**] [**+rst** | **-rst**] [**+psh** | **-psh**] [**+ack** | **-ack**] [**+urg** | **-urg**] [**established**]] [**flow-label** *value*] [**icmp-type** *icmp-type* [**icmp-code** *icmp-code*] | **icmp-message** *icmp-message*] [**routing**] [**fragments**] [**dscp** *dscp*]}} [**log**] [**assign-queue** *queue-id*] [{**mirror** | **redirect**} *unit/slot/port*] [**rate-limit** *rate burst-size*]	• *sequence-number* — Identifies the order of application of the permit/deny statement. If no sequence number is assigned, permit/deny statements are assigned a sequence number beginning at 1000 and incrementing by 10. Statements are applied in hardware beginning with the lowest sequence number. Sequence numbers are applicable only within an access group; i.e., the ordering applies within the access-group scope. The range for sequence numbers is 1–2147483647. • {**deny** | **permit**} — Specifies whether the IP ACL rule permits or denies the matching traffic. • {*ipv6-protocol* | *number*| **every**} — Specifies the protocol to match for the IP ACL rule. – IPv4 protocols: **icmpv6**, **ipv6**, **tcp** and **udp** – **every**: Match any protocol (don't care) • *source-ipv6-prefix/prefixlength* | **any** | **host** *src-ipv6-address* — Specifies a source IP address and netmask to match for the IP ACL rule. – For IPv6 ACLs, **any** implies a 0::/128 prefix and a mask of all ones. – Specifying "host X::X" implies a prefix length as "/128" and a mask of 0::/128. • [{**range** {*portkey* | *startport*} {*portkey* | *endport*} | {**eq** | **neq** | **lt** | **gt**} {*portkey* | *0-65535*}] — Specifies the layer-4 destination port match condition for the IP/TCp/UDP ACL rule. A destination port number, which ranges from 0-65535, can be entered, or a *portkey*, which can be one of the following keywords: **bgp**, **domain**, **echo**, **ftp**, **ftp-data**, **http**, **ntp**, **pop2**, **pop3**, **rip**, **smtp**, **snmp**, **telnet**, **tftp**, **telnet**, **time**, **who**, and www. Each of these keywords translates into its equivalent destination port number. – When **range** is specified, IPv6 ACL rule matches only if the layer-4 port number falls within the specified port range. The *startport* and *endport* parameters identify the first and last ports that are part of the port range. They have values from 0 to 65535. The ending port must have a value equal or greater than the starting port. The starting port, ending port, and all ports in between will be part of the layer-4 port range.

Command	Purpose
(Continued)	– When **eq** is specified, IPv6 ACL rule matches only if the layer-4 port number is equal to the specified port number or portkey. – When **lt** is specified, IPv6 ACL rule matches if the layer-4 destination port number is less than the specified port number or portkey. It is equivalent to specifying the range as 0 to <specified port number – 1>. • *destination ipv6 prefix* — IPv6 prefix in IPv6 global address format. • **flow label** *value* — The value to match in the Flow Label field of the IPv6 header (Range 0–1048575). • **dscp** *dscp* — Specifies the TOS for an IPv6 ACL rule depending on a match of DSCP values using the parameter dscp. • **log** — Specifies that this rule is to be logged. • *time-range-name* — Specifies the named time range to associate with the ACL rule. • **assign-queue** *queue-id* — Specifies particular hardware queue for handling traffic that matches the rule. • **mirror** *interface* — Allows the traffic matching this rule to be copied to the specified interface. • **redirect** *interface* — This parameter allows the traffic matching this rule to be forwarded to the specified interface.
interface *interface*	(Optional) Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.

Command	Purpose
ipv6 traffic-filter *name direction* [**sequence** *seq-num*]	Bind the specified IPv6 ACL to an interface. **NOTE:** To apply this ACL to all interfaces, issue the command in Global Configuration mode. • *name* — Access list name. (Range: Valid IPv6 access-list name up to 31 characters in length) • *direction* — Direction of the ACL. (Range: In or out. Default is *in*.) • *seqnum* — Precedence for this interface and direction. A lower sequence number has higher precedence. Range: 1 – 4294967295. Default is 1.
CTRL + Z	Exit to Privileged EXEC mode.
show ipv6 access-lists [*name*]	Display all IPv6 access lists and all of the rules that are defined for the IPv6 ACL. Use the optional *name* parameter to identify a specific IPv6 ACL to display.

Configuring a Time Range

Beginning in Privileged EXEC mode, use the following commands to create a time range and configure time-based entries for the time range.

Command	Purpose
configure	Enter global configuration mode.
time-range *name*	Create a named time range and enter the Time-Range Configuration mode for the range.
absolute {[**start** *time date*] [**end** *time date]*}	Configure a nonrecurring time entry for the named time range. • **start** *time date* — Time and date the ACL rule starts going into effect. The time is expressed in a 24-hour clock, in the form of hours:minutes. For example, 8:00 is 8:00 am and 20:00 is 8:00 pm. The date is expressed in the format day month year. If no start time and date are specified, the configuration statement is in effect immediately. • **end** *time date* — Time and date the ACL rule is no longer in effect.

Command	Purpose
periodic {*days-of-the-week time*} **to** {[*days-of-the-week] time*}	Configure a recurring time entry for the named time range. • *days-of-the-week* —The first occurrence indicates the starting day(s) the ACL goes into effect. The second occurrence is the ending day(s) when the ACL rule is no longer in effect. If the end *days-of-the-week* are the same as the start, they can be omitted This variable can be any single day or combinations of days: **Monday**, **Tuesday**, **Wednesday**, **Thursday**, **Friday**, **Saturday**, **Sunday**. Other possible values are: – **daily** -- Monday through Sunday – **weekdays** -- Monday through Friday – **weekend** -- Saturday and Sunday • *time* — Time the ACL rule starts going into effect (first occurrence) or ends (second occurrence). The time is expressed in a 24-hour clock, in the form of hours:minutes.
CTRL + Z	Exit to Privileged EXEC mode.
show time-range [*name*]	View information about all configured time ranges, including the absolute/periodic time entries that are defined for each time range. Use the *name* variable to view information about the specified time range.

ACL Configuration Examples

This section contains the following examples:

NOTE: None of these ACL rules are applicable to the OOB interface.

Basic Rules

- Inbound rule allowing all packets sequenced after all other rules. It is recommended that the largest possible sequence number be specified with a permit every rule to ensure that it is the last rule processed in the ACL.

  ```
  2147483647 permit every
  ```

  Administrators should be cautious when using the **permit every** rule in an access list, especially when using multiple access lists. All packets match a **permit every** rule and no further processing is done on the packet. This means that a **permit every** match in an access list will skip processing subsequent rules in the current or subsequent access-lists and allow all packets not previously denied by a prior rule.

- Inbound rule to drop all packets:

  As the last rule in a list, this rule is redundant as an implicit “deny every” is added after the end of the last access-group configured on an interface.

  ```
  10000 deny every
  ```

  Administrators should be cautious when using the **deny every** rule in an access list, especially when using multiple access lists. When a packet matches a rule, no further processing is done on the packet. This means that a **deny every** match in an access list will skip processing subsequent rules in the current or subsequent access-lists and drop all packets not previously allowed by a prior rule.

- Inbound rule allowing access FROM hosts with IP addresses ranging from 10.0.46.0 to 10.0.47.254:

  ```
  permit ip 10.0.46.0 0.0.1.255 any
  ```

- Inbound rule allowing access TO hosts with IP addresses ranging from 10.0.48.0 to 10.0.49.254:

  ```
  permit ip any 10.0.48.0 0.0.1.255
  ```

  As the last rule in an administrator-defined list, the narrower scope of this inbound rule has no effect other than to possibly interfere with switch management access or router operations. The system installs an implicit deny every rule after the end of the last access group bound to an interface:

  ```
  500 deny ip any any
  ```

Internal System ACLs

The switch installs a number of internal ACLS to trap packets to the switch CPU for processing. Examples of these types of packets are IEEE 802.1x EAPOL packets, IP source guard packets, LLPF packets, LLDP packets, IEEE 802.1AD packets, etc. These internal ACLs are generally configured at the lowest priority (higher numerically) so that the switch administrator, through the use of ACLs, can override the default switch behavior. An example is an ACL that matches only on the source MAC address. Some of the system rules are installed when the administrator enables specific protocols; other rules are always present and may have their behaviors altered by enabling or disabling protocols, e.g., iSCSI or LLPF. For example, spanning tree BPDUs, LLDP packets, and IEEE 802.1X packets are never forwarded by the switch by default.

Complete ACL Example

The following example is a complete inbound ACL that allows access for hosts connected to gi1/0/1 with IP address in 10.1.1.x range to send IP packets to 192.168.0.X hosts on gi1/0/2. IP packets not from 10.1.1.x addresses or not addressed to 192.168.0.x hosts are dropped. Packets with protocols other than IP, DNS, ARP, or ICMP are dropped. Allowing ICMP supports the 10.1.1.x hosts in reliably receiving and initiating TCP connections and pinging through the switch. This example also allows ARP and DNS packets to any destination and is suitable for a layer-2 switch. Both administrator-specified and automatic sequence numbering of the ACLs is demonstrated.

```
console#config
console(config)#mac access-list extended Allow-ARP
console(config-mac-access-list)#permit any any arp
console(config-mac-access-list)#exit

console(config)#ip access-list Allow-10-1-1-x
console(config-ip-acl)#10 permit ip 10.1.1.0 0.0.0.255
console(config-ip-acl)#20 permit any 192.168.0.0 0.0.0.255
console(config-ip-acl)#30 permit icmp 10.1.1.0 0.0.0.255 any
console(config-ip-acl)#40 permit ip 0.0.0.0 255.255.255.255 any
console(config-ip-acl)#50 permit udp any any eq domain
console(config-ip-acl)#exit

console(config)#interface gi1/0/1
console(config-if-gi1/0/1)#mac access-group Allow-ARP in 10
console(config-if-gi1/0/1)#ip access-group Allow-10-1-1-x in 20
console(config-if-gi1/0/1)#exit
```

Another list on the 192.168.0.x network attached port (gi1/0/2) is configured for this example. Because the two access lists are complementary/end-to-end, it is necessary to allow ICMP packets to travel between the attached hosts. Specific sequence numbering of the ACLs rules is shown here.

```
console(config)#ip access-list Allow-192-168-0-x
console(config-ip-acl)#10 permit ip 192.168.0.0 0.0.0.255 10.1.1.0
0.0.0.255
console(config-ip-acl)#20 permit icmp 192.168.0.0 0.0.0.255 any
console(config-ip-acl)#30 permit udp any any eq domain
console(config-ip-acl)#exit

console(config)#interface gi1/0/2
console(config-if-gi1/0/2)#mac access-group Allow-ARP in 10
console(config-if-gi1/0/2)#ip access-group Allow-192-168-0-x in 20
```

```
console(config-if-gi1/0/2)#exit
```

Consider the following inbound rules that allow Telnet connections and UDP traffic from the 192.168.0.x network to host 10.1.1.23:

```
ip access-list Host10-1-1-23
! Permit Telnet traffic from 192.168.0.X network to host 10.1.1.23:
permit tcp 192.168.0.0 0.0.0.255 host 10.1.1.23 eq telnet
! Permit TCP traffic from 192.168.0.X network to host 10.1.1.23:
permit tcp 192.168.0.0 0.0.0.255 host 10.1.1.23
! Permit UDP traffic from 192.168.0.X network to host 10.1.1.23
permit udp 192.168.0.0 0.0.0.255 host 10.1.1.23
! Permit IP traffic from 192.168.0.X network to 10.1.1.x network
permit ip 192.168.0.0 0.0.0.255 10.1.1.23 0.0.0.255
```

In the above list, the fourth rule allows all IP packets between the network and host. The narrower scope of the first three rules is redundant, as all IP traffic, including TCP and UDP, is permitted by the fourth rule.

The following list has corrected rules that allow Telnet and UDP packets only and rely on the implicit "deny all" after the end of the last access group to deny other traffic.

```
ip access-list Host10-1-1-23
! Permit Telnet traffic from 192.168.0.X network to host 10.1.1.23
permit tcp 192.168.0.0 0.0.0.255 host 10.1.1.23 eq telnet
! Permit UDP traffic from 192.168.0.X network to host 10.1.1.23
permit udp 192.168.0.0 0.0.0.255 host 10.1.1.23
```

The ACL feature supports TCP and UDP port matching using operators:

```
console(config-ip-acl)#permit tcp 10.1.1.0 0.0.0.255 ?

<dstip>              Enter a Destination IP Address.
any                  Match any Destination IP Address.
eq                   Matches only if port number is equal.
gt                   Matches only if port number is greater.
host                 Enter a destination host.
lt                   Matches only if port number is less.
neq                  Matches only if port number is not equal.
range                Specify the range of ports.
```

The range operator is inclusive of the specified port parameters.

ACLs support TCP flags. If multiple flags are set (**+flag**) in a single rule, only packets with the all the same flags asserted are matched (logical AND). Likewise, if multiple flags are cleared (**–flag**) in a single rule, only packets with the same flags cleared are matched. The established keyword matches TCP packets with either the RST or ACK bits set (logical OR). Flags that are neither set nor cleared in the rule are not checked in the ACL (don't care or wildcard).

```
console(config)#ip access-list flags-demo
console(config-ip-acl)#permit tcp any any flag ?

<value>        Enter a TCP Flag (+fin, -fin, +syn, -syn, +rst, -rst,
               +psh, -psh, +ack, -ack, +urg, -urg, established).
               Enter a flag (+|-) only once. Specifying established
               implies specifying either +rst or +ack
established    Match occurs if either RST or ACK bits are set in the
               TCP header (Only for TCP).
```

The following is an example rule to match TCP packets with the PUSH flag asserted AND the RESET flag cleared. The other flags bits are “don't care”:

```
console(config-ip-acl)#permit tcp any any flag -rst +psh
```

ACLs may also contain a number of shorthand qualifiers for protocols and IP, TCP, and UDP port numbers, as shown below. Note that not all of these qualifiers make sense in the context of any given port number; e.g., ftp and ftp-data only make sense in the context of the IP or UDP protocols, while an HTTP port number only makes sense in terms of the TCP or IP protocols. Refer to RFC 1700 or iana.org/protocols for a list of protocol numbers.

```
console(config-ip-acl)#permit ?

<1-255>      Match the protocol number.
eigrp        Match the EIGRP protocol.
every        Match every packet.
gre          Match the GRE protocol.
icmp         Match the ICMP protocol.
igmp         Match the IGMP protocol.
ip           Match the IP protocol.
ipinip       Match the IPINIP protocol.
ospf         Match the OSPF protocol.
tcp          Match the TCP protocol.
udp          Match the UDP protocol.

console(config-ip-acl)#permit tcp 10.1.1.0 0.0.0.255 eq ?

<0-65535>    Enter the layer 4 port number in the range 0 to 65535.
<portkey>    Enter a keyword { domain | echo | ftp | ftp-data |
             http | smtp | snmp | telnet | tftp | www | bgp |
```

```
pop2 | pop3 | ntp | rip | time | who }.
```

To bind an access-list to an interface, use the **access-group** command. The **in** parameter specifies that the ACL is applied to ingress packets. The **out** parameter specifies that the ACL is applied to egress packets not generated by the switch/router. If no **in**/**out** parameter is specified, the access list default is to apply the ACL to ingress packets.

```
console(config)#interface gi1/0/1
console(config-if-Gi1/0/1)#ip access-group Host10-1-1-23 in
```

Multiple access lists can be configured on an interface. The processing order is determined by the last parameter on the access-group command where the lowest sequence number is processed first, followed by the next higher sequence number, etc.

In this example, access list Host10-1-1-21 is processed first, followed by Host-1-1-23:

```
console(config)#interface gi1/0/1
console(config-if-Gi1/0/1)#ip access-group Host10-1-1-23 in 2
console(config-if-Gi1/0/1)#ip access-group Host10-1-1-21 in 1
```

Advanced Examples

Configuring a Time-Based ACL

The following example configures an ACL that denies HTTP traffic from 8:00 pm to 12:00 pm and 1:00 pm to 6:00 pm on weekdays and from 8:30 am to 12:30 pm on weekends. The ACL affects all hosts connected to ports that are members of VLAN 100. The ACL permits VLAN 100 members to browse the Internet only during lunch and after hours.

To configure the switch:

1 Create a time range called work-hours.

```
console#config
console(config)#time-range work-hours
```

2 Configure an entry for the time range that applies to the morning shift Monday through Friday.

```
console(config-time-range)#periodic weekdays 8:00 to 12:00
```

3 Configure an entry for the time range that applies to the afternoon shift Monday through Friday.

```
console(config-time-range)#periodic weekdays 13:00
to 18:00
```

4 Configure an entry for the time range that applies to Saturday and Sunday.

```
console(config-time-range)#periodic weekend 8:30 to 12:30
console(config-time-range)#exit
```

5 Create an ACL named web-limit that denies HTTP traffic during the work-hours time range.

```
console(config)#ip access-list web-limit
console(config-ip-acl)#deny tcp any any eq http time-range
work-hours
console(config-ip-acl)#permit every
```

6 Enter interface configuration mode for VLAN 100 and apply the ACL to ingress traffic.

```
console(config)#interface vlan 100
console(config-if-vlan100)#ip access-group weblimit
in
console(config-if-vlan100)#exit
console(config)#exit
```

7 Verify the configuration.

```
console#show ip access-lists web-limit

IP ACL Name: web-limit


Rule Number: 1000
Action....................................... deny
Match All.................................... FALSE
Protocol..................................... 6(tcp)
Source IP Address............................ any
Destination IP Address....................... any
Destination Layer 4 Operator................. Equal To
Destination L4 Port Keyword.................. 80(www/http)

Rule Number: 1010
Action....................................... permit
Match All.................................... TRUE
```

Denying FTP Traffic

This example filters (drops) ingress FTP setup and data traffic on interfaces gi1/0/24 to 48. This example is suitable for configuration on a switch or a router where it is desirable to eliminate FTP data traffic on certain interfaces:

```
console#config
console(config)#ip access-list deny-ftp
console(config-ip-acl)#deny tcp any any eq ftp
console(config-ip-access-list)#deny tcp any any eq ftp-data
console(config-ip-access-list)#2147483647 permit every
console(config-ip-access-list)#exit

console(config)#interface range gi1/0/24-48
console(config-if)#ip access-list deny-ftp in
console(config-if)#exit
```

Allow FTP Traffic Only to an FTP Server

This ACL limits traffic from a router to a directly connected FTP server (172.16.0.5) on gi1/0/11. Notice that this is an "out" or egress ACL. Traffic to the router from the FTP server is not affected by this rule. Traffic from the router to the FTP server is limited to ICMP and packets destined to the FTP ports. There is no need to add permit rules for all the protocols the router can send to the host (e.g., ARP, ICMP, LLDP, etc.), as internally generated packets are not limited by ACLs. Routing must be enabled to process ARPs or they must be allowed by an explicit rule. We allow ICMP from remote hosts so that the FTP server can receive ICMP feedback from clients utilizing the FTP service. A better implementation would narrow the scope of the ICMP to eliminate ICMP messages not required for the FTP service, e.g., echo, echo-reply, redirect, timestamp, etc.

```
console#config
console(config)#ip access-list allow-ftp-server
console(config-ip-acl)#permit tcp any host 172.16.0.5 eq ftp-data
flag established
console(config-ip-acl)#permit tcp any host 172.16.0.5 eq ftp
console(config-ip-acl)#permit icmp any any
console(config-ip-acl)#exit

console(config)#interface gi1/0/11
console(config-if-gi1/0/11)#ip access-group allow-ftp-server out
console(config-if-gi1/0/11)#exit
```

Block Incoming Pings

This ingress ACL blocks incoming pings on interface Gi1/0/1 directed to hosts reachable from other ports on the switch.

```
console#config
console(config)#ip access-list no-ping
console(config-ip-acl)#deny icmp any any icmp-message echo
console(config-ip-acl)#2147483647 permit every
console(config-ip-acl)#exit

console(config)#interface gi1/0/1
console(config-if-gi1/0/1)#ip access-group no-ping in
console(config-if-gi1/0/1)#exit
```

Block Incoming Pings and Responses

This example configures an ingress ACL that blocks incoming pings and ping responses. Since packets generated by the CPU are not affected by ACLs, to block pinging from the switch we add a rule to block the ping responses on ingress.

```
console#config
console(config)#ip access-list no-ping
console(config-ip-acl)#deny icmp any any icmp-message echo
console(config-ip-acl)#deny icmp any any icmp-message echo-reply
console(config-ip-acl)#2147483647 permit every
console(config-ip-acl)#exit

console(config)#interface gi1/0/1
console(config-if-gi1/0/1)#ip access-group no-ping in
console(config-if-gi1/0/1)#exit
```

Block RFC 1918 Addresses

This ingress ACL may be useful on connections to ISPs to block traffic from non-routable addresses.

```
console#config
console(config)#ip access-list no-private-internet
console(config-ip-acl)#deny ip 10.0.0.0 0.255.255.255 any
console(config-ip-acl)#deny ip 192.168.0.0 0.0.255.255 any
console(config-ip-acl)#deny ip 172.16.0.0 0.15.255.255 any
console(config-ip-acl)#2147483647 permit every
console(config-ip-acl)#exit

console(config)#interface port-channel 1
console(config-if-Po1)#access-group no-private-internet in
console(config-if-Po1)#exit
```

Assign Ingress Packets to a CoS Queue

Assign a range of source or destination TCP ports to CoS queue 3 to provide elevated service. Two rules are necessary to handle packets that have source or destination ports outside the range.

```
console#config
console(config)#ip access-list elevated-cos
console(config-ip-acl)#permit tcp any range 49152 65535 any assign-
queue 3
console(config-ip-acl)#permit tcp any any range 49152 65535 assign-
queue 3
console(config-ip-acl)#2147483647 permit every
console(config-ip-acl)#exit

console(config)#ip access-group elevated-cos in 25
```

Rewrite CoS on Egress (Diffserv)

This Diffserv policy rewrites the CoS value to 4 for all IPv4 packets with a CoS value of 5.

```
console#config
console(config)#class-map match-all rewrite-cos ipv4
console(config-classmap)#match cos 5
console(config-classmap)#exit

console(config)#policy-map rewrite out
console(config-policy-map)#class rewrite-cos
console(config-policy-classmap)#mark cos 4
console(config-policy-classmap)#exit

console(config-policy-map)#exit

console(config)#interface gi1/0/1
console(config-if-gi1/0/1)#service-policy out rewrite
console(config-if-gi1/0/1)#exit
```

Schedule Forwarding of Packets to a Different Port

This ACL layer-2 forwards matching packets to a different port based on a time schedule. This is not equivalent to Policy-Based Forwarding, as the TTL in the packet is not decremented, nor is a new destination MAC address written into the packet. The access-group policy is globally configured on all switch interfaces.

```
console#config
console(config)#time-range work-hours
console(config-time-range)#periodic weekdays 07:30 to 18:00
console(config-time-range)#exit

console(config)#ip access-list redirect-traffic
console(config-ip-acl)#permit ip any 172.16.1.0 255.255.255.0
redirect te1/0/1 time-range work-hours
console(config-ip-acl)#2147483647 permit every
console(config-ip-acl)#exit

console(config)#ip access-group redirect-traffic in 30
```

Rate Limit WWW Traffic (Diffserv)

This ingress ACL creates a Diffserv policy to rate-limit WWW packets. Limit and burst values require tuning for local traffic patterns and link speeds. Compare this to the next example.

```
console#config
console(config)#class-map match-all rate-limit-control ipv4
console(config-classmap)#match protocol tcp
console(config-classmap)#match srcl4port www
console(config-classmap)#exit

console(config)#policy-map rate-limit-policy in
console(config-policy-map)#class rate-limit-control
console(config-policy-classmap)#police-simple 9216 128 conform-
action transmit violate-action drop
console(config-policy-classmap)#exit

console(config-policy-map)#exit

console(config)#interface te1/0/1
console(config-if-te1/0/1)#service-policy in rate-limit-policy
console(config-if-te1/0/1)#exit
```

Rate limit WWW traffic (ACL)

This example creates an ACL to rate-limit WWW traffic ingressing the switch on te1/0/1. Initial and established values require tuning for local traffic patterns and link speeds. Note that this ACL applies to traffic sent to the switch IP address as well as traffic forwarded by the switch (in rule). Permit rules with a rate-limit parameter do not require a following deny rule as matching packets exceeding the rate limit are discarded. Compare this with the example above.

```
console#config
console(config)#ip access-list rate-limit-www
console(config-ip-acl)#permit tcp any any eq www flag established
rate-limit 9216 128
console(config-ip-acl)#permit tcp any any eq www rate-limit 1024 64
console(config-ip-acl)#2147483647 permit every
console(config-ip-acl)#exit

console(config)#interface te1/0/1
console(config-if-te1/0/1)#ip access-group rate-limit-www in
console(config-if-te1/0/1)#exit
```

Rate Limit In-Band Management Traffic

The following is an example of rate limiting in-band management traffic on a layer-2 switch. The first two rules rate limit Telnet and SSH (22) traffic for established connections. The third and fourth rules set specific limits for in-bound Telnet and SSH connection requests (third and fourth rules). Setting the control plane mode on the access group limits the requests to those packets transferred to the CPU and does not affect packets transiting the switching silicon. Likewise, because this is internally an egress ACL, it rate limits packets egressing the silicon to the CPU and does not affect packets that are routed in software due to layer-3 table lookup failures, nor does it affect packets sent to the CPU via the system rules, as they are applied on ingress.

The established connection rate limit parameters are 1024 Kbits/second and a burst of 128 Kbytes. The non-established rate limits are 12 Kbytes/second with a 2 Kbyte burst.

```
console#config
console(config)#ip access-list rate-limit-inband-mgmt
console(config-ip-acl)#permit tcp any any eq telnet flag
established rate-limit 1024 128
```

```
console(config-ip-acl)#permit tcp any any eq 22 flag established
rate-limit 1024 128
console(config-ip-acl)#permit tcp any any eq telnet rate-limit 12 2
console(config-ip-acl)#permit tcp any any eq 22 rate-limit 12 2
console(config-ip-acl)#2147483647 permit every
console(config-ip-acl)#exit
console(config)#ip access-group rate-limit-inband-mgmt control-
plane
```

The following commands block fragmented traffic from being sent to the CPU:

```
console#config
console(config)#ip access-list no-frag-inband-mgmt
console(config-ip-acl)#deny tcp any any fragments
console(config-ip-acl)#deny udp any any fragments
console(config-ip-acl)#deny ip any any fragments
console(config-ip-acl)#2147483647 permit every
console(config-ip-acl)#exit

console(config)#ip access-group no-frag-inband-mgmt control-plane
```

Stop Bonjour (mDNS) Traffic

This example drops all traffic destined to 224.0.0.251 on ingress. Packets destined to the reserved multicast address 224.0.0.x are normally forwarded in hardware. This ACL is prioritized over the system rules as it is applied on ingress. This has the effect of stopping all Bonjour (mDNS) traffic from crossing the switch. If it is desired to allow Bonjour traffic in the network, a rate limiter might be more appropriate.

```
console#config
console(config)#ip access-list deny-mdns
console(config-ip-acl)#deny ip any host 224.0.0.251
console(config-ip-acl)#exit
console(config)#ip access-group deny-mdns control-plane
```

Expedite DSCP(EF) Traffic/Limit Background Traffic

By default (with no CoS or DSCP configuration), packets are assigned to User Priority 1/CoS queue 0 (see the output from **show classofservice trust** and **show classofservice dot1p-mapping**). When incast occurs (multiple ports sending to a single output port at a rate greater than can be accommodated), the switch buffer capacity can be exhausted. When the buffer capacity is exhausted, the switch is unable to perform QoS properly as the decision on whether to expedite a packet is overridden by the availability of a buffer to receive the packet. If no buffer is available, the packet is dropped on ingress.

The following configuration sets the switch to expedite DSCP EF traffic and limits buffering of background traffic in CoS queue 0.

This configuration sets the switch to trust DSCP on ingress, maps DSCP EF to CoS queue 3, and enables WRED on CoS queue 0. Then, green TCP traffic is set to begin random discard at 75% port capacity with a 5% drop probability. Non-TCP traffic is set to tail drop at 100% of port buffer capacity. The other WRED queue parameters (yellow and red traffic) are kept at their default values.

```
console#config
console(config)#classofservice trust ip-dscp
console(config)#classofservice ip-dscp-mapping 46 3
console(config)#cos-queue random-detect 0
console(config)#cos-queue strict 3
console(config)#random-detect queue-parms 0 min-thresh 75 30 20 100
max-thresh 100 90 80 100 drop-prob-scale 5 10 10 100
```

Configure a VLAN ACL

This example configures a MAC ACL to rate-limit matching traffic. The ACL is configured on the VLAN interface, and multiple ports are made members of the VLAN. As the ACL is the only ACL on the interfaces, a **permit any any** clause is included to allow other traffic to be permitted. Subsequent ACL will never be matched due to this clause.

1 Create VLAN 100:

```
console(config)#vlan 100
console(config-vlan99)#exit
```

2 Declare a MAC access list with the matching criteria:

```
console(config)#mac access-list extended vlan100
```

3 Match source MAC 001E.C9XX.XXXX. Rate limit to 100 Kbps with a burst of 32 Kbytes:

```
console(config-mac-access-list)#permit 001E.C900.0000
0000.00FF.FFFF any rate-limit 100 32
```

4 Let everyone else in:

```
console(config-mac-access-list)#permit any any
console(config-mac-access-list)#exit
```

5 Configure the access group on the VLAN:

```
console(config)#interface vlan 100
console(config-if-vlan100)#mac access-group vlan100 in 1000
console(config-if-vlan100)#exit
```

6 Assign the VLAN to interfaces:

```
console(config)#interface Gi1/0/1
console(config-if-gi1/0/1)#switchport access vlan 100
console(config-if-gi1/0/1)#exit

console(config)#interface Gi1/0/2
console(config-if-gi1/0/2)#switchport access vlan 100
console(config-if-gi1/0/2)#exit

console(config)#interface Gi1/0/3
console(config-if-gi1/0/3)#switchport access vlan 100
console(config-if-gi1/0/3)#exit
```

A Consolidated DoS Example

This example includes some ACL rules to consider to reduce DoS attacks on the switch. It does not represent a complete DoS suite. A firewall with deep packet inspection capabilities should be used for true DoS protection.

NOTE: The rate limits below should be adjusted to match the expected rates of traffic coming to the CPU.

1 Configure an IP access list named "squelch-dos attacks":

```
console#config
console(config)#ip access-list squelch-dos-attacks
```

2 Rate-limit echo requests:

```
console(config-ip-acl)#permit icmp any any icmp-message echo
rate-limit 32 64
```

3 Deny telnet and rate-limit SSH to the CPU:

```
console(config-ip-acl)#deny tcp any any eq telnet flag
established
console(config-ip-acl)#permit tcp any any eq 22 flag
established rate-limit 1024 128
console(config-ip-acl)#deny tcp any any eq telnet
console(config-ip-acl)#permit tcp any any eq 22 rate-limit 12 2
```

4 Rate limit TCP opens:

```
console(config-ip-acl)#permit tcp any any flag +syn rate-limit
8 2
```

5 Rate limit TCP closes:

```
console(config-ip-acl)#permit tcp any any flag +fin rate-limit
8 2
```

6 Block TCP/UDP/IP frag attacks:

```
console(config-ip-acl)#deny ip any any fragments
```

7 Limit SNMP (should set source address to management stations). Must be tuned for SNMP walks. May need to adjust the SNMP client retry count or timeout:

```
console(config-ip-acl)#permit udp any any eq snmp rate-limit
1024 128
```

8 Allow other traffic types to come to CPU:

```
console(config-ip-acl)#permit every
console(config-ip-acl)#exit
```

```
console(config)#ip access-group squelch-dos-attacks control-
plane
```

9 Further limit inbound traffic on in-band management ports. Allow only VLAN 99 SSH and TFTP, no telnet, HTTP, HTTPS, or SNMP. The management access list actions are performed by the switch firmware in addition to the access list actions performed by the switching silicon, e.g., squelch-dos-attacks. Note that the switch forces TFTP accesses to use the well-known TFTP port number 69:

```
console(config)#management access-list mgmt-blocks
console(config-ip-acl)#permit vlan 99 service ssh
console(config-ip-acl)#permit vlan 99 service tftp
console(config-ip-acl)#deny vlan 99
console(config-ip-acl)#permit service any
console(config-ip-acl)#exit
```

10 Create an in-band Management VLAN (99), assign it to two ports (gi1/0/47 and gi1/0/48), and add both ACLs and Management ACLs to ALL ports in global config mode.

```
console(config)#vlan 99
console(config-vlan99)#exit
console(config)#interface vlan 99
console(config-if-vlan99)#ip address dhcp
console(config-if-vlan99)#exit
console(config)#interface gi1/0/47-48
console(config-if-Gi1/0/47-48)#switchport access vlan 99
console(config-if-Gi1/0/47-48)#exit
console(config)#management access-class mgmt-blocks
console(config)#line ssh
console(config-ssh)#login authentication default
console(config-ssh)#exit
console(config)#crypto key generate rsa
console(config)#crypto key generate dsa
console(config)#ip ssh server
```

Policy-Based Routing Examples

Route-Map with Scheduled Redirection of RFC 1918 Addresses to a Different Next-Hop

1 Create a time range named "work-hours" the from 7:30 AM to 6:00 PM:

```
console#config
console(config)#time-range work-hours
console(config-time-range)#periodic weekdays 07:30 to 18:00
console(config-time-range)#exit
```

2 Define an IP ACL named "subnet-172-16" and permit all accesses on the subnet during the work-hours time range:

```
console(config)#ip access-list subnet-172-16
console(config-ip-acl)#permit ip any 172.16.0.0 0.15.255.255
time-range work-hours
console(config-ip-acl)#exit
```

3 Define an IP ACL named "subnet-192-168" and permit all accesses on the subnet during the work-hours time range.

```
console(config)#ip access-list subnet-192-168
console(config-ip-acl)#permit ip any 192.168.0.0 0.0.255.255
time-range work-hours
console(config-ip-acl)#exit
```

4 Define an IP ACL named "subnet-10-0" and permit all accesses on the subnet during the work-hours time range.

```
console(config)#ip access-list subnet-10-0
console(config-ip-acl)#permit ip any 10.0.0.0 0.255.255.255
time-range work-hours
console(config-ip-acl)#exit
```

5 Define a route-map named "redirect-vlan12" that permits routes in the three subnets defined earlier. Specify the next hop addresses for all matching routes.

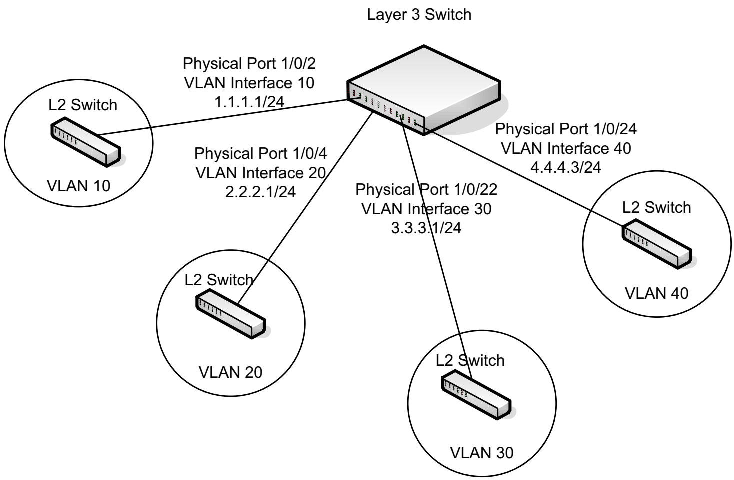
```
console(config)#route-map redirect-vlan12 permit 32
console(config-route-map)#match ip address subnet-172-16
subnet-192-168 subnet-10-0
console(config-route-map)#set ip next-hop 12.1.13.1 12.1.14.1
console(config-route-map)#exit
```

Complete Example of Policy-Based Routing on VLAN Routing Interfaces

In this example, an layer-3 router with four VLAN routing interfaces (VLAN 10, VLAN 20, VLAN 30 and VLAN 40) is configured. Each of these interfaces is connected to layer-2 switches.

Traffic sent to host 2.2.2.2 from host 1.1.1.2 on VLAN interface 10 is normally routed over VLAN interface 20. The steps to override the normal routing decision and policy route traffic from VLAN interface 10 to VLAN interface 30 are described following the figure.

Figure 20-14. Policy-Based Routing on VLAN Interfaces Example



1 Create VLANs 10, 20, 30 and 40

```
console#config
console(config)#vlan 10,20,30,40
console(config-vlan10,20,30,40)#exit
```

2 Add VLAN Membership to Physical Ports. Also, configure the native VLAN on the corresponding interfaces:

```
console(config)#interface gi1/0/2
console(config-if-gi1/0/2)#switchport mode trunk
console(config-if-gi1/0/2)#switchport trunk allowed vlan
remove 1
console(config-if-gi1/0/2)#switchport trunk native vlan 10
```

```
console(config-if-gi1/0/2)#exit

console(config)#interface gi 1/0/4
console(config-if-gi1/0/4)#switchport mode trunk
console(config-if-gi1/0/4)#switchport trunk allowed vlan
remove 1
console(config-if-gi1/0/4)#switchport trunk native vlan 20
console(config-if-gi1/0/4)#exit

console(config)#interface gi1/0/22
console(config-if-gi1/0/22)#switchport mode trunk
console(config-if-gi1/0/22)#switch trunk allowed vlan remove 1
console(config-if-gi1/0/22)#switch trunk native vlan 30
console(config-if-gi1/0/22)#exit

console(config)#interface gi 1/0/24
console(config-if-gi1/0/24)#switchport mode trunk
console(config-if-gi1/0/24)#switchport trunk native vlan 40
console(config-if-gi1/0/24)#switchport trunk allowed vlan
remove 1
console(config-if-gi1/0/24)#exit
```

3 Enable Routing on Each VLAN Interface

```
console(config)#interface vlan 10
console(config-if-vlan10)#ip address 1.1.1.1 255.255.255.0
console(config-if-vlan10)#exit

console(config)#interface vlan 20
console(config-if-vlan20)#ip address 2.2.2.1 255.255.255.0
console(config-if-vlan20)#exit

console(config)#interface vlan 30
console(config-if-vlan30)#ip address 3.3.3.1 255.255.255.0
console(config-if-vlan30)#exit

console(config)#interface vlan 40
console(config-if-vlan40)#ip address 4.4.4.3 255.255.255.0
console(config-if-vlan40)#exit
```

4 Enable IP Routing (Global Configuration)

```
console(config)#ip routing
```

In this configuration, traffic from host 1.1.1.2 to host 2.2.2.2 is routed from VLAN routing interface 10 to VLAN routing interface 20 using the directly connected subnets as they appear in the routing table.

5 Configure Policy Routing. To policy-route such traffic to VLAN routing interface 30, the following additional steps should be performed:

a Create an access-list matching all incoming IP traffic from host 1.1.1.1 destined to host 2.2.2.2:

```
console(config)#ip access-list Match-ip-1_1_1_2-to-2_2_2_2
console(config-ip-acl)#permit ip host 1.1.1.2 host 2.2.2.2
console(config-ip-acl)#exit
```

There is no need to add a **permit every** rule, as would be configured in a normal access list, as this ACL will only be used for PBR. The default for PBR is to route non-matching traffic or traffic which is addressed to a non-connected interface normally.

b Create a route-map and add match/set rules to the route-map:

```
console(config)#route-map Redirect_to_3_3_3_3 permit 100
console(route-map)#match ip address Match-ip-1_1_1_2-to-
2_2_2_2
console(route-map)#set ip next-hop 3.3.3.3
console(route-map)#exit
```

c Assign the route-map to VLAN routing interface 10:

```
console(config)#interface vlan 10
console(config-if-vlan10)#ip policy route-map
Redirect_to_3_3_3_3
console(config-if-vlan10)#exit
```

Traffic matching ACL *Match-ip-1_1_1_2-to-2_2_2_2* is now policy-routed to VLAN interface 30 when an interface in VLAN 30 is connected via policy *Redirect_to_3_3_3_3*. Counters are incremented in the **show route-map** command indicating that traffic is being policy routed.

```
console(config)#show route-map Redirect_to_3_3_3_3

route-map "Redirect_to_3_3_3_3" permit 10
   Match clauses:
      ip address (access-lists) : match-subnet-1_1_1_X
   Set clauses:
      ip next-hop 3.3.3.3
Policy routing matches: 19922869 packets, 1275063872 bytes
```

21

VLANs

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure VLANs, including port-based VLANs, protocol-based VLANs, double-tagged VLANs, subnet-based VLANs, and Voice VLANs.

The topics covered in this chapter include:

- VLAN Overview
- Default VLAN Behavior
- Configuring VLANs (Web)
- Configuring VLANs (CLI)
- VLAN Configuration Examples

VLAN Overview

By default, all ports on Dell Networking N1500, N2000, N3000, and N4000 Series switches are in the same broadcast domain (VLAN 1). This means when any host connected to the switch broadcasts traffic, every other device connected to the switch receives that broadcast. All ports in a broadcast domain also forward multicast and unknown unicast traffic to every directly connected device. Large broadcast domains can result in network congestion, and end users might complain that the network is slow. In addition to latency, large broadcast domains are a greater security risk since all hosts receive all broadcasts.

Virtual Local Area Networks (VLANs) allow the administrator to divide a broadcast domain into smaller, logical networks. Like a bridge, a VLAN switch forwards traffic based on the layer-2 header, which is fast, and like a router, it partitions the network into logical segments, which provides better administration, security, and management of multicast traffic.

Network administrators have many reasons for creating logical divisions, such as department or project membership. Because VLANs enable logical groupings, group members do not need to be physically connected to the same switch or network segment. Some network administrators use VLANs to segregate traffic by type so that the time-sensitive traffic, like voice traffic, has priority over other traffic, such as data. Administrators also use VLANs to protect network resources. Traffic sent by authenticated clients might be assigned to one VLAN, while traffic sent from unauthenticated clients might be assigned to a different VLAN that allows limited network access.

When one host in a VLAN sends a broadcast, the switch forwards traffic only to other members of that VLAN. For traffic to go from a host in one VLAN to a host in a different VLAN, the traffic must be forwarded by a layer-3 device, such as a router. VLANs work across multiple switches and switch stacks, so there is no requirement for the hosts to be located near each other to participate in the same VLAN.

NOTE: Dell Networking N-Series switches support VLAN routing. When you configure VLAN routing, the switch acts as a layer-3 device and can forward traffic between VLANs. For more information, see "What Are VLAN Routing Interfaces? " on page 1141.

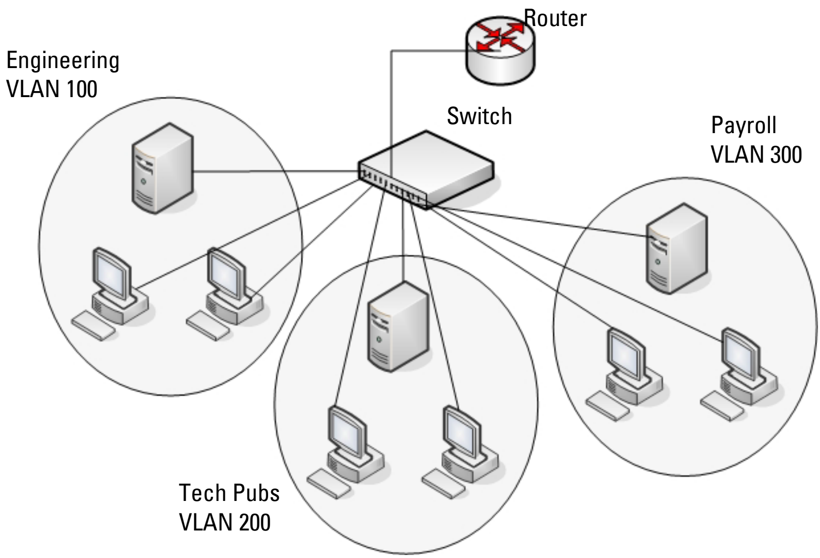
Each VLAN has a unique number, called the VLAN ID. The Dell Networking N-Series switches support a configurable VLAN ID range of 1–4093. A VLAN with VLAN ID 1 is configured on the switch by default. VLAN 1 is named *default*, which cannot be changed. However, names can be associated with any other VLANs that are created.

In a tagged frame, the VLAN is identified by the VLAN ID in the tag. In an untagged frame, the VLAN identifier is the Port VLAN ID (PVID) specified for the port that received the frame. For information about tagged and untagged frames, see "VLAN Tagging " on page 704.

The Dell Networking N-Series switches support adding individual ports and Link Aggregation Groups (LAGs) as VLAN members.

Figure 21-1 shows an example of a network with three VLANs that are department-based. The file server and end stations for the department are all members of the same VLAN.

Figure 21-1.   Simple VLAN Topology



In this example, each port is manually configured so that the end station attached to the port is a member of the VLAN configured for the port. The VLAN membership for this network is port-based or static.

Dell Networking N-Series switches also support VLAN assignment based on any of the following criteria:

- MAC address of the end station
- IP subnet of the end station
- Protocol of the packet transmitted by the end station

Table 21-1 provides an overview of the types of VLANs that can be used to logically divide the network.

Table 21-1. VLAN Assignment

VLAN Assignment	Description
Port-based (Static)	This is the most common way to assign hosts to VLANs. The port where the traffic enters the switch determines the VLAN membership.
IP Subnet	Hosts are assigned to a VLAN based on their IP address. All hosts in the same subnet are members of the same VLAN.
MAC-Based	The MAC address of the device determines the VLAN assignment. This type of VLAN is useful when a host might not always connect to the network through the same port but needs to be on the same VLAN.
Protocol	Protocol-based VLANs were developed to separate traffic based on the layer-2 Ethertype before IP traffic became the de facto standard in the LAN. Use a protocol-based VLAN on networks where you might have a group of hosts that use IPX or another legacy protocol. With protocol-based VLANs, traffic can be segregated based on the EtherType value in the frame.

VLAN Tagging

Dell Networking N-Series switches support IEEE 802.1Q tagging. Ethernet frames on a tagged VLAN have a 4-byte VLAN tag in the header. VLAN tagging is required when a VLAN spans multiple switches, which is why trunk ports transmit and receive only tagged frames.

NOTE: A stack of switches behaves as a single switch, so VLAN tagging is not required for packets traversing different stack members.

Tagging may be required when a single port supports multiple devices that are members of different VLANs. For example, a single port might be connected to an IP phone, a PC, and a printer (the PC and printer are connected via ports on the IP phone). IP phones are typically configured to use a tagged VLAN for voice traffic, while the PC and printers typically use the untagged VLAN.

Trunk ports can receive tagged and untagged traffic. Untagged traffic is tagged internally with the native VLAN. Native VLAN traffic received untagged is transmitted untagged on a trunk port.

By default, trunk ports are members of all existing VLANs and will automatically participate in any newly created VLANs. The administrator can restrict the VLAN membership of a trunk port. VLAN membership for tagged frames received on a trunk port is configured separately from the membership of the native VLAN. To configure a trunk port to accept frames only for a single VLAN, both the native VLAN and the tagged VLAN membership settings must be configured. If the native VLAN for a trunk port is deleted, the trunk port drops untagged packets.

Access ports accept untagged traffic and traffic tagged with the access port PVID. Untagged ingress traffic is considered to belong to the VLAN identified by the PVID. If the PVID for an access port is deleted, the PVID is set to VLAN 1.

GVRP

The GARP VLAN Registration Protocol (GVRP) helps to dynamically manage VLAN memberships on trunk ports. When GARP is enabled, switches can dynamically register (and de-register) VLAN membership information with other switches attached to the same segment.

Information about the active VLANs is propagated across all networking switches in the bridged LAN that support GVRP. Ports can be configured to forbid dynamic VLAN assignment through GVRP.

The operation of GVRP relies upon the services provided by the Generic Attribute Registration Protocol (GARP). GVRP can create up to 1024 VLANs. For information about GARP timers, see "What Are GARP and GMRP? " on page 874.

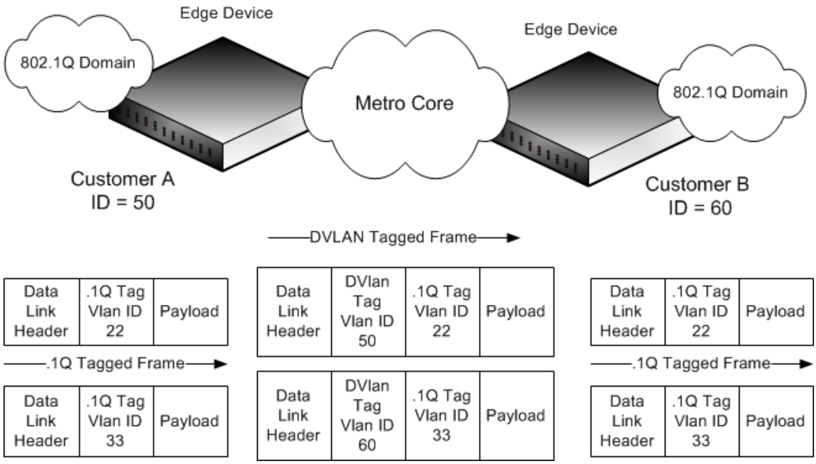
Double-VLAN Tagging

For trunk ports, which are ports that connect one switch to another switch, the Dell Networking N-Series switches support double-VLAN tagging as an option. This feature allows service providers to connect to Virtual Metropolitan Area Networks (VMANs). With double-VLAN tagging, service providers can pass VLAN traffic from one customer domain to another through a metro core in a simple and cost-effective manner. By using an additional tag on the traffic, the switch can differentiate between customers in the MAN while preserving an individual customer's VLAN identification when the traffic enters the customer's 802.1Q domain.

With the introduction of this second tag, customers are no longer required to divide the 4-byte VLAN ID space to send traffic on a Ethernet-based MAN. In short, every frame that is transmitted from an interface has a double-VLAN tag attached, while every packet that is received from an interface has a tag removed (if one or more tags are present).

In Figure 21-2, two customers share the same metro core. The service provider assigns each customer a unique ID so that the provider can distinguish between the two customers and apply different rules to each. When the configurable EtherType is assigned to something different than the 802.1Q (0x8100) EtherType, it allows the traffic to have added security from misconfiguration while exiting the metro core. For example, if the edge device on the other side of the metro core is not stripping the second tag, the packet would never be classified as a 802.1Q tag, so the packet would be dropped rather than forwarded in the incorrect VLAN.

Figure 21-2. Double VLAN Tagging Network Example



Voice VLAN

The Voice VLAN feature enables switch ports to carry voice traffic from IP phones with an administrator-defined priority. When multiple devices, such as a PC and an IP phone, are connected to the same port, the port can be configured to use one VLAN for voice traffic and another VLAN for data traffic. Multiple IP phones per port are supported.

Voice over IP (VoIP) traffic is inherently time-sensitive: for a network to provide acceptable service, low latency is vital. The priority level enables the separation of voice and data traffic coming onto the port and can provide expedited forwarding of voice VLAN traffic.

A primary benefit of using Voice VLAN is to ensure that the sound quality of an IP phone is safeguarded from deteriorating when the data traffic on the port is high. The switch uses the source MAC address of the traffic traveling through the port to identify the IP phone data flow.

The Voice VLAN feature can be enabled on a per-port basis. This feature supports a configurable voice VLAN DSCP or IEEE 802.1p value. This value is later transmitted by LLDP when the LLDPDU is transmitted, if LLDP has been enabled on the port, the DSCP/802.1p value is configured, and the required TLV is configured for the port.

NOTE: Voice VLAN must be configured on general mode ports. It is not supported on access mode or trunk mode ports.

Identifying Voice Traffic

Some VoIP phones contain full support for IEEE 802.1X. When these phones are connected to a port that uses 802.1X port-based authentication, these phones authenticate and receive their VLAN information from LLDP-MED. However, if a VoIP phone has limited support for 802.1X authentication it might try to authenticate and fail. A phone with no 802.1X support would not attempt to authenticate at all. Instead of placing these phones on an unauthenticated or guest VLAN, the switch can automatically direct the VoIP traffic to the Voice VLAN without manual configuration.

The switch identifies the device as a VoIP phone by one of the following protocols:

- Cisco Discovery Protocol (CDP) or Industry Standard Discovery Protocol (ISDP) for Cisco VoIP phones
- DHCP vendor-specific options for Avaya VoIP phones
- LLDP-MED for most VoIP phones

NOTE: By default, ISDP is enabled globally and per-interface on the switch. LLDP-MED is disabled on each interface by default. Port-based authentication using 802.1X is also disabled on each port by default.

After the VoIP phone receives its VLAN information, all traffic is tagged with the VLAN ID of the Voice VLAN. The phone is considered to be authorized to send traffic but not necessarily authenticated.

Segregating Traffic with the Voice VLAN

The switch can be configured to support Voice VLAN on a port that is connected to the VoIP phone. Both of the following methods segregate the voice traffic and the data traffic in order to provide better service to the voice traffic.

- When a VLAN is associated with a Voice VLAN port instead of an 802.1p priority, then the VLAN ID information is passed onto the VoIP phone using either the LLDP-MED or the CDP protocol, depending on how the phone is identified: if it is identified via CDP, then the VLAN assignment is via CDP and if it is identified via LLDP-MED, then the VLAN assignment is via LLDP-MED. In either case, the voice data coming from the VoIP phone is tagged with the exchanged VLAN ID. Untagged data arriving on the switch is given the default PVID of the port. As a result, both kinds of traffic may be segregated by operator configuration in order to provide better service to the voice traffic. Traffic on the Voice VLAN can be assigned to a specific CoS queue or otherwise given priority using the normal policy configuration mechanisms.

- When an 802.1p priority is associated with a Voice VLAN port instead of, or in conjunction with, a VLAN ID, then the priority information is passed onto the VoIP phone using the LLDP-MED or CDP protocol. With this method, the voice data coming from the VoIP phone is tagged with VLAN 0 (or the configured VLAN) and with the configured priority; regular data arriving on the switch is given the default priority of the port, and the voice traffic is received with an operator-configured priority.

By default, the switch is configured to trust the 802.1p priority for traffic received from any device, including IP phones. The voice traffic is identified by the IP phone's MAC address when Voice VLAN is configured. This helps to ensure that the Voice VLAN traffic is resilient in the presence of other devices on the network such as a desktop computer.

Voice VLAN is incompatible with Auto-VoIP. Disable Auto-VoIP before enabling Voice VLAN.

Voice VLAN and LLDP-MED

The interactions with LLDP-MED are important for Voice VLAN:

- LLDP-MED notifies the Voice VLAN component of the presence and absence of a VoIP phone on the network.
- The Voice VLAN component interacts with LLDP-MED for applying VLAN ID, priority, and tag information to the VoIP phone traffic.

Private VLANs

Private VLANs partition a standard VLAN domain into two or more subdomains. Each subdomain is defined by a primary VLAN and a secondary VLAN. The primary VLAN ID is the same for all subdomains that belong to a particular private VLAN instance. The secondary VLAN ID differentiates the subdomains from each other and provides layer-2 isolation between ports on the same private VLAN.

The following types of VLANs can be configured in a private VLAN:

- **Primary VLAN**—Forwards the traffic from the promiscuous ports to isolated ports, community ports and other promiscuous ports in the same private VLAN. Only one primary VLAN can be configured per private VLAN. All ports within a private VLAN share the same primary VLAN.
- **Isolated VLAN**—A secondary VLAN. It carries traffic from isolated ports to promiscuous ports. Only one isolated VLAN can be configured per private VLAN.
- **Community VLAN**—A secondary VLAN. It forwards traffic between ports which belong to the same community and to the promiscuous ports. There can be multiple community VLANs per private VLAN.

A port may be designated as one of the following types in a private VLAN:

- **Promiscuous port**—A port associated with a primary VLAN that is able to communicate with all interfaces in the private VLAN, including other promiscuous ports, community ports and isolated ports.
- **Host port**—A port associated with a secondary VLAN that can either communicate with the promiscuous ports in the VLAN and with other ports in the same community (if the secondary VLAN is a community VLAN) or can communicate only with the promiscuous ports (if the secondary VLAN is an isolated VLAN).

Private VLANs may be configured across a stack and on physical and port-channel interfaces.

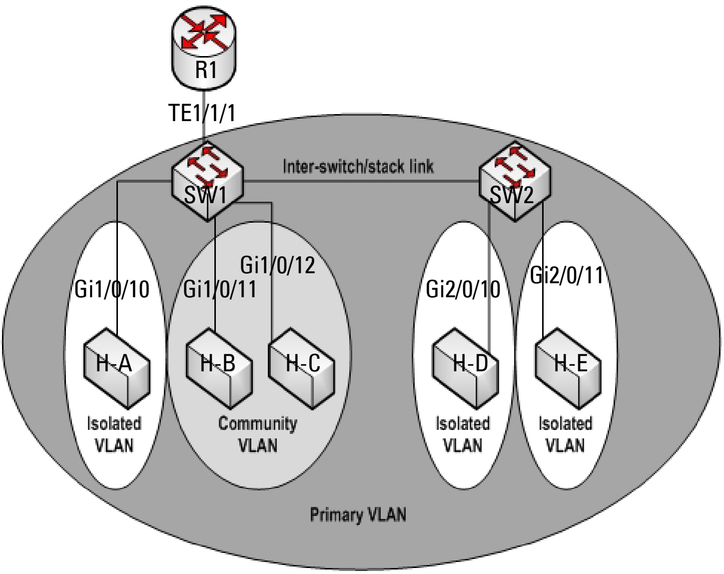
Private VLAN Usage Scenarios

Private VLANs are typically implemented in a DMZ for security reasons. Servers in a DMZ are generally not allowed to communicate with each other but they must communicate to a router, through which they are connected to the users. Such servers are connected to host ports, and the routers are attached to promiscuous ports. Then, if one of the servers is compromised, the intruder cannot use it to attack another server in the same network segment.

The same traffic isolation can be achieved by assigning each port with a different VLAN, allocating an IP subnet for each VLAN, and enabling layer-3 routing between them. In a private VLAN domain, on the other hand, all members can share the common address space of a single subnet, which is associated with a primary VLAN. So, the advantage of the private VLANs feature is that it reduces the number of consumed VLANs, improves IP addressing space utilization, and helps to avoid layer-3 routing.

Figure 21-3 shows an example Private VLAN scenario, in which five hosts (H-A through H-E) are connected to a stack of switches (SW1, SW2). The switch stack is connected to router R1. Port references shown are with reference to the stack.

Figure 21-3. Private VLAN Domain



Promiscuous Ports

An endpoint connected to a promiscuous port is allowed to communicate with any endpoint within the private VLAN. Multiple promiscuous ports can be defined for a single private VLAN domain.

In the configuration shown in Figure 21-3, the port connected from SW1 to R1 (TE1/1/1) is configured as a promiscuous port. It is possible to configure a port-channel as a promiscuous port in order to provide a level of redundancy on the private VLAN uplink.

Isolated Ports

An endpoint connected to an isolated port is allowed to communicate with endpoints connected to promiscuous ports only. Endpoints connected to adjacent isolated ports cannot communicate with each other.

Community Ports

An endpoint connected to a community port is allowed to communicate with the endpoints within a community and can also communicate with any configured promiscuous port. The endpoints that belong to one community cannot communicate with endpoints that belong to a different community, or with endpoints connected to isolated ports.

Private VLAN Operation in the Switch Stack and Inter-switch Environment

The Private VLAN feature is supported in a stacked switch environment. The stack links are transparent to the configured VLANs; thus, there is no need for special private VLAN configuration beyond what would be configured for a single switch. Any private VLAN port can reside on any stack member.

To enable private VLAN operation across multiple switches that are not stacked, trunk ports must be configured between the switches to transport the private VLANs. The trunk ports must be configured with the promiscuous, isolated, and community VLANs. Trunk ports must also be configured on all devices separating the switches.

In regular VLANs, ports in the same VLAN switch traffic at layer 2. However, for a private VLAN, the promiscuous port forwards received traffic to secondary ports in the VLAN (isolated and community). Community ports forward received traffic to the promiscuous ports and other community ports using the same secondary VLAN. Isolated ports transmit received traffic to the promiscuous ports only.

The ports to which the broadcast traffic is forwarded depend on the type of port on which the traffic was received. If the received port is a host port, traffic is broadcast to all promiscuous and trunk ports. If the received port is a community port, the broadcast traffic is forwarded to all promiscuous, trunk, and community ports in the same secondary VLAN. A promiscuous port broadcasts traffic to other promiscuous ports, isolated ports, and community ports.

Table 21-2. Forwarding Rules for Traffic in Primary VLAN

	To				
From	**promiscuous**	**community 1**	**community 2**	**isolated**	**stack (trunk)**
promiscuous	allow	allow	allow	allow	allow
community 1	N/A	N/A	N/A	N/A	N/A
community 2	N/A	N/A	N/A	N/A	N/A
isolated	N/A	N/A	N/A	N/A	N/A
stack (trunk)	allow	allow	allow	allow	allow

Table 21-3. Forwarding Rules for Traffic in Community 1 VLAN

	To				
From	**promiscuous**	**community 1**	**community 2**	**isolated**	**stack (trunk)**
promiscuous	N/A	N/A	N/A	N/A	N/A
community 1	allow	allow	deny	deny	allow
community 2	N/A	N/A	N/A	N/A	N/A
isolated	N/A	N/A	N/A	N/A	N/A
stack (trunk)	allow	allow	deny	deny	allow

Table 21-4. Forwarding Rules for Traffic in Isolated VLAN

	To				
From	**promiscuous**	**community 1**	**community 2**	**isolated**	**stack (trunk)**
promiscuous	N/A	N/A	N/A	N/A	N/A
community 1	N/A	N/A	N/A	N/A	N/A
community 2	N/A	N/A	N/A	N/A	N/A

From	To				
	promiscuous	community 1	community 2	isolated	stack (trunk)
isolated	allow	deny	deny	deny	allow
stack (trunk)	allow	deny	deny	deny	Allow

Limitations and Recommendations

- Only a single isolated VLAN can be associated with a primary VLAN. Multiple community VLANs can be associated with a primary VLAN.
- Trunk and general modes are not supported on private VLAN ports.
- Do not configure access ports using the VLANs participating in any of the private VLANs.
- Multiple primary VLANs may be configured. Each primary VLAN must be unique and each defines a separate private VLAN domain. The operator must take care to use only the secondary VLANs associated with the primary VLAN of a domain.
- Private VLANs cannot be enabled on a pre-configured interface. The interface must physically exist in the switch.
- Secondary (community and isolated) VLANS are associated to the same multiple spanning tree instance as the primary VLAN.
- GVRP/MVRP cannot be enabled after the private VLAN is configured. The administrator will need to disable both before configuring the private VLAN.
- DHCP snooping can be configured on the primary VLAN. If it is enabled for a secondary VLAN, the configuration does not take effect if a primary VLAN is already configured.
- If IP source guard is enabled on private VLAN ports, then DHCP snooping must be enabled on the primary VLAN.
- Do not configure private VLAN ports on interfaces configured for voice VLAN.
- If static MAC addresses are added for the host port, the same static MAC address entry must be added to the associated primary VLAN. This does not need to be replicated for dynamic MAC addresses.
- A private VLAN cannot be enabled on a management VLAN.

- A private VLAN cannot be enabled on the default VLAN.
- VLAN routing can be enabled on private VLANs. It is not very useful to enable routing on secondary VLANs, as the access to them is restricted. However, primary VLANs can be enabled for routing.
- It is recommended that the private VLAN IDs be removed from the trunk ports connected to devices that do not participate in the private VLAN traffic.

Private VLAN Configuration Example

See "Configuring a Private VLAN " on page 763.

Additional VLAN Features

The Dell Networking N-Series switches also support the following VLANs and VLAN-related features:

- VLAN routing interfaces — See "Routing Interfaces " on page 1141.
- Guest VLAN — See "Port and System Security " on page 623.

Default VLAN Behavior

One VLAN is configured on the Dell Networking N-Series switches by default. The VLAN ID is 1, and all ports are included in the VLAN as access ports, which are untagged. This means when a device connects to any port on the switch, the port forwards the packets without inserting a VLAN tag. If a device sends a tagged frame to a port with a VLAN ID other than 1, the frame is dropped. Since all ports are members of this VLAN, all ports are in the same broadcast domain and receive all broadcast and multicast traffic received on any port.

When a new VLAN is created, all trunk ports are members of the VLAN by default. The configurable VLAN range is 2–4093. VLANs 4094 and 4095 are reserved for internal system use.

Ports in trunk and access mode have the default behavior shown in Table 18-2 and cannot be configured with different tagging or ingress filtering values. When adding a VLAN to a port in general mode, the VLAN has the behavior shown in Table 21-5.

Table 21-5. General Mode Default Settings

Feature	Default Value
Frames accepted	Untagged Incoming untagged frames are classified into the VLAN whose VLAN ID is the currently configured PVID.
Frames sent	Untagged
Ingress Filtering	On
PVID	1

Table 21-6 shows the default values or maximum values for VLAN features.

Table 21-6. Additional VLAN Default and Maximum Values

Feature	Value
Default VLAN	VLAN 1
VLAN Name	No VLAN name is configured except for VLAN 1, whose name "default" cannot be changed.
VLAN Range	2–4093
Switchport mode	Access
Double-VLAN tagging	Disabled If double-VLAN tagging is enabled, the default EtherType value is 802.1Q
Maximum number of configurable MAC-to-VLAN bindings	128
Maximum number of configurable IP Subnet-to-VLAN bindings	64
GVRP	Disabled If GVRP is enabled, the default per-port parameters are: • GVRP State: Disabled • Dynamic VLAN Creation: Enabled • GVRP Registration: Enabled
Number of dynamic VLANs that can be assigned through GVRP	1024
Voice VLAN	Disabled
Voice VLAN DSCP value	46
Voice VLAN authentication mode	Enabled

Configuring VLANs (Web)

This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring VLANs on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

VLAN Membership

Use the **VLAN Membership** page to create VLANs and define VLAN groups stored in the VLAN membership table.

To display the **VLAN Membership** page, click **Switching** → **VLAN** → **VLAN Membership** in the navigation panel.

The **VLAN Membership** tables display which Ports and LAGs are members of the VLAN, and whether they're tagged (T), untagged (U), or forbidden (F). The tables have two rows: **Static** and **Current**. Only the **Static** row is configurable. The **Current** row is updated either dynamically through GVRP or when the **Static** row is changed and **Apply** is clicked.

There are two tables on the page:

- **Ports** — Displays and assigns VLAN membership to ports. To assign membership, click in **Static** for a specific port. Each click toggles between U, T, and blank. See Table 21-7 for definitions.
- **LAGs** — Displays and assigns VLAN membership to LAGs. To assign membership, click in **Static** for a specific LAG. Each click toggles between U, T, and blank. See Table 21-7 for definitions.

Table 21-7. VLAN Port Membership Definitions

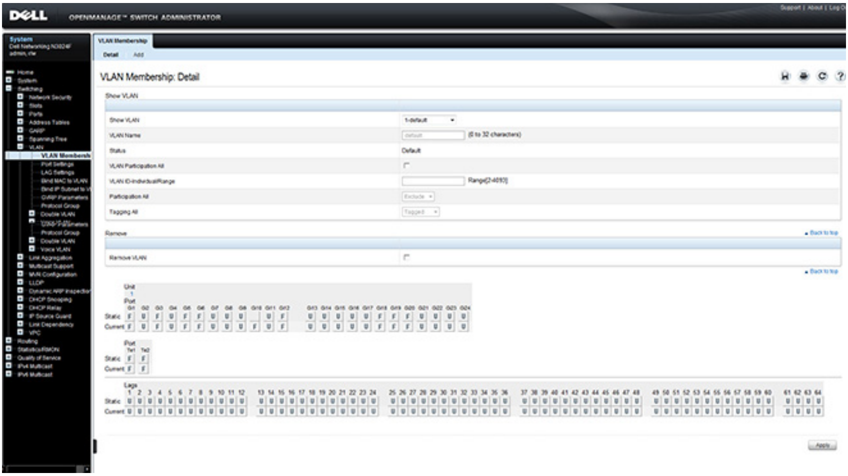
Port Control	Definition
T	Tagged: the interface is a member of a VLAN. All packets forwarded by the interface in this VLAN are tagged. The packets contain VLAN information.
U	Untagged: the interface is a VLAN member. Packets forwarded by the interface in this VLAN are untagged.
F	Forbidden: indicates that the interface is forbidden from becoming a member of the VLAN. This setting is primarily for GVRP, which enables dynamic VLAN assignment.

Table 21-7. VLAN Port Membership Definitions

Port Control	Definition
Blank	Blank: the interface is not a VLAN member. Packets in this VLAN are not forwarded on this interface.

To perform additional port configuration, such as making the port a trunk port, use the **Port Settings** page.

Figure 21-4. VLAN Membership

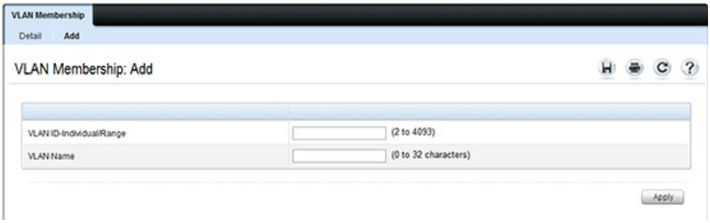


Adding a VLAN

To create a VLAN:

1. Open the **VLAN Membership** page.
2. Click **Add** to display the **Add VLAN** page.
3. Specify a VLAN ID and a VLAN name.

Figure 21-5. Add VLAN



4 Click **Apply**.

Configuring Ports as VLAN Members

To add member ports to a VLAN:

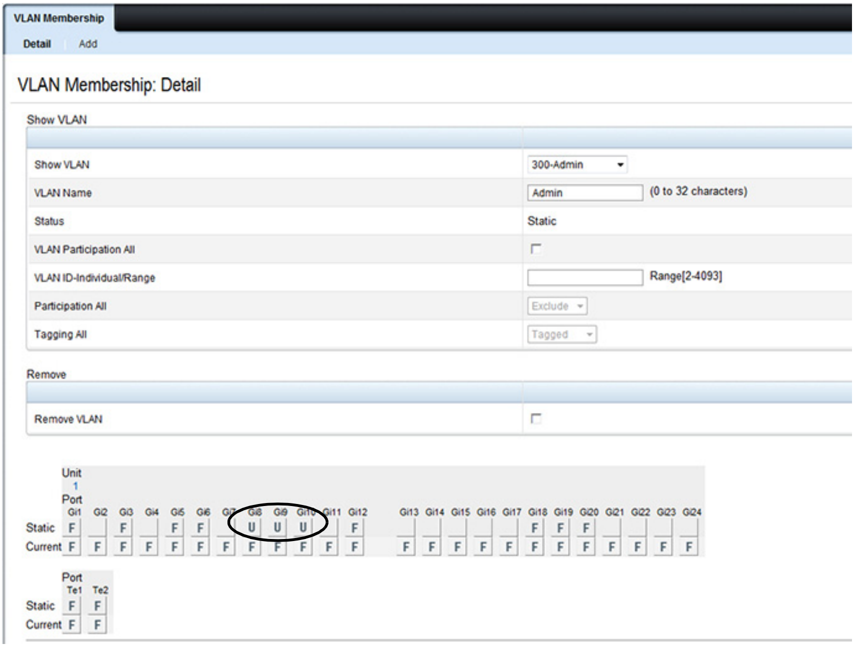
1 Open the **VLAN Membership** page.

2 From the **Show VLAN** menu, select the VLAN to which you want to assign ports.

3 In the **Static** row of the **VLAN Membership** table, click the blank field to assign the port as an untagged member.

Figure 21-6 shows Gigabit Ethernet ports 8–10 being added to VLAN 300.

Figure 21-6. Add Ports to VLAN

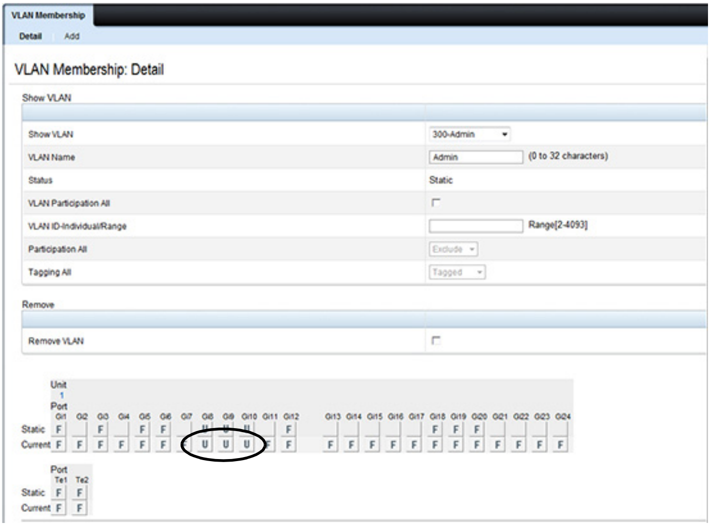


4 Click **Apply**.

5 Verify that the ports have been added to the VLAN.

In Figure 21-7, the presence of the letter **U** in the **Current** row indicates that the port is an untagged member of the VLAN.

Figure 21-7. Add Ports to VLAN



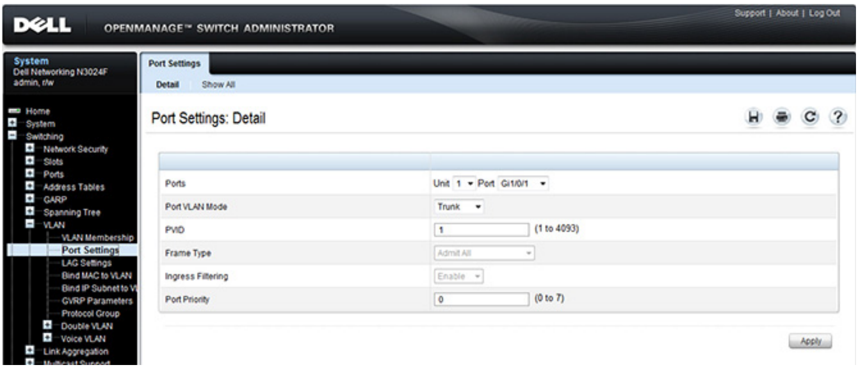
VLAN Port Settings

Use the **VLAN Port Settings** page to add ports to an existing VLAN and to configure settings for the port. If you select Trunk or Access as the **Port VLAN Mode**, some of the fields are not configurable because of the requirements for that mode.

NOTE: Ports can be added to a VLAN through the table on the **VLAN Membership** page or through the **PVID** field on the **Port Settings** page. The PVID is the VLAN that untagged received packets are assigned to. To include a general-mode port in multiple VLANs, use the **VLAN Membership** page.

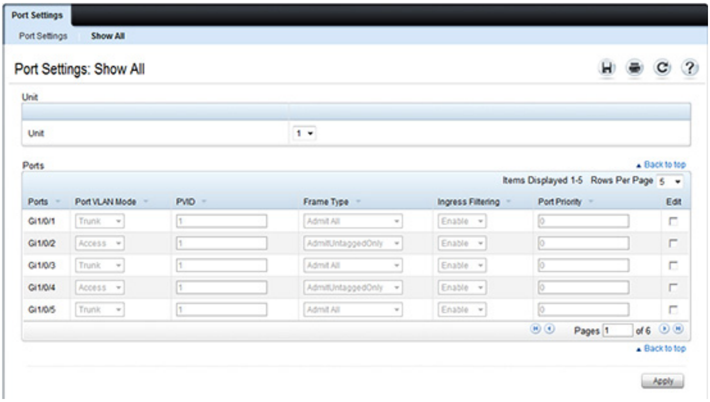
To display the **Port Settings** page, click **Switching** → **VLAN** → **Port Settings** in the navigation panel.

Figure 21-8. VLAN Port Settings



From the **Port Settings** page, click **Show All** to see the current VLAN settings for all ports. To change the settings for one or more ports, click the **Edit** option for a port and select or enter new values.

Figure 21-9. VLAN Settings for All Ports

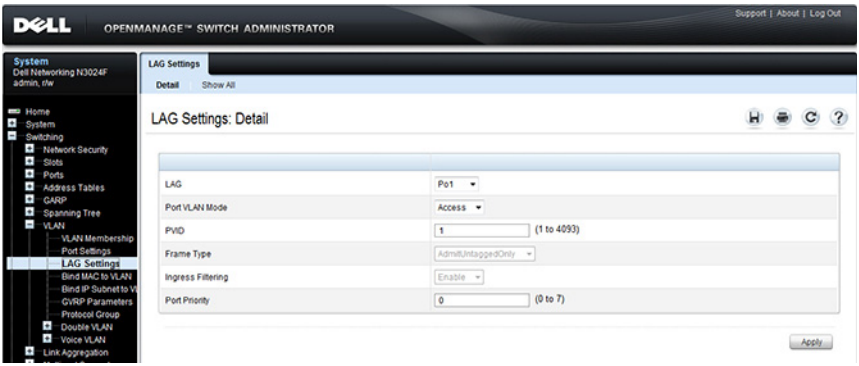


VLAN LAG Settings

Use the **VLAN LAG Settings** page to map a LAG to a VLAN and to configure specific VLAN settings for the LAG.

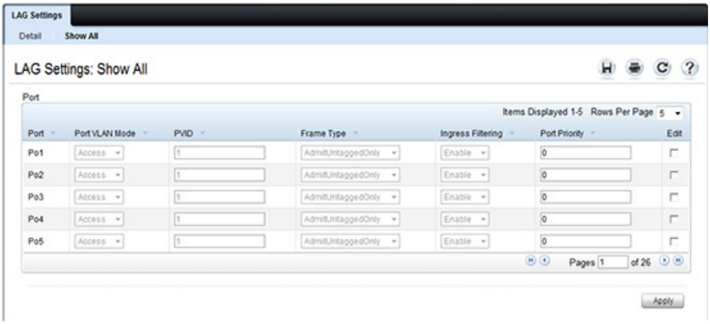
To display the **LAG Settings** page, click **Switching** → **VLAN** → **LAG Settings** in the navigation panel.

Figure 21-10. VLAN LAG Settings



From the **LAG Settings** page, click **Show All** to see the current VLAN settings for all LAGs. To change the settings for one or more LAGs, click the **Edit** option for a port and select or enter new values.

Figure 21-11. VLAN LAG Table

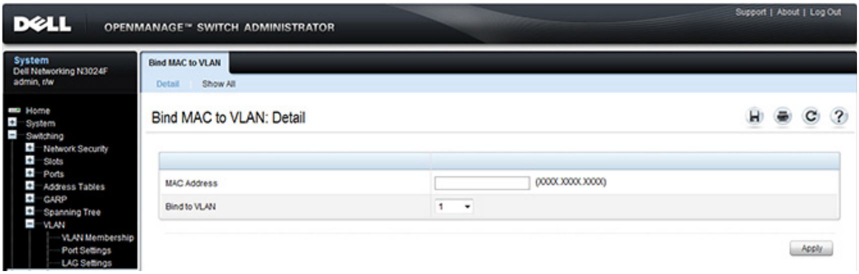


Bind MAC to VLAN

Use the **Bind MAC to VLAN** page to map a MAC address to a VLAN. After the source MAC address and the VLAN ID are specified, the MAC to VLAN configurations are shared across all ports of the switch. The MAC to VLAN table supports up to 128 entries.

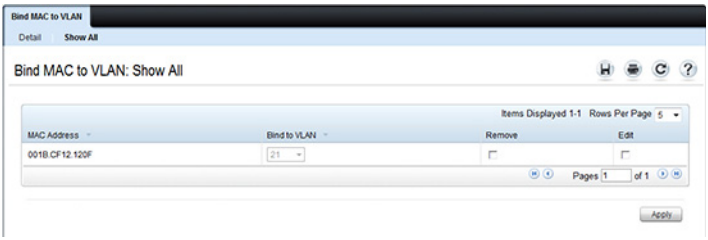
To display the **Bind MAC to VLAN** page, click **Switching** → **VLAN** → **Bind MAC to VLAN** in the navigation panel.

Figure 21-12. Bind MAC to VLAN



From the **Bind MAC to VLAN** page, click **Show All** to see the MAC addresses that are mapped to VLANs. From this page, settings can be changed for one or more entries or entries can be removed.

Figure 21-13. MAC-VLAN Bind Table

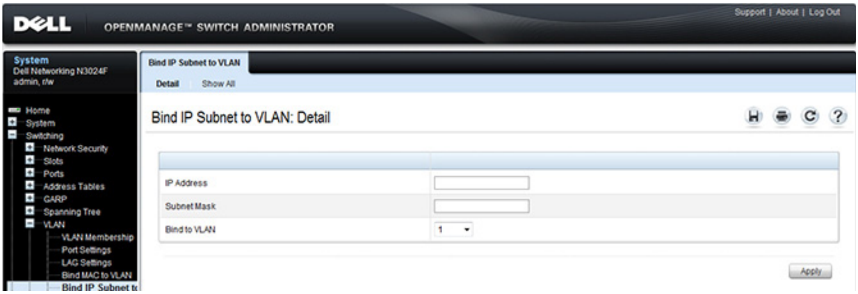


Bind IP Subnet to VLAN

Use the **Bind IP Subnet to VLAN** page to assign an IP Subnet to a VLAN. The IP Subnet to VLAN configurations are shared across all ports of the switch. There can be up to 128 entries configured in this table.

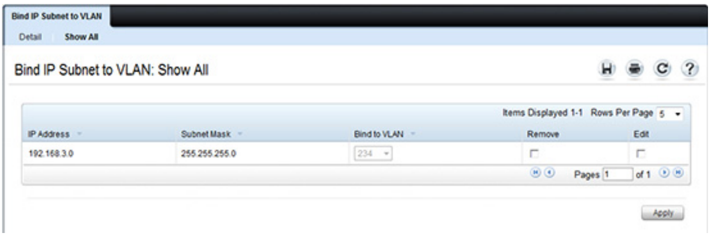
To display the **Bind IP Subnet to VLAN** page, click **Switching** → **VLAN** → **Bind IP Subnet to VLAN** in the navigation panel.

Figure 21-14. Bind IP Subnet to VLAN



From the **Bind IP Subnet to VLAN** page, click **Show All** to see the IP subnets that are mapped to VLANs. From this page, settings can be changed for one or more entries or entries can be removed.

Figure 21-15. Subnet-VLAN Bind Table

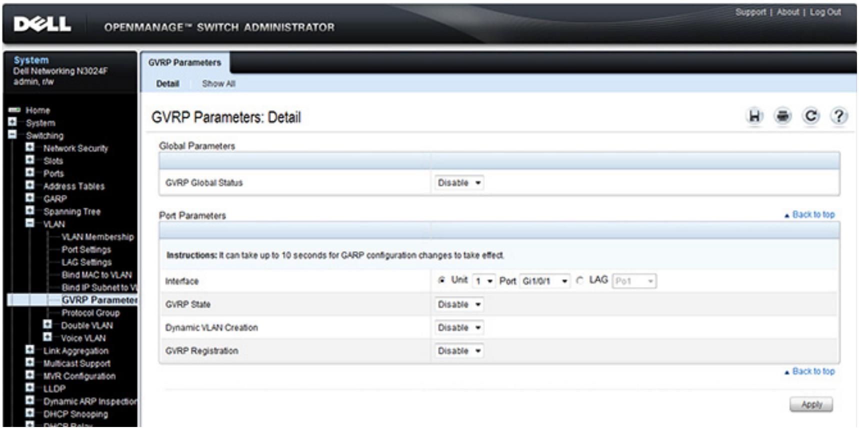


GVRP Parameters

Use the **GVRP Parameters** page to enable GVRP globally and configure the port settings.

To display the **GVRP Parameters** page, click **Switching** → **VLAN** → **GVRP Parameters** in the navigation panel.

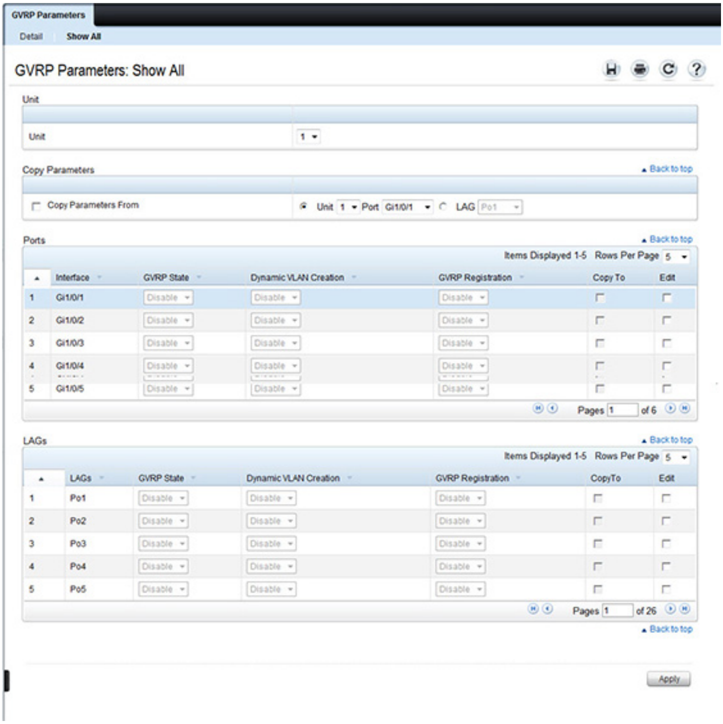
Figure 21-16. GVRP Parameters



From the **GVRP Parameters** page, click **Show All** to see the GVRP configuration for all ports. From this page, settings can be changed for one or more entries.

NOTE: Per-port and per-LAG GVRP Statistics are available from the Statistics/RMON page. For more information, see "Monitoring Switch Traffic " on page 519.

Figure 21-17. GVRP Port Parameters Table

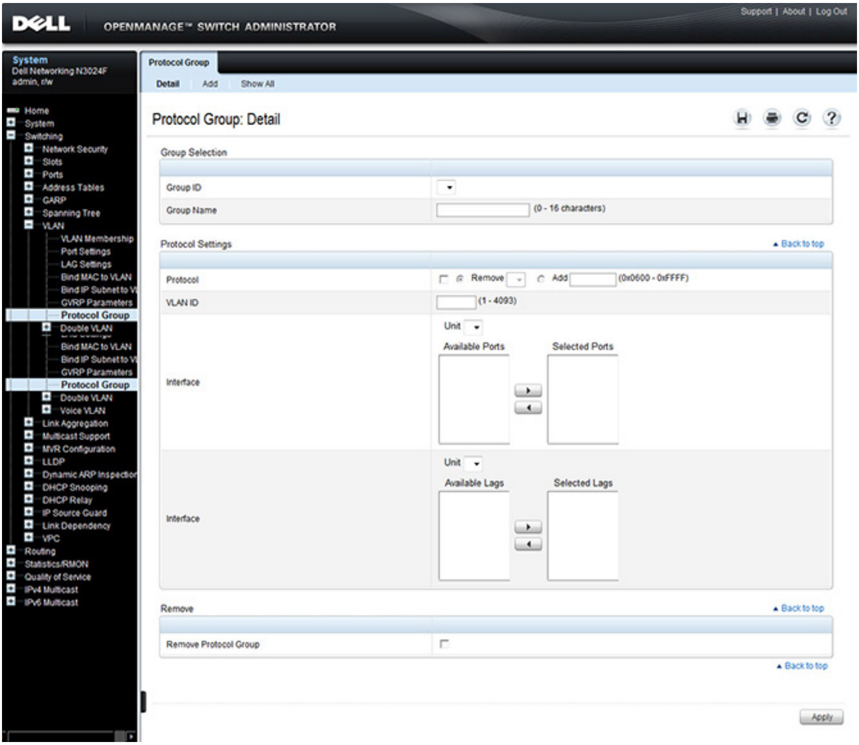


Protocol Group

Use the **Protocol Group** page to configure which EtherTypes go to which VLANs, and then enable certain ports to use these settings. Protocol-based VLANs are most often used in situations where network segments contain hosts running multiple protocols.

To display the **Protocol Group** page, click **Switching** → **VLAN** → **Protocol Group** in the navigation panel.

Figure 21-18. Protocol Group



Adding a Protocol Group

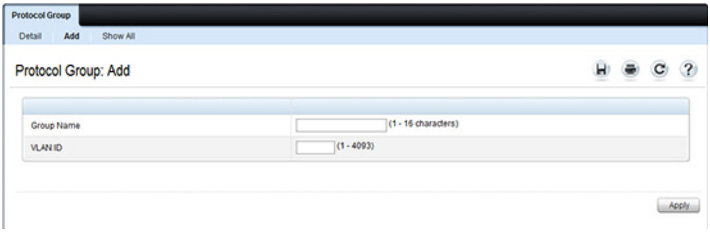
To add a protocol group:

1 Open the **Protocol Group** page.

2 Click **Add** to display the **Add Protocol Group** page.

3 Create a name for the group and associate a VLAN with the group.

Figure 21-19. Add Protocol Group



4 Click **Apply**.

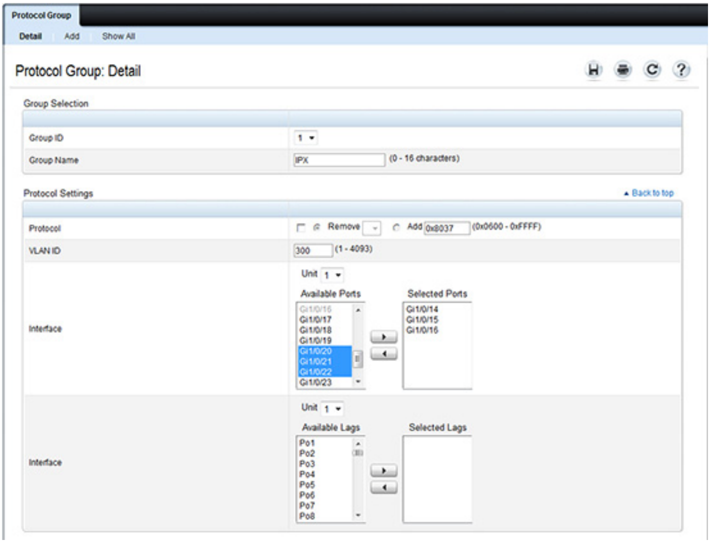
5 Click **Protocol Group** to return to the main **Protocol Group** page.

6 From the **Group ID** field, select the group to configure.

7 In the **Protocol Settings** table, select the protocol and interfaces to associate with the protocol-based VLAN.

In Figure 21-20, the Protocol Group 1 (named IPX) is associated with the IPX protocol and ports 14–16. Ports 20-22 are selected in **Available Ports** list. After clicking the right arrow, they will be added to the **Selected Ports** list.

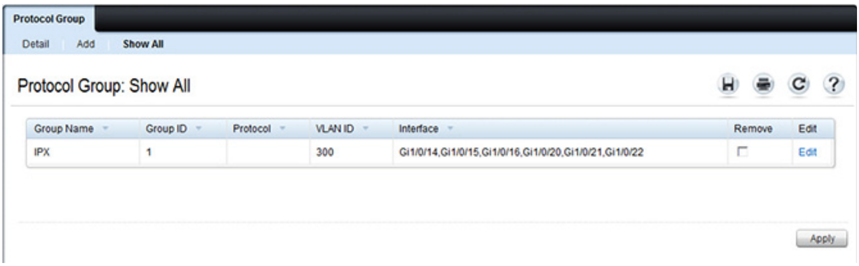
Figure 21-20. Configure Protocol Group



8 Click **Apply**.

9 Click **Show All** to see the protocol-based VLANs and their members.

Figure 21-21. Protocol Group Table

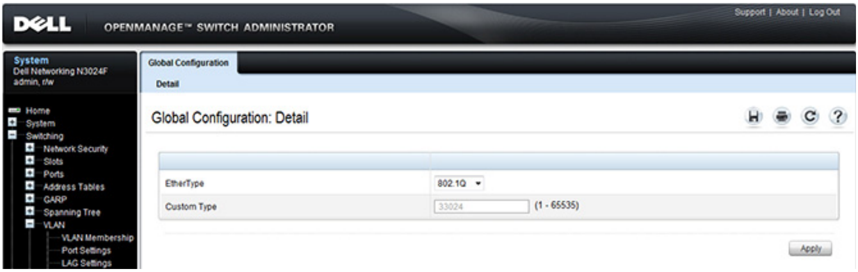


Double VLAN Global Configuration

Use the **Double VLAN Global Configuration** page to specify the value of the **EtherType** field in the first EtherType/tag pair of the double-tagged frame.

To display the **Double VLAN Global Configuration** page, click **Switching** → **VLAN** → **Double VLAN** → **Global Configuration** in the navigation panel.

Figure 21-22. Double VLAN Global Configuration

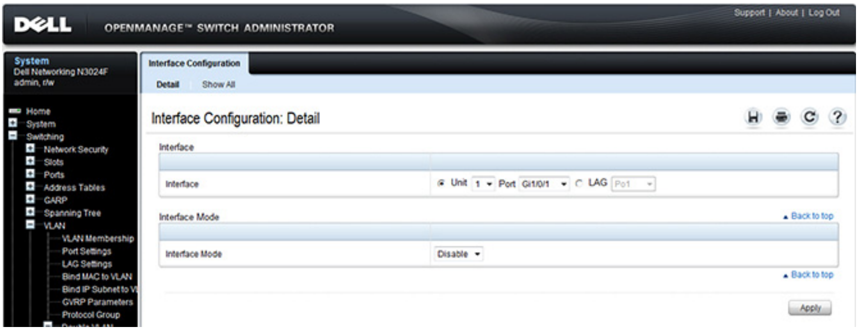


Double VLAN Interface Configuration

Use the **Double VLAN Interface Configuration** page to specify the value of the **EtherType** field in the first EtherType/tag pair of the double-tagged frame.

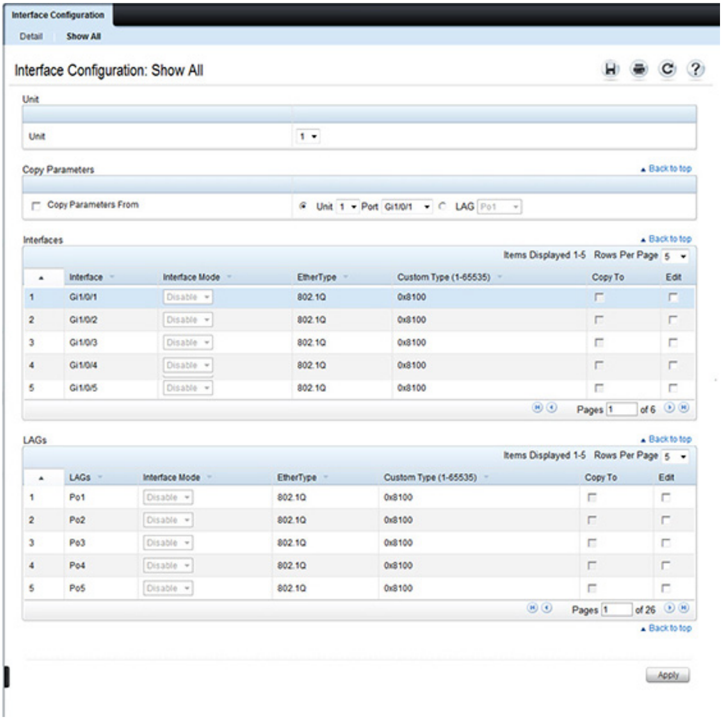
To display the **Double VLAN Interface Configuration** page, click **Switching** → **VLAN** → **Double VLAN** → **Interface Configuration** in the navigation panel.

Figure 21-23. Double VLAN Interface Configuration



To view a summary of the double VLAN configuration for all interfaces and to edit settings for one or more interfaces, click **Show All**.

Figure 21-24. Double VLAN Port Parameter Table

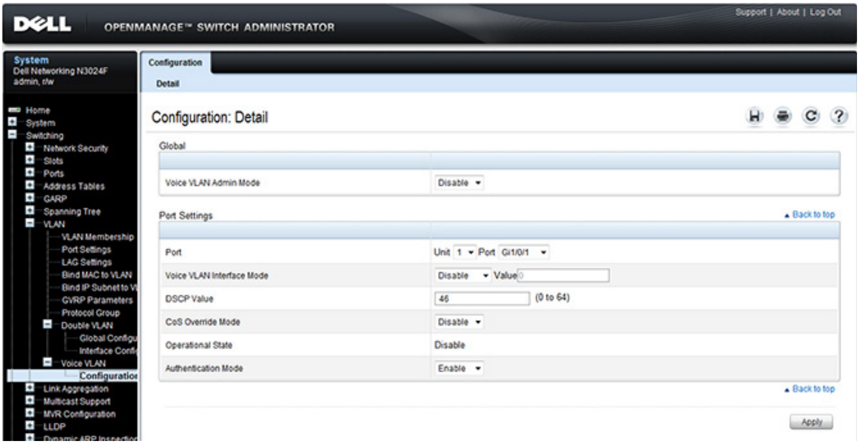


Voice VLAN

Use the **Voice VLAN Configuration** page to configure and view voice VLAN settings that apply to the entire system and to specific interfaces.

To display the page, click **Switching** → **VLAN** → **Voice VLAN** → **Configuration** in the navigation panel.

Figure 21-25. Voice VLAN Configuration



NOTE: IEEE 802.1X must be enabled on the switch before you disable voice VLAN authentication. Voice VLAN authentication can be disabled in order to allow VoIP phones that do not support authentication to send and receive unauthenticated traffic on the Voice VLAN.

Configuring VLANs (CLI)

This section provides information about the commands you use to create and configure VLANs. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Creating a VLAN

Beginning in Privileged EXEC mode, use the following commands to configure a VLAN and associate a name with the VLAN.

Command	Purpose
configure	Enter global configuration mode.
vlan {*vlan-id* | *vlan-range*}	Create a new VLAN or a range of VLANs and enter the interface configuration mode for the specified VLAN or VLAN range. • *vlan-id*—A valid VLAN IDs (Range: 2–4093). • *vlan-range* — A list of valid VLAN IDs to be added. List separate, non-consecutive VLAN IDs separated by commas (without spaces); use a hyphen to designate a range of IDs. (Range: 2–4093)
name *string*	Add a name to the specified VLAN. *string* — Comment or description to help identify a specific VLAN (Range: 1–32 characters).
CTRL + Z	Exit to Privileged EXEC mode.
show vlan [**id** *vlan-id* | **name** *vlan-name*]	Display VLAN information. • *vlan-id* — A valid VLAN ID. (Range: 1–4093) • *vlan-name* — A valid VLAN name string. (Range: 1–32 characters)

Configuring VLAN Settings for a LAG

The VLAN mode and memberships settings you configure for a port are also valid for a LAG (port-channel). Beginning in Privileged EXEC mode, use the following commands to configure the VLAN mode for a LAG. Once the switchport mode settings are specified for a LAG, other VLAN memberships settings can be specified that are valid for the switchport mode.

Command	Purpose
configure	Enter global configuration mode.
interface port-channel *channel-id*	Enter interface configuration mode for the specified interface. *channel-id* — Specific port-channel. (Range 1–48). A range of LAGs can be specified using the **interface range port-channel** command. For example, **interface range port-channel 4-**8.
switchport mode [**access**|**general**|**trunk**]	Configure the interface as an untagged layer-2 VLAN interface.
CTRL + Z	Exit to Privileged EXEC mode.
show interfaces switchport port-channel *channel-id*	Display information about the VLAN settings configured for the specified LAG.

Configuring Double VLAN Tagging

Dell Networking N-Series switches use switchport dot1q-tunnel mode to configure an interface as a customer edge (CE) interface. The dot1q-tunnel mode is an overlay on switchport access mode. In particular, configuring the access mode PVID sets the outer dot1q-tunnel VLAN ID. Changing the switchport mode on a CE port to access, general, or trunk, effectively disables tunneling on the interface.

CE interfaces can be physical ports or port-channels. Untagged frames received on the CE interface are processed as if they belong to the PVID and are transmitted out the service provider (SP) interface with a single VLAN tag (presuming that the destination MAC address has been learned on the SP interface). Tagged frames received on the CE interface are transmitted out the service provider (SP) interface with an outer tag containing the access mode native VLAN ID and the inner tag as received on the CE interface.

CE interfaces MUST be configured in dot1q-tunnel mode with the PVID configured with the outer tag (native) VLAN ID for the associated service provider (SP) interface. Configure the outer VLAN ID using the **switchport mode access vlan** command. All MAC address learning and forwarding occurs on the outer VLAN tag MAC addresses. The VLAN ID must be common to both the SP port and the CE ports.

The service provider interface MUST be configured for egress tagging (trunk or general mode) with a native VLAN identical to the PVID of the associated CE ports. SP interfaces SHOULD be configured with a single VLAN ID. Be aware that a trunk mode port accepts untagged packets on the native VLAN and be a member of any existing or newly created VLANs by default.

It is not possible to configure an inner VLAN TPID value. The inner VLAN TPID value is always 802.1Q (0x8100). Up to four unique outer TPIDs may be configured in the system. An outer TPID/EtherType (other than 802.1Q) must be configured in Global Configuration mode prior to configuration on an interface. The outer TPID/Ethertype must be configured on the interface prior to putting the interface into tunnel mode.

Multiple groups of associated CE and SP ports can be defined by configuring the groups with unique VLAN IDs. An outer TPID/EtherType (other than 802.1Q) must be configured in Global Configuration mode prior to configuration on an interface. The outer TPID/Ethertype must be configured on the interface prior to putting the interface into tunnel mode.

DVLAN CE interfaces must be configured for tagging (dot1q-tunnel mode) for double tags to be observed on frames egressing the service provider (SP) interface. The DVLAN uplink interface should be configured to accept tagged frames for the DVLAN or outer VLAN (trunk or general mode). Ensure that the native (access mode) VLAN on the customer edge (CE) port is set to the DVLAN ID. MAC address learning on DVLAN enabled ports occurs on the DVLAN CE port's native VLAN.

If it is desirable to restrict propagation of spanning tree topology changes from CE interfaces into the service provider network, enable **spanning-tree tcnguard** on the CE interfaces. Optionally, use the **spanning-tree guard root** on CE ports to eliminate the possibility that a CE interface becomes a root port. Be aware that root guard may cause spanning-tree to stop forwarding if a superior BPDU is received on the interface.

Perform the following steps to configure an interface as a CE interface. The DVLAN VLAN must also be configured on an interface with tagging enabled (trunk or general mode), and the native VLAN must be set to the DVLAN VLAN identifier.. That interface will act as the SP interface. It is advisable to restrict the allowed VLAN on the SP interface to the DVLAN VLAN only.

Command	Purpose
configure	Enter global configuration mode.
vlan 100	Create the DVLAN (outer) VLAN.
exit	Exit VLAN configuration mode
switchport dot1q ethertype vman	Define the VMAN Ethertype for use on the CE port.
interface *interface-id*	Enter interface configuration mode for the specified CE interface. The *interface-id* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
switchport mode dot1q-tunnel	Configure the interface as a CE tunnel port.
switchport access vlan 100	Configure the DVLAN VLAN

Command	Purpose
spanning-tree guard root	(Optional) Disable the ability of the CE port to become spanning tree root.
spanning-tree tcnguard	(Optional) Ignore topology changes received from CE ports.
exit	Exit to global configuration mode
CTRL + Z	Exit to Privileged EXEC mode.
show dot1q-tunnel	Display all interfaces enabled for Double VLAN Tunneling
show dot1q-tunnel interface {*interface-id*| **all**}	Display detailed information about Double VLAN Tunneling for the specified interface or all interfaces.
show switchport ethertype [interface *interface-id* | **all**]	Display the configured Ethertype for each interface.

Beginning in global configuration mode, perform the following steps to configure the SP port using the VMAN (0x88A8) outer tag. In this configuration, packets received on CE ports will have the VMAN tag pushed onto the frame before being transmitted out the SP interface. Packets received on the SP interface will have the VMAN tag stripped before forwarding to the CE interface.

Command	Purpose
switchport dot1q ethertype {**802.1Q** | **vman** | **custom** *0-65535*} [*primary-tpid*]	Configure the EtherType to use for an SP interface using one of the previously configured EtherTypes. • **802.1Q** — Configures the EtherType as 0x8100 (default). • **vman** — Configures the EtherType as 0x88A8. • **custom** — Configure a custom EtherType for the DVLAN tunnel. The value must be 0–65535. • *primary-tpid* — Configure the primary (outer) TPID. Up to four unique outer VLAN tag TPIDs may be configured.
interface *interface-id*	Enter interface configuration mode for the SP uplink port.
switchport mode trunk	Configure the interface in trunk mode.

Command	Purpose
switchport trunk allowed vlan 100	Only allow VLAN 100 packets on the interface.
switchport trunk native vlan 100	Configure untagged packets to be members of VLAN 100.

Configuring MAC-Based VLANs

Beginning in Privileged EXEC mode, use the following commands to associate a MAC address with a configured VLAN. The VLAN does not need to be configured on the system to associate a MAC address with it. However, the associated VLAN must be configured on a port in order for the system to map packets matching the MAC address to the associated VLAN and to learn the associated MAC address on the associated VLAN so that packets addressed to the associated MAC address are forwarded properly. Up to 256 VLAN to MAC address associations can be created. VLAN associations operate on untagged packets on access and trunk ports. Tagged traffic is associated with the VLAN identified in the VLAN tag.

Command	Purpose
configure	Enter global configuration mode.
interface gi1/0/3	Enter Interface Config mode for port gil/0/3.
switchport mode trunk	Configure gi1/0/3 as a trunk port associated with the selected VLAN ID. Note that the native VLAN for gi1/0/3 is still VLAN 1. Untagged traffic with the associated MAC address is learned on the associated VLAN ID, not VLAN 1.
exit	Exit to Global Config mode.
interface gi1/0/4	Enter Interface Config mode for port gi1/0/4.
switchport access vlan *vlanid*	Configure gi1/0/4 as an access port. The PVID for Gi1/0/4 is the associated VLAN ID. It will receive the MAC associated traffic.
exit	Exit to Global Config mode.
vlan *vlanid*	Enter VLAN configuration mode.

Command	Purpose
vlan association mac *mac-address*	Associate a MAC address with a VLAN. • *mac-address* — MAC address to associate. (Range: Any MAC address in the format xxxx.xxxx.xxxx or xx:xx:xx:xx:xx:xx)
CTRL + Z	Exit to Privileged EXEC mode.
show vlan association mac [*mac-address*]	Display the VLAN associated with a specific configured MAC address. If no MAC address is specified, the VLAN associations of all the configured MAC addresses are displayed.

Configuring IP-Based VLANs

Beginning in Privileged EXEC mode, use the following commands to associate an IP subnet with a configured VLAN. The VLAN does not need to be configured on the system to associate an IP subnet with it. However, the subnet VLAN must be configured on a port in order for the system to map packets matching the IP address to the subnet VLAN and to learn the associated MAC address on the subnet VLAN so that packets addressed to the associated IP address are forwarded properly. Up to 256 VLAN-to-IP address associations can be created.

It is not necessary to assign IP addresses to VLANs in order to utilize subnet associations. Untagged packets are switched into the subnet VLAN using the defined subnet address and from the IP subnet VLAN using the learned MAC addresses.

VLAN associations operate on untagged packets on access and trunk ports. Tagged traffic is associated with the VLAN identified in the VLAN tag.

Command	Purpose
configure	Enter global configuration mode.
vlan *vlanid*	Enter VLAN configuration mode.
vlan association subnet *ip-address subnet-mask vlanid*	Associate an IP subnet with a VLAN. • *ip-address* — Source IP address. (Range: Any valid IP address) • *subnet-mask* — Subnet mask. (Range: Any valid subnet mask) • *vlanid* — VLAN to associated with subnet. (Range: 1-4093)
exit	Exit to Global Config mode.
interface gi1/0/3	Enter Interface Config mode for gi1/0/3.
switchport mode trunk	Configure gi1/0/3 as a trunk member of the subnet VLAN. The Native VLAN is 1 but the port is a member of the subnet VLAN.
exit	Exit to Global Config mode.
interface gi1/0/4	Enter Interface Config mode for gi1/0/4.
switchport mode access	Configure gi1/0/4 as an access port.

Command	Purpose
switchport access vlan *vlanid*	Specify the subnet VLAN ID of which gi1/0/4 is an access port member.
exit	Exit to Global Config mode.
CTRL + Z	Exit to Privileged EXEC mode.
show vlan association subnet [*ip-address ip-mask*]	Display the VLAN associated with a specific configured IP-Address and netmask. If no IP Address and net mask are specified, the VLAN associations of all the configured IP-subnets are displayed.

Configuring a Protocol-Based VLAN

Beginning in Privileged EXEC mode, use the following commands to create and name a protocol group, and associate VLANs with the protocol group. When you create a protocol group, the switch automatically assigns it a unique group ID number. The group ID is used for both configuration and script generation to identify the group in subsequent commands.

A protocol group may have more than one interface associated with it, but each interface and protocol combination can be associated with one group only. If adding an interface to a group causes any conflicts with protocols currently associated with the group, adding the interface(s) to the group fails and no interfaces are added to the group. Ensure that the referenced VLAN is created prior to the creation of the protocol-based group except when GVRP is expected to create the VLAN.

VLAN associations operate on untagged packets on access and trunk ports. Tagged traffic is associated with the VLAN identified in the VLAN tag.

Command	Purpose
configure	Enter Global Config mode.
vlan protocol group *groupid*	Create a new protocol group.
vlan protocol group name *groupid name-string*	Name the protocol group.
exit	Exit to Global Config mode.
interface gi1/0/3	Enter Interface Config mode for gi1/0/3.
switchport mode trunk	Configure Gi1/0/3 as a trunk member of the associated VLAN. The Native VLAN is 1 but the port is member of the protocol VLAN.
exit	Exit to Global Config mode.
interface gi1/0/4	Enter Interface Config mode for gi1/0/4.
switchport mode access	Configure gi1/0/4 as an access port.
switchport access vlan *vlanid*	Specify the subnet VLAN ID of which gi1/0/4 is an access port member.

Command	Purpose
exit	Exit to Global Config Mode
show port protocol all	Obtain the group ID for the newly configured group.
configure	Enter global configuration mode.
vlan protocol group add protocol *groupid* **ethertype** *value*	Add any EtherType protocol to the protocol-based VLAN groups identified by *groupid*. A group may have more than one protocol associated with it. Each interface and protocol combination can be associated with one group only. If adding a protocol to a group causes any conflicts with interfaces currently associated with the group, this command fails and the protocol is not added to the group. • *groupid* — The protocol-based VLAN group ID. • *protocol* — The protocol you want to add. The ethertype can be any valid number in the range 0x0600-0xffff.
protocol vlan group all *groupid*	(Optional) Add all physical interfaces to the protocol-based group identified by *groupid*. Individual interfaces can be added to the protocol-based group as shown in the next two commands. *groupid* — The protocol-based VLAN group ID.
interface *interface*	Enter interface configuration mode for the specified interface. *interface* — Specific interface type and number, such as gi1/0/8.
protocol vlan group *groupid*	Add the physical unit/port interface to the protocol-based group identified by groupid. *groupid* — The protocol-based VLAN group ID.
exit	Exit to global configuration mode.
vlan *vlanid*	Enter VLAN configuration mode.

Command	Purpose
protocol group *groupid vlanid*	Attach a VLAN ID to the protocol-based group identified by groupid. A group may only be associated with one VLAN at a time. However, the VLAN association can be changed. • *groupid* — The protocol-based VLAN group ID, which is automatically generated when you create a protocol-based VLAN group with the **vlan protocol group** command. To see the group ID associated with the name of a protocol group, use the **show port protocol all** command. • *vlanid* — A valid VLAN ID.
CTRL + Z	Exit to Privileged EXEC mode.
show port protocol [**all** | *groupid*]	Display the Protocol-Based VLAN information for either the entire system or for the indicated group.

Configuring GVRP

Beginning in Privileged EXEC mode, use the following commands to enable GVRP on the switch and on an interface, and to configure various GVRP settings.

Command	Purpose
configure	Enter global configuration mode.
gvrp enable	Enable GVRP on the switch.
interface *interface*	Enter interface configuration mode for the specified port or LAG. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3** or **port-channel 3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
gvrp enable	Enable GVRP on the interface.
switchport general forbidden vlan {**add** *vlan-list*|**remove** *vlan-list*} —or— **switchport trunk allowed vlan** {**add** *vlan-list* | **remove** *vlan-list*}	(Optional) Forbids dynamically adding the VLANs specified by the **remove** parameter to a port. To revert to allowing the addition of specific VLANs to the port, use the **add** parameter of this command. **add** *vlan-list* — List of valid VLAN IDs to remove from the forbidden list. Separate nonconsecutive VLAN IDs with a comma and no spaces. Use a hyphen to designate a range of IDs. **remove** *vlan-list* — List of valid VLAN IDs to add to the forbidden list. Separate nonconsecutive VLAN IDs with a comma and no spaces. Use a hyphen to designate a range of IDs.
gvrp registration-forbid	(Optional) Deregister all VLANs on a port and prevent any dynamic registration on the port.
gvrp vlan-creation-forbid	(Optional) Disable dynamic VLAN creation.
exit	Exit to global configuration mode.

Command	Purpose
vlan makestatic *vlan-id*	(Optional) Change a dynamically created VLAN (one that is created by GVRP registration) to a static VLAN (one that is permanently configured and defined). *vlan-id* — Valid vlan ID. Range is 2-4093.
CTRL + Z	Exit to Privileged EXEC mode.
show gvrp configuration	Display GVRP configuration information. Timer values are displayed. Other data shows whether GVRP is enabled and which ports are running GVRP.
show vlan	Display the VLAN configuration, including the VLAN configuration type and the associated ports.

Configuring Voice VLANs

Beginning in Privileged EXEC mode, use the following commands to enable the Voice VLAN feature on the switch and on an interface.

Command	Purpose
configure	Enter global configuration mode.
voice vlan	Enable the voice VLAN capability on the switch.
interface *interface*	Enter interface configuration mode for the specified interface. *interface* — Specific interface, such as gi1/0/8. A range of interfaces can be specified using the **interface range** command. For example, **interface range gi1/0/8-12** enters Interface Configuration mode for ports 8–12.

Command	Purpose
voice vlan {*vlanid* | **dot1p** *priority* | **none** | **untagged** | **data** *priority* {**trust** | **untrust**} | **auth** {**enable** | **disable**} | **dscp** *value*}	Enable the voice VLAN capability on the interface. • *vlanid*—The voice VLAN ID. This VLAN ID is sent to IP phones via LLDP. • *priority*—The IEEE 802.1p priority sent to IP phones on the port. This value is transmitted to the IP phone via LLDP. The switch must be configured locally to give packets using the transmitted priority the appropriate QoS. • **none**—Allow the phone to use it's own configuration. The administrator must configure the switch appropriately to give voice packets the required QoS. • **untagged**—Configure the phone to send untagged traffic using LLDP. • **trust**—Trust the dot1p priority or DSCP values contained in packets arriving on the voice VLAN. • **untrust**—Do not trust the dot1p priority or DSCP values contained in packets arriving on the voice VLAN. • **auth** {**enable** | **disable**} — Use **enable** to allow voice traffic on an 802.1x unauthorized port. Use **disable** to prevent voice traffic on an 802.1x unauthorized port. • **dscp** *value*—The DSCP value (Range: 0–64). This value is transmitted to the IP phone via LLDP. The switch must be configured locally to give packets using the transmitted DSCP value the appropriate QoS.
CTRL + Z	Exit to Privileged EXEC mode.
show voice vlan [**interface** {*interface* |**all**}]	Display voice VLAN configuration information for the switch, for the specified interface, or for all interfaces.

Configuring a Voice VLAN (Extended Example)

The commands in this example create a VLAN for voice traffic with a VLAN ID of 25 using an IP phone that does not support 802.1X authentication. Port gi1/0/10 is set to an 802.1Q VLAN. In this example, there are multiple devices connected to port gi1/0/10, so the port must be in general mode in order to enable MAC-based 802.1X authentication. Next, Voice VLAN is enabled on the port with the Voice VLAN ID set to 25. Finally, Voice VLAN authentication is disabled on port gi1/0/10 because the phone connected to that port does not support 802.1X authentication. All other devices connected to the port are required to use 802.1X authentication for network access. For more information about 802.1X authentication, see "Port and System Security " on page 623.

This example shows the configuration for a switch with directly connected IP phones. The interior of the network will still require configuration of QoS on the selected voice VLAN in order to ensure service.

NOTE: In an environment where the IP phone uses LLDP-MED to obtain configuration information, an additional step to enable LLDP-MED on the interface would be required by issuing the **lldp med** command in Interface Configuration mode.

To configure the switch:

1 Create the voice VLAN.

```
console#configure
console(config)#vlan 25
console(config-vlan25)#exit
```

2 Enable the Voice VLAN feature on the switch.

```
console(config)#voice vlan
```

3 Configure port 10 to be in access mode.

```
console(config)#interface gi1/0/10
console(config-if-Gi1/0/10)#switchport mode access
```

4 Enable MAC-based 802.1X authentication on the port. This step is required only if there are multiple devices that use 802.1X authentication connected to the port. The authentication server will need to be configured with the MAC address of the IP phone.

```
console(config-if-Gi1/0/10)#dot1x port-control mac-based
```

5 Enable the voice VLAN feature on the interface

```
console(config-if-Gi1/0/10)#voice vlan 25
```

6 Disable authentication for the voice VLAN on the port. This step is required only if the voice phone does not support port-based authentication.

```
console(config-if-Gi1/0/10)#voice vlan auth disable
```

7 Exit to Privileged Exec mode.

```
console(config-if-Gi1/0/10)#<CTRL+Z>
```

8 View the voice VLAN settings for port 10.

```
console#show voice vlan interface gi1/0/10

Interface............................ Gi1/0/10
Voice VLAN Interface Mode............ Enabled
Voice VLAN ID........................ 25
Voice VLAN COS Override.............. False
Voice VLAN DSCP Value................ 46
Voice VLAN Port Status............... Disabled
Voice VLAN Authentication............ Disabled
```

Enterprise Voice VLAN Configuration With QoS

In this example, voice VLAN traffic is transmitted and received tagged on VLAN 25 using IEEE 802.1p user priority 5. Background traffic is carried on the default VLAN. The 802.1p user priority 5 tagged packets are mapped onto internal CoS queue 2. CoS queue 2 is additionally configured as strict priority to ensure that the latency-sensitive voice traffic is transmitted first. This is to help overcome the quantization effect of IP traffic, where the first voice sample in an IP packet is typically delayed 10 or 20 ms while the voice samples are collected. A rate-limiting ACL is applied to ensure that 802.1p priority 5 packets are limited in their ability to disrupt lower-priority traffic via a denial-of-service attack.

To configure the switch on the IP phone facing interface:

1 Create the voice VLAN.

```
console#configure
console(config)#vlan 25
console(config-vlan25)#exit
```

2 Globally enable the voice VLAN feature on the switch.

```
console(config)#voice vlan
```

3 Configure a rate-limiting ACL to ensure that the voice VLAN does not present a denial-of-service threat. A G.711 voice stream generates 64 Kbps, which translates to 80 bytes of uncompressed voice every 10 ms. Overhead adds 40 bytes, so the phone will generate 100 to 120 byte packets every second per voice stream, or about 96 Kbps. The rate limit below will permit a single voice stream.

```
console(config)#mac access-list extended dot1p-5-limit
console(config-mac-access-list)#permit any any cos 5 rate-
limit 100 64
console(config-mac-access-list)#permit any any
console(config-mac-access-list)#exit
```

4 Configure port 10 to be in access mode.

```
console(config)#interface gi1/0/10
console(config-if-Gi1/0/10)#switchport mode access
```

5 Configure the switch to tell the IP phone to use VLAN 25 and 802.1p priority 5 for voice traffic.

```
console(config-if-Gi1/0/10)#voice vlan 25
console(config-if-Gi1/0/10)#voice vlan dot1p 5
```

6 Enable IEEE 802.1p trust mode on the interface. This is the switch default.

```
console(config-if-Gi1/0/10)#classofservice dot1p trust
```

7 Configure internal CoS queue 2 as strict priority to ensure that egressing voice traffic is transmitted first on this interface. This reduces latency for transmitted voice traffic.

```
console(config-if-Gi1/0/10)#cos-queue strict 2
```

8 Map 802.1p priority 5 onto internal CoS queue 2. This is the switch default mapping.

```
console(config-if-Gi1/0/10)#classofservice dot1p-mapping 5 2
```

9 Rate limit incoming IEEE 802.1p priority 5 traffic

```
console(config-if-Gi1/0/10)#mac access-group dot1p-5-limit in
```

Steps 6–8 are required to be configured on all ports that carry voice traffic end-to-end, including the switch ports connected to other switches and the ports on other switches that will carry voice traffic. It may be desirable to configure steps 6–8 globally.

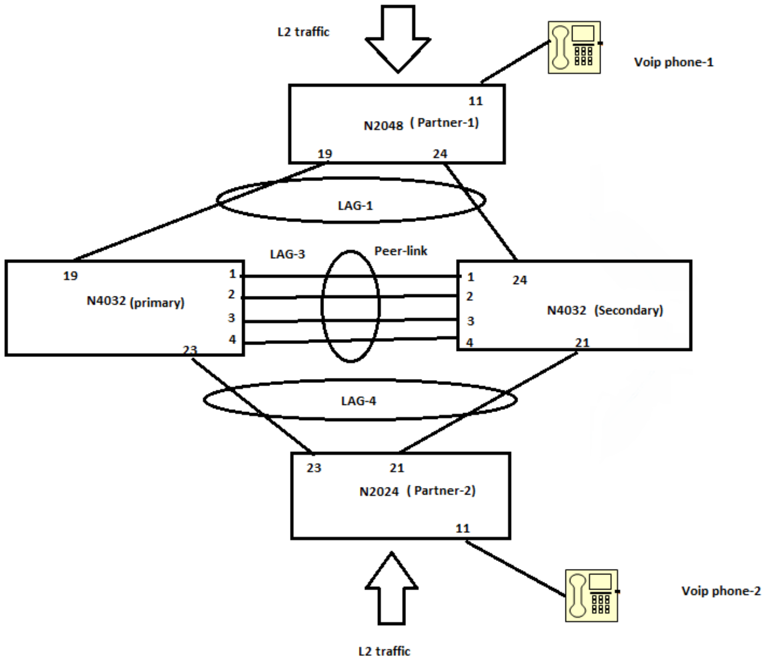
Step 9 should be configured on all ports connected to IP phones if using strict priority or perhaps on all host facing ports if IP phones are moved frequently. Do not configure steps 3 or 9 on inter-switch connections as they will be used to aggregate voice traffic.

When configuring an MLAG for transport of voice VLAN traffic, remember to configure steps 6-8 on the corresponding MLAG/Voice VLAN and both ends of the MLAG peer link (or configure them globally on both peers and the partner switches).

MLAG with RPVST and Voice VLAN

Figure 21-26 describes the reference topology. It is presumed that a call manager of some type is connected to the network in the voice VLAN.

Figure 21-26. Network Topology for LAG with RPVST and Voice VLAN



MLAG Primary Peer Configuration

1 Configure the MLAG primary switch.

Keepalives are disabled on the peer links (optional). The four peer-links are placed in port-channel 3. Port-channel 1 is the northbound (partner 1) MLAG interface in VPC 1 and port-channel 4 is the southbound (partner 2) interface in VPC 4. Finally, VPC is enabled and the VPC domain is set to 1.

```
console#config
console(config)#interface Te1/0/1
console(config-if-Te1/0/1)#channel-group 3 mode active
console(config-if-Te1/0/1)#no keepalive
console(config-if-Te1/0/1)#exit

console(config)#interface Te1/0/2
console(config-if-Te1/0/2)#channel-group 3 mode active
console(config-if-Te1/0/2)#no keepalive
console(config-if-Te1/0/2)#exit

console(config)#interface Te1/0/3
console(config-if-Te1/0/3)#channel-group 3 mode active
console(config-if-Te1/0/3)#no keepalive
console(config-if-Te1/0/3)#exit

console(config)#interface Te1/0/4
console(config-if-Te1/0/4)#channel-group 3 mode active
console(config-if-Te1/0/4)#no keepalive
console(config-if-Te1/0/4)#exit

console(config)#interface Te1/0/19
console(config-if-Te1/0/19)#channel-group 1 mode active
console(config-if-Te1/0/19)#no keepalive
console(config-if-Te1/0/19)#exit

console(config)#interface Te1/0/23
console(config-if-Te1/0/23)#channel-group 4 mode active
console(config-if-Te1/0/23)#no keepalive
console(config-if-Te1/0/23)#exit

console(config)#interface port-channel 1
console(config-if-Po1)#vpc 1
console(config-if-Po1)#switchport mode trunk
console(config-if-Po1)#exit

console(config)#interface port-channel 3
```

```
console(config-if-Po3)#vpc peer-link
console(config-if-Po3)#switchport mode trunk
console(config-if-Po3)#exit

console(config)#interface port-channel 4
console(config-if-Po4)#vpc 4
console(config-if-Po4)#switchport mode trunk
console(config-if-Po4)#exit

console(config)#feature vpc
console(config)#vpc domain 1
console(config-vpc 1)#peer-keepalive enable
console(config-vpc 1)#exit
```

2 Disable loop protect on all interfaces.

```
console(config)#interface range tengigabitethernet all
console(config-if)#no keepalive
console(config-if)#exit
```

3 Configure spanning-tree mode as RPVST.

```
console(config)#spanning-tree mode rapid-pvst
```

4 Create VLAN-2 for voice traffic.

```
console(config)#vlan 2
console(config)#exit
```

5 Enable voice VLAN globally.

```
console(config)#voice vlan
```

6 Configure CoS queue 2 as strict. By default, the VoIP phone sends voice traffic with 802.1p priority 5, which is mapped to CoS queue 2 by default.

```
console(config)#cos-queue strict 2
```

MLAG Secondary Peer Device Configuration

1 Configure the secondary MLAG peer device. The peer links, up links, and down links correspond to those configured on the primary MLAG peer.

```
console#config
console(config)#interface Te1/0/1
console(config-if-Te1/0/1)#channel-group 3 mode active
console(config-if-Te1/0/1)#exit

console(config)#interface Te1/0/2
console(config-if-Te1/0/2)#channel-group 3 mode active
console(config-if-Te1/0/2)#no keepalive
console(config-if-Te1/0/2)#exit
```

```
console(config)#interface Te1/0/3
console(config-if-Te1/0/3)#channel-group 3 mode active
console(config-if-Te1/0/3)#no keepalive
console(config-if-Te1/0/3)#exit

console(config)#interface Te1/0/4
console(config-if-Te1/0/4)#channel-group 3 mode active
console(config-if-Te1/0/4)#no keepalive
console(config-if-Te1/0/4)#exit

console(config)#interface Te1/0/21
console(config-if-Te1/0/21)#channel-group 4 mode active
console(config-if-Te1/0/21)#no keepalive
console(config-if-Te1/0/21)#exit

console(config)#interface Te1/0/24
console(config-if-Te1/0/24)#channel-group 1 mode active
console(config-if-Te1/0/24)#no keepalive
console(config-if-Te1/0/24)#exit

console(config)#interface port-channel 1
console(config-if-Po1)#vpc 1
console(config-if-Po1)#switchport mode trunk
console(config-if-Po1)#exit

console(config)#interface port-channel 3
console(config-if-Po3)#vpc peer-link
console(config-if-Po3)#switchport mode trunk
console(config-if-Po3)#exit

console(config)#interface port-channel 4
console(config-if-Po4)#vpc 4
console(config-if-Po4)#switchport mode trunk
console(config-if-Po4)#exit

console(config)#feature vpc
console(config)#vpc domain 1
console(config-vpc 1)#peer-keepalive enable
console(config-vpc 1)#exit
```

2 Disable loop protect feature on all the interfaces.

```
console(config)#interface range tengigabitethernet all
console(config-if)#no keepalive
console(config-if)#exit
```

3 Configure spanning-tree mode as RPVST.

```
console(config)#spanning-tree mode rapid-pvst
```

4 Create VLAN 2 for voice traffic. This configuration must be identical on both MLAG peers.

```
console(config)#vlan 2
console(config-vlan-2)#exit
```

5 Enable voice VLAN globally.

```
console(config)#voice vlan
```

6 Configure egress queue 2 as strict. By default, the VoIP phone sends voice traffic with 802.1p priority 5, which is mapped to egress queue 2 by default. This configuration must be identical on both MLAG peers.

```
console(config)#cos-queue strict 2
```

MLAG Partner Switch Configuration

1 Configure partner switch 1 with a port-channel connected to the MLAG aware switches.

```
console#config
console(config)#interface Gi1/0/19
console(config-if-Gi1/0/19)#channel-group 1 mode active
console(config-if-Gi1/0/19)#no keepalive
console(config-if-Gi1/0/19)#exit

console(config)#interface Gi1/0/24
console(config-if-Gi1/0/24)#channel-group 1 mode active
console(config-if-Gi1/0/24)#no keepalive
console(config-if-Gi1/0/24)#exit

console(config)#interface port-channel 1
console(config-if-Po1)#switchport mode trunk
console(config-if-Po1)#exit
```

2 Disable the loop protect feature on all the interfaces (optional).

```
console(config)#interface range gigabitethernet all
console(config-if)#no keepalive
console(config-if)#exit
```

3 Configure spanning-tree mode as RPVST.

```
console(config)#spanning-tree mode rapid-pvst
```

4 Create VLAN-2 for voice traffic. This configuration must be the same as on the MLAG aware switches.

```
console(config)#vlan 2
```

```
console(config-vlan-2)#exit
```

5 Enable voice VLAN globally.

```
console(config)#voice vlan
```

6 Configure the VoIP phone connected port. The voice VLAN assignment must be the same on all switches.

```
console(config)#interface Gi2/0/11
console(config-if-Gi2/0/11)#switchport mode access
console(config-if-Gi2/0/11)#voice vlan 2
console(config-if-Gi2/0/11)#exit
```

7 Configure egress queue 2 as strict. By default, the VoIP phone sends voice traffic with 802.1p priority 5, which is mapped to egress queue 2 by default.

```
console(config)#cos-queue strict 2
```

8 Configure an ACL to rate-limit the voice traffic in case of DoS attacks and apply the ACL on the phone-connected interfaces. The administrator should consider whether to apply this configuration on all perimeter ports.

```
console(config)#mac access-list extended dot1p-5-limit
console(config-mac-access-list)#1000 permit any any cos 5
console(config-mac-access-list)#rate-limit 1024 128
console(config-mac-access-list)#1010 permit any any
console(config-mac-access-list)#exit

console(config)#interface Gi2/0/11
console(config-if-Gi2/0/11)#mac access-group dot1p-5-limit in
1
console(config-if-Gi2/0/11)#exit
```

Non-MLAG aware device-2 (Partner-2)

1 Configure partner-2 with the following configuration. This configuration is highly similar to the partner 1 configuration.

```
console#config
console(config)#interface Gi1/0/21
console(config-if-Gi2/0/21)#channel-group 4 mode active
console(config-if-Gi2/0/21)#no keepalive
console(config-if-Gi2/0/21)#exit

console(config)#interface Gi1/0/23
console(config-if-Gi1/0/23)#channel-group 4 mode active
console(config-if-Gi1/0/23)#no keepalive
console(config-if-Gi1/0/23)#exit
```

```
console(config)#interface port-channel 4
console(config-if-Po4)#switchport mode trunk
console(config-if-Po4)#exit
```

2 Disable loop protect on all the interfaces (optional).

```
console(config)#interface range gigabitethernet all
console(config-if)#no keepalive
console(config-if)#exit
```

3 Configure spanning-tree mode as RPVST.

```
console(config)#spanning-tree mode rapid-pvst
```

4 Create VLAN 2 for voice traffic. All switches must be configured identically for the voice VLAN.

```
console(config)#vlan 2
console(config-vlan-2)#exit
```

5 Enable voice VLAN globally.

```
console(config)#voice vlan
```

6 Configure the VoIP phone connected port as follows:

```
console(config)#interface Gi2/0/11
console(config-if-Gi2/0/11)#switchport mode access
console(config-if-Gi2/0/11)#voice vlan 2
console(config-if-Gi2/0/11)#exit
```

7 Configure CoS queue 2 as strict. By default, the VoIP phone sends voice traffic with 802.1p priority 5, which is mapped to egress queue 2 by default.

```
console(config)#cos-queue strict 2
```

8 Configure an ACL to rate-limit the voice traffic in case of DoS attacks and apply the ACL on the port-channel interfaces. The administrator should consider applying this configuration to all perimeter ports.

```
console(config)#mac access-list extended dot1p-5-limit
console(config-mac-access-list)#1000 permit any any cos 5
console(config-mac-access-list)#rate-limit 1024 128
console(config-mac-access-list)#1010 permit any any
console(config-mac-access-list)#exit

console(config)#interface Gi2/0/11
console(config-if-Gi2/0/11)#mac access-group dot1p-5-limit in
100
console(config-if-Gi2/0/11)#exit
```

NOTE: Spanning-tree status is shown accurately on the MLAG primary switch and on the partner switches. On the MLAG secondary switch, interfaces may show as spanning-tree disabled, but will remain in and are shown in the forwarding state.

Assigning an 802.1p Priority to VLAN Traffic

The following example assigns all traffic on VLAN 25 to internal CoS queue 4. This might be useful when assigning voice traffic a higher priority than normal data traffic. Note that CoS queue 4 shares scheduling with the other CoS queues, albeit more frequently than the lower-number CoS queues.

To ensure that CoS queue 4 packets are always transmitted first, CoS queue 4 could be made a strict-priority queue. In this case, it would be prudent to rate limit CoS queue 4 traffic.

1 Create an access list that permits all traffic and assign it to CoS queue 4.

```
console#config
console(config)#ip access-list voice-vlan
console(config-ip-acl)#permit every assign-queue 4
console(config-ip-acl)#exit
```

2 Assign the access list to VLAN 25. The access-group is given sequence number 100.

```
console(config)#interface vlan 25
console(config-if-vlan25)#ip access-group voice-vlan in 100
console(config-if-vlan25)#exit
```

Configuring a Private VLAN

1 Configure the VLANs and their roles. This example configures VLAN 100 as the primary VLAN, secondary VLAN 101 as the community VLAN and secondary VLANs 102 and 103 as the isolated VLANs:

```
switch#configure
switch(config)#vlan 100
switch(config-vlan-100)#private-vlan primary
switch(config-vlan-100)#exit
switch(config)#vlan 101
switch(config-vlan-101)#private-vlan community
switch(config-vlan-101)#exit
switch(config)#vlan 102
switch(config-vlan-102)#private-vlan isolated
```

```
switch(config-vlan-102)#exit
switch(config)#vlan 103
switch(config-vlan-103)#private-vlan isolated
switch(config-vlan-103)#exit
```

2 Associate the community and isolated VLANs with the primary VLAN.

```
switch(config)#vlan 100
switch(config-vlan-100)#private-vlan association 101-102
switch(config-vlan-100)#exit
```

This completes the configuration of the private VLAN. The only remaining step is to assign the ports to the private VLAN.

3 Assign the router connected port to the primary VLAN:

```
console(config)#interface te1/1/1
console(config-if-Te1/1/1)#switchport mode private-vlan
promiscuous
console(config-if-Te1/1/1)#switchport private-vlan mapping 100
101-102
console(config-if-Te1/1/1)#exit
```

4 Assign the community VLAN ports:

```
console(config)#interface gi1/0/11
console(config-if-Gi1/0/11)#switchport mode private-vlan host
console(config-if-Gi1/0/11)#switchport private-vlan host-
association 100 101
console(config-if-Gi1/0/11)#interface gi1/0/12
console(config-if-Gi1/0/12)#switchport mode private-vlan host
console(config-if-Gi1/0/12)#switchport private-vlan host-
association 100 101
```

5 Assign the isolated VLAN ports:

```
console(config)#interface gi1/0/10
console(config-if-Gi1/0/10)#switchport mode private-vlan host
console(config-if-Gi1/0/10)#switchport private-vlan host-
association 100 102
console(config-if-Gi1/0/10)#interface gi2/0/10
console(config-if-Gi2/0/10)#switchport mode private-vlan host
console(config-if-Gi2/0/10)#switchport private-vlan host-
association 100 102
console(config-if-Gi2/0/10)#interface gi2/0/11
console(config-if-Gi2/0/11)#switchport mode private-vlan host
console(config-if-Gi2/0/11)#switchport private-vlan host-
association 100 102
```

6 Show the configuration:

```
console(config)#show vlan private-vlan type

VLAN Type
---- -----------------------
100  primary
101  community
102  isolated
103  isolated

console#show vlan private-vlan

Primary VLAN Secondary VLAN Community
------------ -------------- -------------------
100          102            101

console(config)#show vlan

VLAN   Name        Ports          Type
-----  ----------- -------------  -------------
1      default     Po1-128,       Default
                   Te1/1/1,
                   Gi1/0/1-10,
                   Gi1/0/13-24
100    VLAN0100    Te1/1/1,       Static
                   Gi1/0/11-12
101    VLAN0101    Gi1/0/11       Static
102    VLAN0102    Gi1/0/12       Static
```

VLAN Configuration Examples

This section contains the following examples:

- Configuring VLANs Using The Dell OpenManage Administrator
- Configuring VLANs Using the CLI
- Configuring a Voice VLAN (Extended Example)

NOTE: For an example that shows how to use a RADIUS server to provide VLAN information, see "Controlling Authentication-Based VLAN Assignment " on page 305. For an example that shows how to allow the switch to dynamically create RADIUS-assigned VLANS, see "Allowing Dynamic Creation of RADIUS-Assigned VLANs " on page 308.

Configuring VLANs Using The Dell OpenManage Administrator

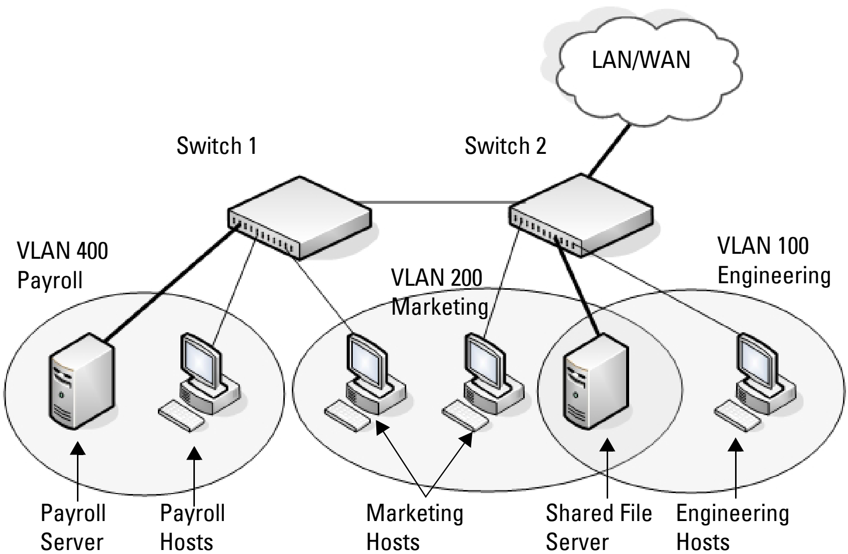
This example assumes that network administrator wants to create the VLANs in Table 21-8:

Table 21-8. Example VLANs

VLAN ID	VLAN Name	VLAN Type	Purpose
100	Engineering	Port-based	All employees in the Engineering department use this VLAN. Confining this department's traffic to a single VLAN helps reduce the amount of traffic in the broadcast domain, which increases bandwidth.
200	Marketing	Port-based	All employees in the Marketing department use this VLAN.
300	Sales	MAC-based	The sales staff works remotely but occasionally comes to the office. Since these employees do not have assigned work areas, they typically plug their laptops into a network port in an available cubicle, office, or conference room.
400	Payroll	Port-based	The payroll department has sensitive traffic and needs its own VLAN to help keep that traffic private.

Figure 21-27 shows the network topology for this example. As the figure shows, there are two switches, two file servers, and many hosts. One switch has an uplink port that connects it to a layer-3 device and the rest of the corporate network.

Figure 21-27. Network Topology for Port-Based VLAN Configuration



The network in Figure 21-27 has the following characteristics:

- Each connection to a host represents multiple ports and hosts.
- The Payroll and File servers are connected to the switches through a LAG.
- Some of the Marketing hosts connect to Switch 1, and some connect to Switch 2.
- The Engineering and Marketing departments share the same file server.
- Because security is a concern for the Payroll VLAN, the ports and LAG that are members of this VLAN will accept and transmit only traffic tagged with VLAN 400.
- The Sales staff might connect to a port on Switch 1 or Switch 2.

Table 21-9 shows the port assignments on the switches.

Table 21-9. Switch Port Connections

Port/LAG	Function
Switch 1	
1	Connects to Switch 2
2–15	Host ports for Payroll
16–20	Host ports for Marketing
LAG1 (ports 21–24)	Connects to Payroll server
Switch 2	
1	Connects to Switch 1
2–10	Host ports for Marketing
11–30	Host ports for Engineering
LAG1 (ports 35–39)	Connects to file server
LAG2 (ports 40–44)	Uplink to router.

This example shows how to perform the configuration by using the web-based interface.

Configure the VLANs and Ports on Switch 1

Use the following steps to configure the VLANs and ports on Switch 1. None of the hosts that connect to Switch 1 use the Engineering VLAN (VLAN 100), so it is not necessary to create it on that switch.

To configure Switch 1:

1 Create the Marketing, Sales, and Payroll VLANs.

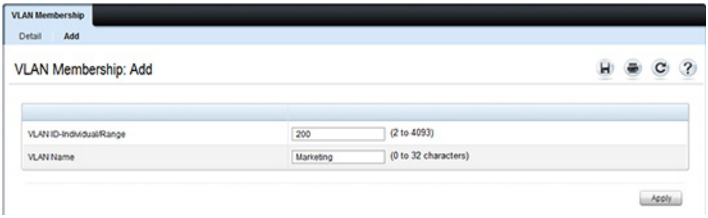
  a From the **Switching** → **VLAN** → **VLAN Membership** page, click **Add**.

  b In the **VLAN ID** field, enter 200.

  c In the **VLAN Name** field, enter Marketing.

  d Click **Apply**.

Figure 21-28. Add VLANs



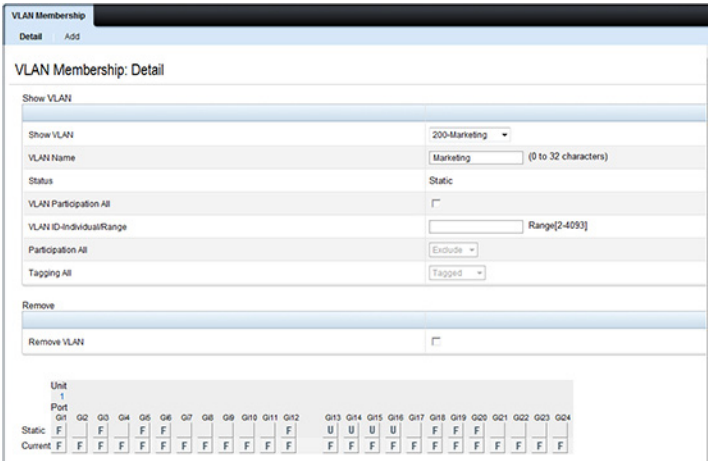
e Repeat steps b–d to create VLANs 300 (Sales) and 400 (Payroll).

2 Assign ports 16–20 to the Marketing VLAN.

a From the **Switching → VLAN → VLAN Membership** page, select 200-Marketing from the **Show VLAN** field.

b In the **Static** row, click the space for ports 13–16 so the U (untagged) displays for each port.

Figure 21-29. VLAN Membership - VLAN 200



3 Click **Apply**.

4 Assign ports 2–15 and LAG1 to the Payroll VLAN.

a From the **Switching → VLAN → VLAN Membership** page, select 400-Payroll from the Show VLAN field.

b In the Static row, click the space for ports 2–15 and LAG 1 so the U (untagged) displays for each port, and then click **Apply**.

5. Configure LAG 1 to be in general mode and specify that the LAG will accept tagged or untagged frames, but that untagged frames will be transmitted tagged with PVID 400.

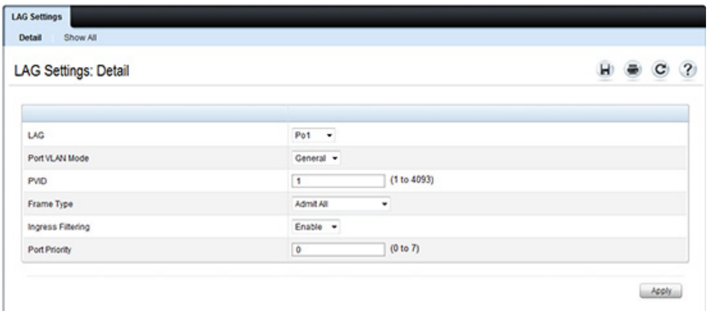
a. From the **Switching → VLAN → LAG Settings** page, make sure Po1 is selected.

b. Configure the following settings:

- Port VLAN Mode — General
- PVID — 400
- Frame Type — AdmitAll

c. Click **Apply**.

Figure 21-30. LAG Settings



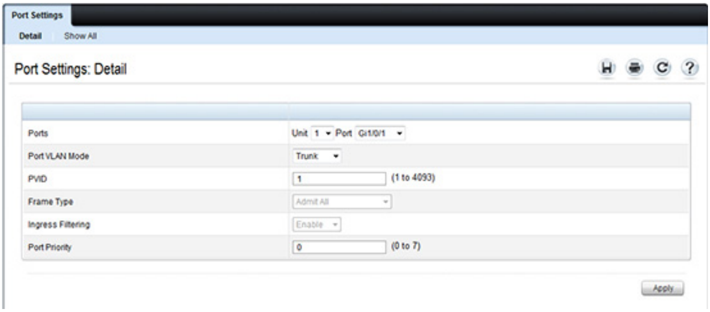
6 Configure port 1 as a trunk port.

a From the **Switching → VLAN → Port Settings** page, make sure port Gi1/0/1 is selected.

b From the **Port VLAN Mode** field, select Trunk.

c Click **Apply**.

Figure 21-31. Trunk Port Configuration



7 From the **Switching → VLAN → VLAN Membership** page, verify that port 1 is marked as a tagged member (T) for each VLAN.

Figure 21-32 shows VLAN 200, in which port 1 is a tagged member, and ports 13–16 are untagged members.

Figure 21-32. Trunk Port Configuration

8 Configure the MAC-based VLAN information.

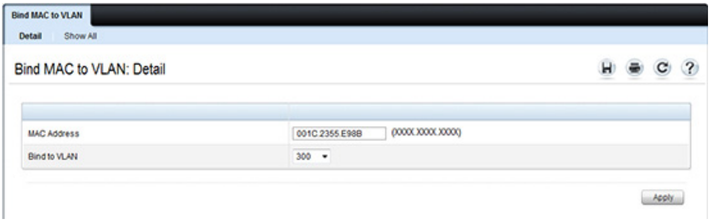
a Go to the **Switching → VLAN → Bind MAC to VLAN** page.

b In the **MAC Address** field, enter a valid MAC address, for example 00:1C:23:55:E9:8B.

c In the **Bind to VLAN** field, enter 300, which is the Sales VLAN ID.

d Click **Apply**.

Figure 21-33. Trunk Port Configuration



e Repeat steps b–d to add additional MAC address-to-VLAN information for the Sales department.

9 To save the configuration so that it persists across a system reset, use the following steps:

a Go to the **System** → **File Management Copy Files** page

b Select Copy Configuration and ensure that Running Config is the source and Startup Config is the destination.

c Click **Apply**.

Configure the VLANs and Ports on Switch 2

Use the following steps to configure the VLANs and ports on Switch 2. Many of the procedures in this section are the same as procedures used to configure Switch 1. For more information about specific procedures, see the details and figures in the previous section.

To configure Switch 2:

1. Create the Engineering, Marketing, Sales, and Payroll VLANs.

Although the Payroll hosts do not connect to this switch, traffic from the Payroll department must use Switch 2 to reach the rest of the network and Internet through the uplink port. For that reason, Switch 2 must be aware of VLAN 400 so that traffic is not rejected by the trunk port.

2. Configure LAG 1 as a general port so that it can be a member of multiple VLANs.

a. From the **Switching** → **VLAN** → **LAG Settings** page, make sure Po1 is selected.

b. From the **Port VLAN Mode** field, select General.

c. Click **Apply**.

3. Configure port 1 as a trunk port.
4. Configure LAG2 as a trunk port.
5. Assign ports 1–10 to VLAN 200 as untagged (U) members.
6. Assign ports 11–30 to VLAN 100 as untagged (U) members.
7. Assign LAG1 to VLAN 100 and 200 as a tagged (T) member.
8. Assign port 1 and LAG2 to VLAN 100, VLAN 200, VLAN 300, and VLAN 400 as a tagged (T) member.
9. Configure the MAC-based VLAN information.
10. If desired, copy the running configuration to the startup configuration.

Configuring VLANs Using the CLI

This example shows how to perform the same configuration by using CLI commands.

Configure the VLANs and Ports on Switch 1

Use the following steps to configure the VLANs and ports on Switch 1. None of the hosts that connect to Switch 1 use the Engineering VLAN (VLAN 100), so it is not necessary to create it on that switch.

To configure Switch 1:

1. Create VLANs 200 (Marketing), 300 (Sales), and 400 (Payroll), and associate the VLAN ID with the appropriate name.

```
console#configure
console(config)#vlan 200,300,400
console(config)#vlan 200
console(config-vlan200)#name Marketing
console(config-vlan200)#exit
console(config)#vlan 300
console(config-vlan300)#name Sales
console(config-vlan300)#exit
console(config)#vlan 400
console(config-vlan400)#name Payroll
console(config-vlan400)#exit
```

2. Assign ports 16–20 to the Marketing VLAN.

```
console(config)#interface range tengigabitEthernet 1/0/16-20
console(config-if)#switchport mode access
console(config-if)#switchport access vlan 200
console(config-if)#exit
```

3. Assign ports 2–15 to the Payroll VLAN

```
console(config)#interface range tengigabitEthernet 1/0/2-15
console(config-if)#switchport mode access
console(config-if)#switchport access vlan 400
console(config-if)#exit
```

4. Assign LAG1 to the Payroll VLAN and specify that frames will always be transmitted tagged with a VLAN ID of 400. By default, all VLANs are members of a trunk port.

```
console(config)#interface port-channel 1
console(config-if-Po1)#switchport mode trunk
console(config-if-Po1)#switchport trunk native vlan 400
console(config-if-Po1)#exit
```

5. Configure port 1 as a trunk port and add VLAN 200, VLAN 300, and VLAN 400 as members. All VLANs are added to trunk ports by default, including those created after the trunk port has been created.

```
console(config)#interface tengigabitEthernet 1/0/1
console(config-if-Te1/0/1)#switchport mode trunk
console(config-if-Te1/0/1)#exit
```

6. Configure the MAC-based VLAN information.

   The following commands show how to associate a system with a MAC address of 00:1C:23:55:E9:8B with VLAN 300. Repeat the **vlan association mac** command to associate additional MAC addresses with VLAN 300.

```
console(config)#vlan 10
console(config-vlan10)#vlan association mac 00:1C:23:55:E9:8B
300
console(config-vlan10)#exit
console(config)#exit
```

7. To save the configuration so that it persists across a system reset, use the following command:

```
console#copy running-config startup-config
```

8. View the VLAN settings.

```
console#show vlan

VLAN     Name        Ports          Type        Authorization
-----    ---------   ------------   ---------   -------------
1        Default     Po1-12,        Default     Required
                     Te1/0/2-15,
                     Te1/0/21-24
                     Te1/12
200      Marketing   Te1/0/1,       Static      Required
                     Te1/0/16-20
300      Sales       Te1/0/1        Static      Required
400      Payroll     Te1/0/1-15     Static      Required
```

9. View the VLAN membership information for a port.

```
console#show interfaces switchport te1/0/1

Port: Te1/0/1
VLAN Membership mode:Trunk Mode

Operating parameters:
PVID: 1
Ingress Filtering: Enabled
Acceptable Frame Type: VLAN Only
Default Priority: 0
GVRP status:Disabled
Protected:Disabled

Port Te1/0/1 is member in:

VLAN    Name                Egress rule    Type
----    ----------------- -----------    --------
200     Marketing           Tagged         Static
300     Sales               Tagged         Static
400     Payroll             Tagged         Static
```

Configure the VLANs and Ports on Switch 2

Use the following steps to configure the VLANs and ports on Switch 2. Many of the procedures in this section are the same as procedures used to configure Switch 1. For more information about specific procedures, see the details and figures in the previous section.

To configure Switch 2:

1. Create the Engineering, Marketing, Sales, and Payroll VLANs.

   Although the Payroll hosts do not connect to this switch, traffic from the Payroll department must use Switch 2 to reach the rest of the network and Internet through the uplink port. For that reason, Switch 2 must be aware of VLAN 400 so that traffic is not rejected by the trunk port.

2. Configure ports 2-10 as access ports and add VLAN 200 to the ports.
3. Configure ports 11–30 as access ports and add VLAN 100 to the ports.
4. Configure LAG 1 as a general port so that it can be a member of multiple untagged VLANs and add VLAN 100 and VLAN 200 to the LAG.
5. Configure port 1 and LAG 2 trunk ports and add VLAN 100, VLAN 200, VLAN 300, and VLAN 400 to the port and LAG.
6. Configure the MAC-based VLAN information.
7. If desired, copy the running configuration to the startup configuration.
8. View VLAN information for the switch and ports.

22

Spanning Tree Protocol

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure the Spanning Tree Protocol (STP) settings on the switch.

The topics covered in this chapter include:

- STP Overview
- RSTP-PV
- Default STP Values
- Configuring Spanning Tree (Web)
- Configuring Spanning Tree (CLI)
- STP Configuration Examples

STP Overview

STP is a layer-2 protocol that provides a tree topology for switches on a bridged LAN. STP allows a network to have redundant paths without the risk of network loops. STP uses the spanning tree algorithm to provide a single path between end stations on a network.

Dell Networking N-Series switches support Classic STP, Multiple STP, and Rapid STP over point-to-point full-duplex links. Half-duplex associated states are not supported in Dell Networking spanning-tree. Dell Networking spanning tree presumes that all links are full-duplex and acts accordingly.

What Are Classic STP, Multiple STP, and Rapid STP?

Classic STP provides a single path between end stations, avoiding and eliminating loops.

Multiple Spanning Tree Protocol (MSTP) supports multiple instances of Spanning Tree to efficiently channel VLAN traffic over different interfaces. Each instance of the Spanning Tree behaves in the manner specified in IEEE 802.1w, Rapid Spanning Tree (RSTP), with slight modifications in the working but not the end effect (chief among the effects, is the rapid transitioning of the port to Forwarding). The difference between RSTP and the traditional STP (IEEE 802.1d) is the ability to recognize full-duplex connectivity and ports which are connected to end stations, resulting in rapid transitioning of the port to the Forwarding state and the suppression of Topology Change Notifications.

MSTP is compatible with both RSTP and STP. It behaves appropriately when connected to STP and RSTP bridges. A MSTP bridge can be configured to behave entirely as a RSTP bridge or a STP bridge.

How Does STP Work?

The switches (bridges) that participate in the spanning tree elect a switch to be the root bridge for the spanning tree. The root bridge is the switch with the lowest bridge ID, which is computed from the unique identifier of the bridge and its configurable priority number. When two switches have an equal bridge ID value, the switch with the lowest MAC address is the root bridge.

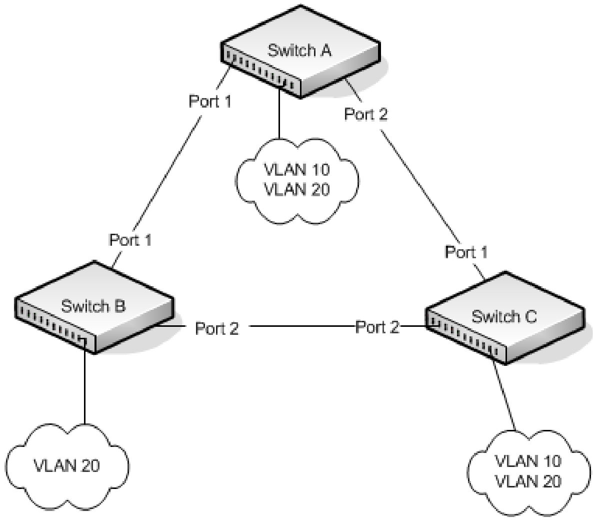
After the root bridge is elected, each switch finds the lowest-cost path to the root bridge. The port that connects the switch to the lowest-cost path is the root port on the switch. The switches in the spanning tree also determine which ports have the lowest-path cost for each segment. These ports are the designated ports. Only the root ports and designated ports are placed in a forwarding state to send and receive traffic. All other ports are put into a blocked state to prevent redundant paths that might cause loops. Both internal and external path costs can be configured. For STP, RSTP, and the MSTP CIST, only the external path costs are utilized in the lowest path cost calculation. The internal path cost is used by the MST instances.

To determine the root path costs and maintain topology information, switches that participate in the spanning tree use Bridge Protocol Data Units (BPDUs) to exchange information.

How Does MSTP Operate in the Network?

In the following diagram of a small 802.1d bridged network, STP is necessary to create an environment with full connectivity and without loops.

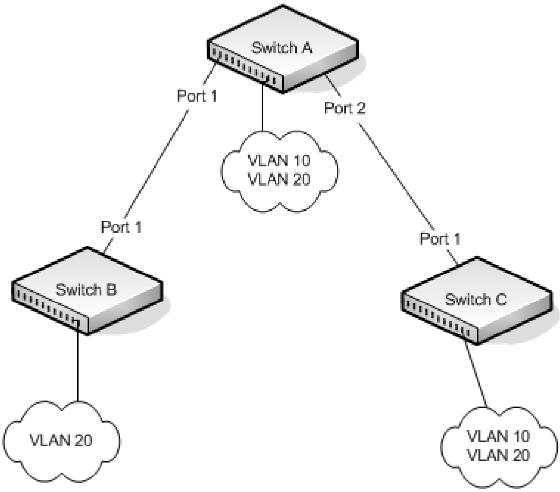
Figure 22-1. Small Bridged Network



Assume that Switch A is elected to be the Root Bridge, and Port 1 on Switch B and Switch C are calculated to be the root ports for those bridges, Port 2 on Switch B and Switch C would be placed into the Blocking state. This creates a loop-free topology. End stations in VLAN 10 can talk to other devices in VLAN 10, and end stations in VLAN 20 have a single path to communicate with other VLAN 20 devices.

Figure 22-2 shows the logical single STP network topology.

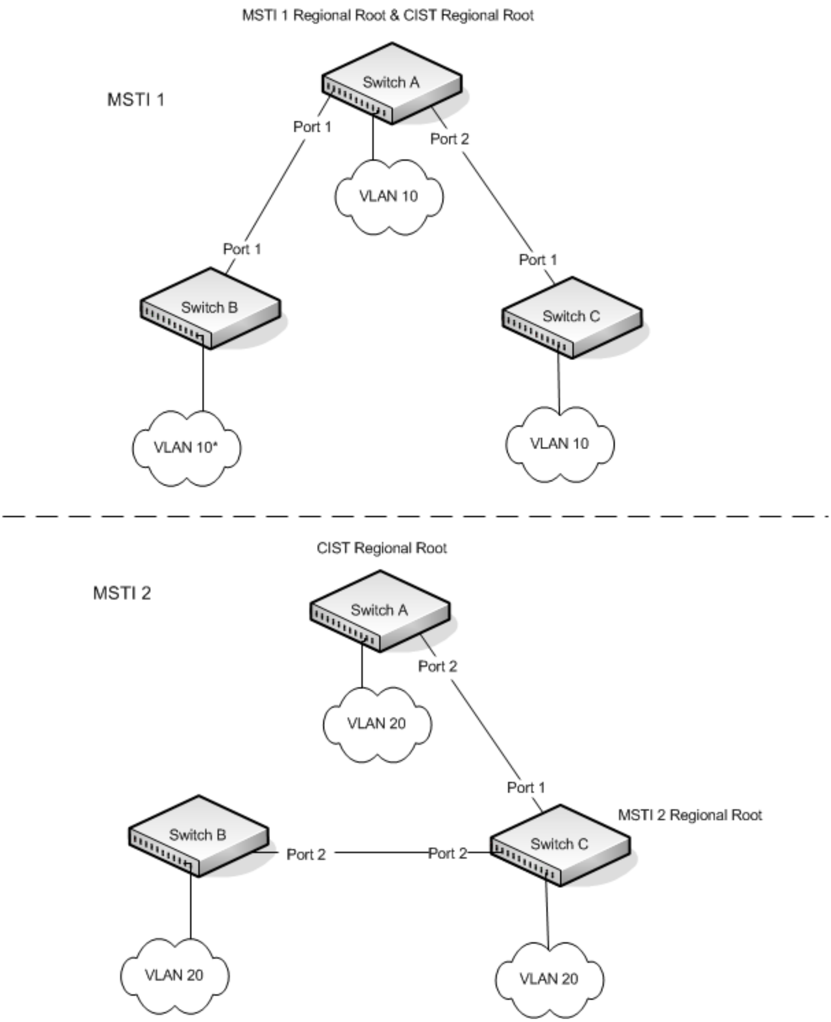
Figure 22-2. Single STP Topology



For VLAN 10 this single STP topology is fine and presents no limitations or inefficiencies. On the other hand, VLAN 20's traffic pattern is inefficient. All frames from Switch B will have to traverse a path through Switch A before arriving at Switch C. If the Port 2 on Switch B and Switch C could be used, these inefficiencies could be eliminated. MSTP does just that, by allowing the configuration of MSTIs based upon a VLAN or groups of VLANs. In this simple case, VLAN 10 could be associated with Multiple Spanning Tree Instance (MSTI)1 with an active topology similar to Figure 22-2 and VLAN 20 could be associated with MSTI 2 where Port 1 on both Switch A and Switch B begin discarding and all others forwarding. This simple modification creates an active topology with a better distribution of network traffic and an increase in available bandwidth.

The logical representation of the MSTP environment for these three switches is shown in Figure 22-3.

Figure 22-3. Logical MSTP Environment



In order for MSTP to correctly establish the different MSTIs as above, some additional changes are required. For example, the configuration would have to be the same on each and every bridge. That means that Switch B would have to add VLAN 10 to its list of supported VLANs (shown in Figure 22-3 with a *). This is necessary with MSTP to allow the formation of Regions made up of all switches that exchange the same MST Configuration Identifier. It is within only these MST Regions that multiple instances can exist. It will also allow the election of Regional Root Bridges for each instance. One common and internal spanning tree (CIST) Regional Root for the CIST and an MSTI Regional Root Bridge per instance will enable the possibility of alternate paths through each Region. Above Switch A is elected as both the MSTI 1 Regional Root and the CIST Regional Root Bridge, and after adjusting the Bridge Priority on Switch C in MSTI 2, it would be elected as the MSTI 2 Regional Root.

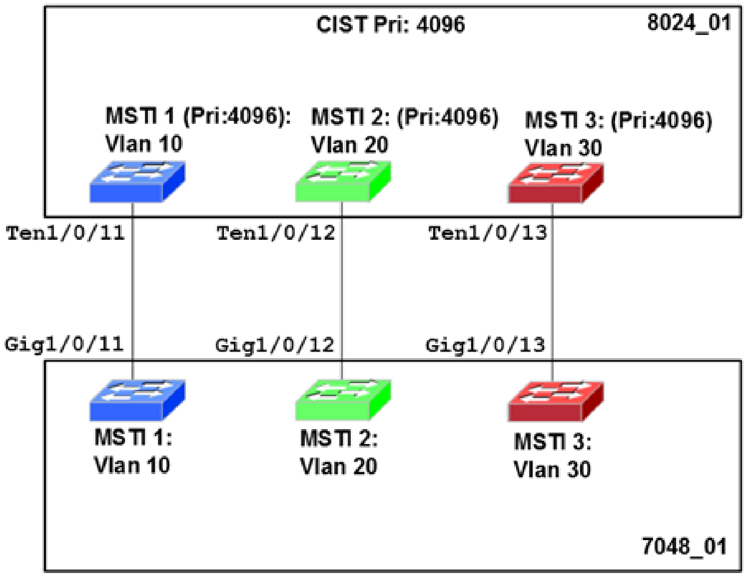
To further illustrate the full connectivity in an MSTP active topology, the following rules apply:

1. Each Bridge or LAN is in only one Region.
2. Every frame is associated with only one VID.
3. Frames are allocated either to the IST or MSTI within any given Region.
4. The internal spanning tree (IST) and each MSTI provides full and simple connectivity between all LANs and Bridges in a Region.
5. All Bridges within a Region reach a consistent agreement as to which ports interconnect that Region to a different Region and label those as Boundary Ports.
6. At the Boundary Ports, frames allocated to the CIST or MSTIs are forwarded or not forwarded alike.
7. The CIST provides full and simple connectivity between all LANs and Bridges in the network.

MSTP with Multiple Forwarding Paths

Consider the physical topology shown in Figure 22-4. It might be assumed that MSTI 2 and MSTI 3 would follow the most direct path for VLANs 20 and 30. However, using the default path costs, this is not the case. MSTI operates without considering the VLAN membership of the ports. This results in unexpected behavior if the active topology of an MSTI depends on a port that is not a member of the VLAN assigned to the MSTI and the port is selected as root port. In this configuration, port TE 1/0/11 is selected as the root port and ports TE1/0/12 and TE1/0/13 are blocked. To resolve the issue, set the port path cost of the directly connected links to allow the MSTIs to connect directly.

Figure 22-4. MSTP with Multiple Forwarding Paths



MSTP and VLAN IDs

MSTP allows VLAN 4094 to be configured in the MD5 digest of an MSTI region for compatibility purposes. However, the switch reserves VLAN 4094 internally for use in stacking and will drop received packets tagged with VLAN 4094.

What are the Optional STP Features?

The Dell Networking N-Series switches support the following optional STP features:

- BPDU flooding
- PortFast
- BPDU filtering
- Root guard
- Loop guard
- BPDU protection

BPDU Flooding

The BPDU flooding feature determines the behavior of the switch when it receives a BPDU on a port that is disabled for spanning tree. If BPDU flooding is configured, the switch will flood the received BPDU to all the ports on the switch which are similarly disabled for spanning tree.

Port Fast

The PortFast feature reduces the STP convergence time by allowing edge ports that are connected to end devices (such as a desktop computer, printer, or file server) to transition to the forwarding state without going through the listening and learning states.

BPDU Filtering

Ports that have the PortFast feature enabled continue to transmit BPDUs. The BPDU filtering feature prevents PortFast-enabled ports from sending BPDUs.

If BPDU filtering is configured globally on the switch, the feature is automatically enabled on all operational PortFast-enabled ports. These ports are typically connected to hosts that drop BPDUs. However, if an operational edge port receives a BPDU, the BPDU filtering feature disables PortFast and allows the port to participate in the spanning tree calculation.

Enabling BPDU filtering on a specific port prevents the port from sending BPDUs and allows the port to drop any BPDUs it receives.

Root Guard

Root guard is another way of controlling the spanning-tree topology other than setting the bridge priority or path costs. Root guard ensures that a port does not become a root port or a blocked port. When a switch is elected as the root bridge, all ports are assigned roles as designated ports unless two or more ports of the root bridge are connected in a loop. If the switch receives a superior STP BPDU on a root-guard enabled port, the root guard feature moves the port to a root-inconsistent spanning-tree state. No traffic is forwarded across the port, but it continues to receive BPDUs, discards received traffic, and is included in the active topology. Essentially, this is equivalent to the IEEE 802.1D listening state. By not transitioning the port on which the superior BPDU has been received to the forwarding state (designated role), root guard helps maintain the existing spanning-tree topology.

When the STP mode is configured as MSTP, the port may be a designated port in one MSTI and an alternate port in the CIST, etc. Root guard is a per port (not a per port instance command) configuration, so all the MSTP instances this port participates in should not be expected to take on a root role.

Loop Guard

Loop guard protects a network from forwarding loops induced by BPDU packet loss. The reasons for failing to receive packets are numerous, including heavy traffic, software problems, incorrect configuration, and unidirectional link failure. When a non-designated port no longer receives BPDUs, the spanning tree algorithm considers the link to be loop free and transitions the link from blocking to forwarding. Once in the forwarding state, the link may create a loop in the network.

Enabling loop guard prevents such accidental loops. When a port is no longer receiving BPDUs and the max age timer expires, the port is moved to a *loop-inconsistent blocking state*. In the loop-inconsistent blocking state, traffic is not forwarded so the port behaves as if it is in the blocking state; that is, it discards received traffic, does not learn MAC addresses, and is not part of the active topology. The port will remain in this state until it receives a BPDU. It will then transition through the normal spanning tree states based on the information in the received BPDU.

NOTE: Loop Guard should be configured only on non-designated ports. These include ports in alternate or backup roles. Root ports and designated ports should not have loop guard enabled so that they can forward traffic.

BPDU Protection

When the switch is used as an access-layer device, most ports function as edge ports that connect to a device such as a desktop computer or file server. The port has a single, direct connection and is configured as an edge port to implement the fast transition to a forwarding state. When the port receives a BPDU packet, the system sets it to non-edge port and recalculates the spanning tree, which causes network topology flapping. In normal cases, these ports do not receive any BPDU packets. However, someone may forge BPDU to maliciously attack the switch and cause network flapping.

BPDU protection can be enabled in RSTP to prevent such attacks. When BPDU protection is enabled, the switch disables an access port that has received a BPDU and notifies the network manager about it.

RSTP-PV

Dell Networking N-Series switches support both Rapid Spanning Tree Per VLAN (RSTP-PV) and Spanning Tree Per VLAN (STP-PV) with a high degree of interoperability with other vendor implementations, such as Cisco's PVST+ and RPVST+. RSTP-PV is the IEEE 802.1w (RSTP) standard implemented per VLAN. A single instance of rapid spanning tree (RSTP) runs on each configured VLAN. Each RSTP instance on a VLAN has a root switch. The RSTP-PV protocol state machine, port roles, port states, and timers are similar to those defined for RSTP. RSTP-PV embeds the DRC and IndirectLink Fast Rapid Convergence (IRC) features, which cannot be disabled.

STP-PV is the IEEE 802.1s (STP) standard implemented per VLAN. The STP-PV-related state machine, roles, and timers are similar to those defined for STP. STP-PV does not have the DirectLink Rapid Convergence (DRC) or IndirectLink Rapid Convergence (IRC) features enabled by default. These features can be enabled by the switch administrator.

The switch spanning tree configuration is global in nature. Enabling RSTP-PV disables other spanning tree modes on the switch. The switch cannot operate with some ports configured to operate in standard spanning tree mode and others to operate in RSTP-PV mode. However, RSTP-PV has fallback modes for compatibility with standards-based versions of spanning tree.

Access Ports—For an access port, normal IEEE BPDUs will be received and sent, though STP-PV or RSTP-PV is enabled on the switch. BPDUs received on the access port will be associated with the CST instance.

Trunk Ports—If the native VLAN on an IEEE 802.1Q trunk is VLAN 1:

- VLAN 1 STP BPDUs are sent to the IEEE STP MAC address (0180.c200.0000), untagged.
- VLAN 1 STP BPDUs are also sent to the SSTP MAC address, untagged.
- Non-VLAN 1 STP BPDUs are sent to the SSTP MAC address (also called the Shared Spanning Tree Protocol [SSTP] MAC address, 0100.0ccc.cccd), tagged with a corresponding IEEE 802.1Q VLAN tag.

If the native VLAN on an IEEE 802.1Q trunk is not VLAN 1:

- VLAN 1 STP BPDUs are sent to the SSTP MAC address, tagged with a corresponding IEEE 802.1Q VLAN tag.
- VLAN 1 STP BPDUs are also sent to the IEEE STP MAC address on the Native VLAN of the IEEE 802.1Q trunk, untagged.
- Non-VLAN 1 STP BPDUs are sent to the SSTP MAC address, tagged with a corresponding IEEE 802.1Q VLAN tag.

DirectLink Rapid Convergence

The DirectLink Rapid Convergence (DRC) feature is designed for an access-layer switch that has redundant blocked uplinks. It operates on ports blocked by spanning tree. DRC can be configured for the entire switch; it cannot be enabled for individual VLANs.

The DRC feature is based on the concept of an uplink group. An uplink group consists of all the ports that provide a path to the root bridge (the root port and any blocked ports). If the root port fails, the blocked port with next lowest cost from the uplink group is selected and immediately put in the forwarding state without going through the standard spanning tree listening and learning states.

To accelerate convergence time once DRC has switched over to a new root port, STP-PV transmits dummy packets out the new root port, with the source MAC addresses taken from its forwarding table. The destination address is an SSTP MAC address that ensures that the packet is flooded on the whole network. The packets update the forwarding tables on the other upstream switches. The rate at which the dummy multicasts are sent can be configured by the administrator. RSTP-PV has a different mechanism adopted from IEEE 802.1w that handles the update of the forwarding database and the fast transition to a new uplink. DRC can be enabled on RSTP-PV enabled switches but has no effect.

DRC is disabled when the administrator modifies the spanning-tree priority of a VLAN and is re-enabled only when the default priority is restored.

DRC and Link Up Events

In the event of failure of the primary uplink, a replacement uplink is immediately selected from the uplink group and put into the forwarding state. If another port is enabled that, in accordance with STP rules, should become the primary uplink (root port), the switch delays migrating to the new port for twice the forwarding delay. The purpose of this delay is two-fold:

- Stability—If the primary uplink is flapping, reenabling the link immediately can introduce additional instability into the network.
- Reduced Traffic Loss—DRC moves a port into the forwarding state as soon as it is up, but the connected port obeys the usual STP rules; i.e. it goes through the listening and learning stages, which take 15 seconds each by default. Delaying the switchover allows the connected port to go to through the listening and learning states while the switch is still transmitting packets on the original uplink.

The optimal behavior is to keep the current uplink active and hold the new port in the blocked state for twice the forwarding delay.

IndirectLink Rapid Convergence Feature

To handle indirect link failure, the STP standard requires that a switch passively wait for "max_age" seconds once a topology change has been detected. IndirectLink Rapid Convergence (IRC) handles these failures in two phases:

- Rapid detection of an indirect link failure. Tracking the inferior BPDUs that a designated bridge detects when it transmits a direct link failure indicates that a failure has occurred elsewhere in the network.
- Performing an immediate check if the BPDU information stored on a port is still valid. This is implemented with a new protocol data unit (PDU) and the Root Link Query message (RLQ).

Receiving an inferior BPDU on a port from the designated bridge indicates that one of the following has occurred on the designated bridge:

- The path to the root has been lost and the switch starts to advertise a root with a numerically higher bridge ID (worse root) than the local switch.
- The path cost to the root has increased above the path cost of the local switch.

IEEE 802.1s behavior is to ignore inferior BPDUs. IRC retains the inferior BPDUs sent by the designated bridge and processes them to determine if a failure has occurred on the path to the root. In this case, it must age-out at least one port. This process occurs only in the case that a bridge in the network detects a direct link failure.

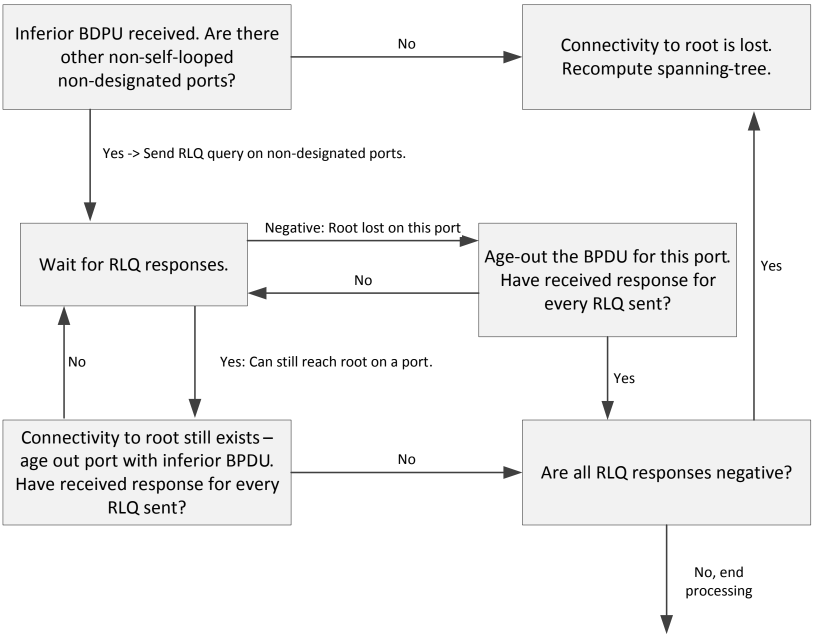
The switch tracks inferior BPDUs sent by the designated bridge only, since this is the BPDU that is stored for the port. If, for instance, a newly inserted bridge starts to send inferior BPDUs, it does not start the IRC feature.

Similar to DRC, the IEEE 802.1w standard incorporated the IRDC feature. RSTP-PV enabled switches allow IRC to be enabled or disabled, but ignore the setting as the RSTP-PV state machines already implement IRC.

Reacting to Indirect Link Failures

When an inferior BPDU is received on a non-designated port, phase 2 of IRC processing starts. An RLQ PDU is transmitted on all non-designated ports except the port where the inferior BPDU was received and self-looped ports. This action is intended to verify that the switch can still receive from the root on ports that should have a path to the root. The port where the switch received the inferior BPDU is excluded because it already failed; self-looped and designated ports are eliminated as they do not have a path to the root.

Figure 22-5. IRC Flow



Upon receiving a negative RLQ response on a port, the port has lost connection to the root and the switch ages-out its BPDU. If all other non-designated ports received a negative answer, the switch has lost the root and restarts the STP calculation.

If the response confirms the switch can still access the root bridge via a particular port, it immediately ages-out the port on which the inferior BPDU was received.

If the switch only received responses with a root different from the original root, it has lost the root port and restarts the STP calculation immediately.

Interoperability Between STP-PV and RSTP-PV Modes

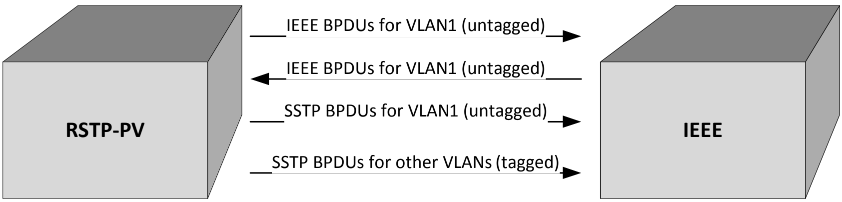
STP-PV is derived from 802.1D and RSTP-PV is derived from 802.1w. The fallback mechanism is the same as between a standard 802.1D switch and a standard 802.1w switch. When a lower protocol version BPDU is received on a switch that runs a higher protocol version, the latter falls back to the lower version after its migration delay timer expires.

For example, an RSTP-PV switch, when connected to STP-PV switch, falls back to the STP-PV protocol after the migration delay timer expires.

Interoperability With IEEE Spanning Tree Protocols

When a switch configured with RSTP-PV receives IEEE standard RSTP BPDUs on a port, it responds with two versions of BPDUs on the port: SSTP formatted BPDUs and IEEE standard STP BPDUs. The IEEE standard BPDUs are processed by the peer switch running MSTP/RSTP, and the SSTP format BPDUs are flooded across the MSTP/RSTP domain.

Figure 22-6. RSTP-PV and IEEE Spanning Tree Interoperability



Common Spanning Tree

There are differences between the ways that MSTP and RSTP-PV map spanning tree instances to VLANs: RSTP-PV creates a spanning tree instance for each VLAN, and MSTP maps one or more VLANs to each MST instance. Where an RSTP-PV region is connected to an MSTP region, the set of RSTP-PV instances does not generally match the set of MST instances. Therefore, the RSTP-PV region and the MSTP region communicate with each other on a single common spanning tree instance.

For the MSTP region, the MSTP instance communicates to the RSTP-PV region using the CIST. For the RSTP-PV region, switches use the VLAN 1 RSTP-PV instance as the common spanning tree. On the link between the

RSTP-PV region and the MSTP region, the RSTP-PV switch sends VLAN1 BPDUs in IEEE standard format, so they can be interpreted by the MSTP peers. Similarly, the RSTP-PV switch processes incoming MSTP BPDUs as though they were BPDUs for the VLAN 1 RSTP-PV instance.

If the RSTP-PV switch ports connected to the MSTP switches are configured with a native VLAN, the RSTP-PV switches are able to detect IEEE standard format BPDUs arriving from peer switches, incorporate them into the common spanning tree that operates in the native VLAN (VLAN 1), and transmit untagged STP or RSTP packets to the STP/RSTP peers, in addition to the SSTP format BPDUs.

SSTP BPDUs Flooding Across MST (CST) Regions

In addition to the IEEE standard RSTP or STP BPDUs that the RSTP-PV switch sends to the MSTP (or RSTP or STP) region, the switch sends SSTP format BPDUs for VLAN 1 untagged. The MSTP switch does not interpret the SSTP BPDUs as standard BPDUs because they do not use the standard destination MAC address, so it makes no spanning tree decisions based on them. Instead, it floods the SSTP BPDUs over all ports in the corresponding VLAN. These SSTP BPDUs may be multicast over the MSTP region to other RSTP-PV switches, which use them to maintain the VLAN 1 spanning tree topology across the MSTP (non-RSTP-PV) switches.

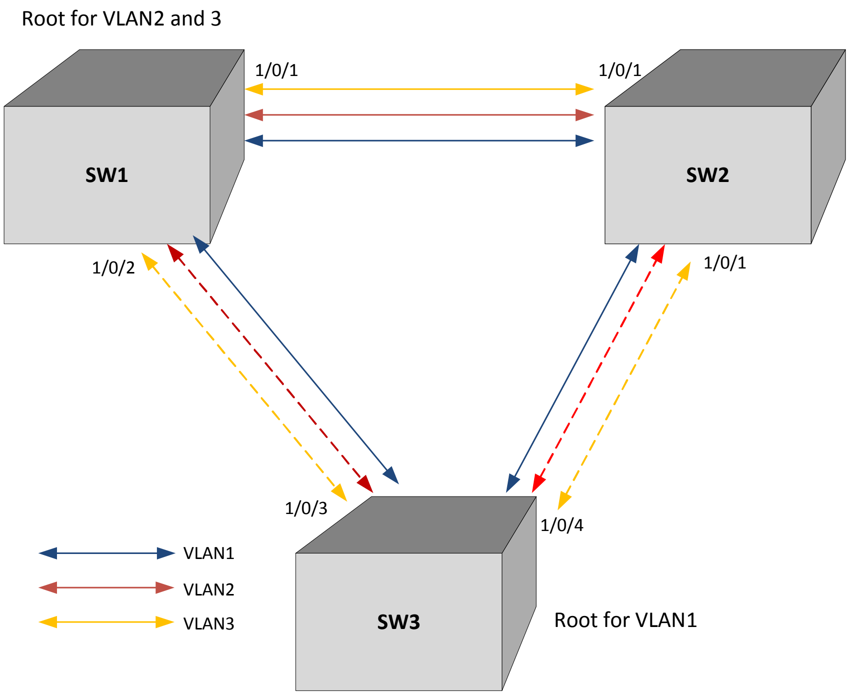
The RSTP-PV switches also send SSTP format BPDUs for the other (non-VLAN 1) RSTP-PV instances into the MSTP region, tagged with the VID of their associated VLANs. These SSTP packets are also be multicast by the switches in the MSTP region, and will reach any other RSTP-PV regions connected to the MSTP region. The switches in the remote RSTP-PV regions receive and process them as normal RSTP-PV BPDUs. Thus, RSTP-PV instances are transparently expanded across the MSTP region and their spanning trees span the MSTP region. For RSTP-PV, the MSTP region is treated as a single hub.

Interoperability with RSTP

In Figure 22-7:

- SW1 and SW2 are Dell Networking N-Series switches running RSTP-PV with default bridge priority 32768.
- SW3 is a Dell Networking N-Series switch running RSTP with default bridge priority 32768.

Figure 22-7. RSTP-PV and RSTP Interoperability



SW3 sends IEEE STP BPDUs to the IEEE multicast MAC address as untagged frames. These BPDUs are processed by the VLAN 1 STP instance on the RSTP-PV switch as part of the VLAN 1 STP instance.

The RSTP-PV side sends IEEE STP BPDUs corresponding to the VLAN 1 STP to the IEEE MAC address as untagged frames across the link. At the same time, SSTP BPDUs are sent as untagged frames. IEEE switches simply flood the SSTP BPDUs throughout VLAN 1. This facilitates RSTP-PV connectivity in case there are other RSTP-PV switches connected to the IEEE STP domain.

For non-native VLANs (VLANs 2–4093), the RSTP-PV switch sends SSTP BPDUs, tagged with their VLAN number. The VLAN STP instances are multicast across the RSTP region, as if it were a hub switch.

The VLAN 1 STP instance of SW1 and SW2 are joined with the STP instance running in SW3. VLANs 2 and 3 consider the path across SW3 as another segment linking SW1 and SW2, and their SSTP information is multicast across SW3.

The bridge priority of SW1 and SW2 for VLAN1 instance is 32769 (bridge priority + VLAN identifier).

The bridge priority of SW3 is 32768, per the IEEE 802.w standard.

SW3 is selected as Root Bridge for the VLAN1 instance that is CST, and SW1 is selected as Root Bridge for VLAN2 and VLAN3 (based on the low MAC address of SW1).

Interoperability with MSTP

RSTP-PV runs an individual RSTP instance for each VLAN. MSTP maps VLANs to MSTIs, so one-to-one mapping between VLAN and STP instance is not possible.

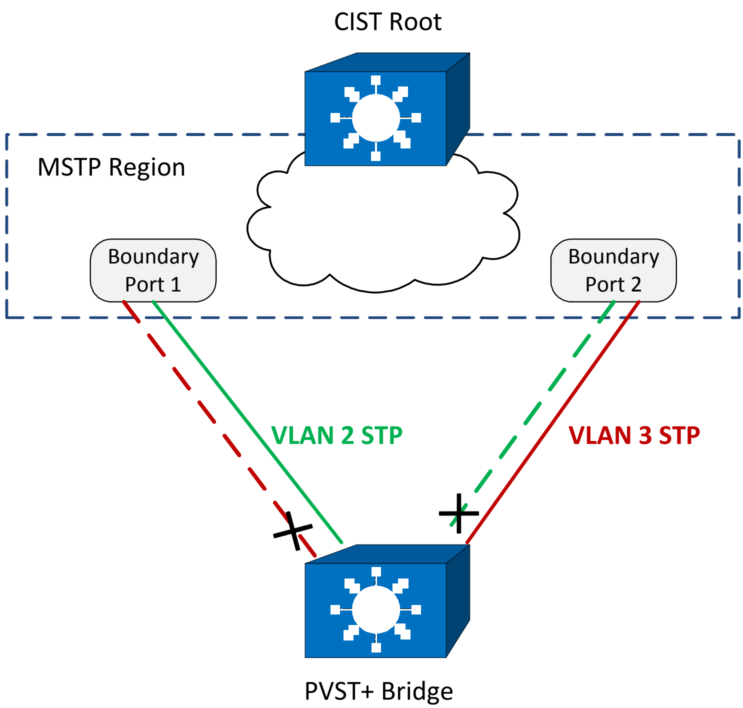
MSTP runs multiple MSTIs inside a region and maps them to the CIST on the border ports. The interoperability model must ensure that internal MSTIs are aware of changes to any of the RSTP-PV trees. Therefore, the simplest way to ensure the correct behavior is to join ALL RSTP-PV trees to the CST.

Connecting RSTP-PV trees to the CST ensures that changes in any of the RSTP-PV STP instances will affect the CST and all MSTIs. This approach ensures that no changes go unnoticed and no black holes occur in a single VLAN. As with IEEE STP, every tree in the RSTP-PV domain views the MSTP regions as virtual bridges with multiple boundary ports. A topology change in any of the RSTP-PV spanning trees will affect the CST and propagate through every MSTI instance in all MSTP regions. This behavior, consequently, makes the MSTP topology less stable.

The MSTP implementation simulates RSTP-PV by replicating CIST BPDUs on the link facing the RSTP-PV domain and sending those BPDUs on ALL VLANs active on the trunk. The MSTP switch processes IEEE STP VLAN 1 BPDUs received from the RSTP-PV domain using the CIST instance. The RSTP-PV+ domain interprets the MSTP domain as an RSTP-PV bridge with all per-VLAN instances claiming the CIST Root as the root of their individual spanning tree. For the common STP Root elected between MSTP and RSTP-PV, two options are possible:

- The MSTP domain contains the root bridge for ALL VLANs. This implies that the CIST Root Bridge ID is configured to be better than any RSTP-PV STP root Bridge ID. If there is only one MSTP region connected to the RSTP-PV domain, then all boundary ports on the virtual-bridge will be unblocked and used by RSTP-PV. This is the only supported topology, as the administrator can manipulate uplink costs on the RSTP-PV side and obtain optimal traffic engineering results. In Figure 22-8, VLANs 2 and 3 have their STP costs configured to select different uplinks connected to the MSTP region's boundary ports. Since the CIST Root is inside the MSTP region, both boundary ports are non-blocking designated and the load balancing scheme operates as expected.

Figure 22-8. MSTP and RSTP-PV Interoperability



- The alternative is that the RSTP-PV domain contains the root bridges for ALL VLANs. This is only true if all RSTP-PV root bridges' Bridge IDs for all VLANs are better than the MSTP CIST Root Bridge ID. This is not a supported topology, because all MSTIs map to CIST on the border link, and it is not possible to load-balance the MSTIs as they enter the RSTP-PV domain.

The Dell Networking RSTP-PV implementation does not support the second option. The MSTP domain must contain the bridge with the best Bridge ID to ensure that the CIST Root is also the root for all RSTP-PV trees. In any other case, the MSTP border switch will place the ports that receive superior BPDUs from the RSTP-PV region in the root-inconsistent state. To resolve this issue, ensure that the RSTP-PV domain does not have any bridges with Bridge IDs better than the CIST Root Bridge ID.

Native VLAN Inconsistent State

This occurs if a trunk port receives an untagged SSTP BPDU with a VLAN type, length, value (TLV) that does not match the VLAN where the BPDU was received. In this case, the port transitions to the blocked state.

Configuration Examples

See "RSTP-PV Access Switch Configuration Example " on page 820.

Default STP Values

Spanning tree is globally enabled on the switch and on all ports and LAGs. Table 22-1 summarizes the default values for STP.

Table 22-1. STP Defaults

Parameter	Default Value
Enable state	Enabled (globally and on all ports)
Spanning tree mode	RSTP (Classic STP, STP-PV, RSTP-PV and MSTP are disabled)
Switch priority	32768
BPDU flooding	Disabled
PortFast mode	Disabled
PortFast BPDU filter	Disabled
Loop guard	Disabled
BPDU protection	Disabled
Spanning tree port priority	128
Maximum-aging time	20 seconds
Forward-delay time	15 seconds
Maximum hops	20
Spanning tree transmit hold count	6
MSTP region name	MAC address of switch
MSTP included VLANs	1

Configuring Spanning Tree (Web)

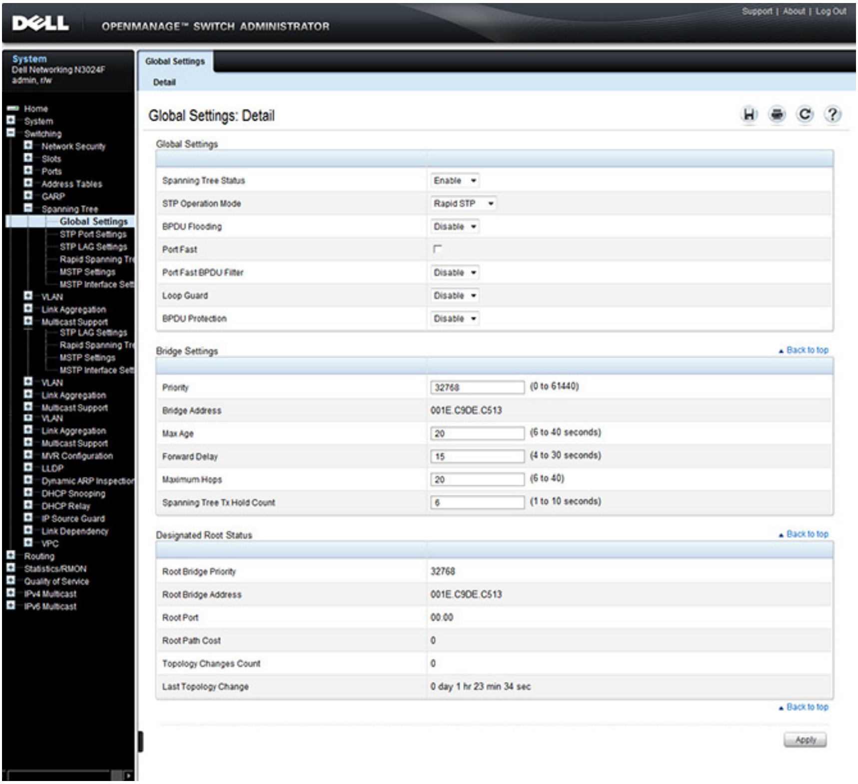
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring STP settings on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click  at the top of the page.

STP Global Settings

The **STP Global Settings** page contains fields for enabling STP on the switch.

To display the **STP Global Settings** page, click **Switching** → **Spanning Tree** → **Global Settings** in the navigation panel.

Figure 22-9. Spanning Tree Global Settings

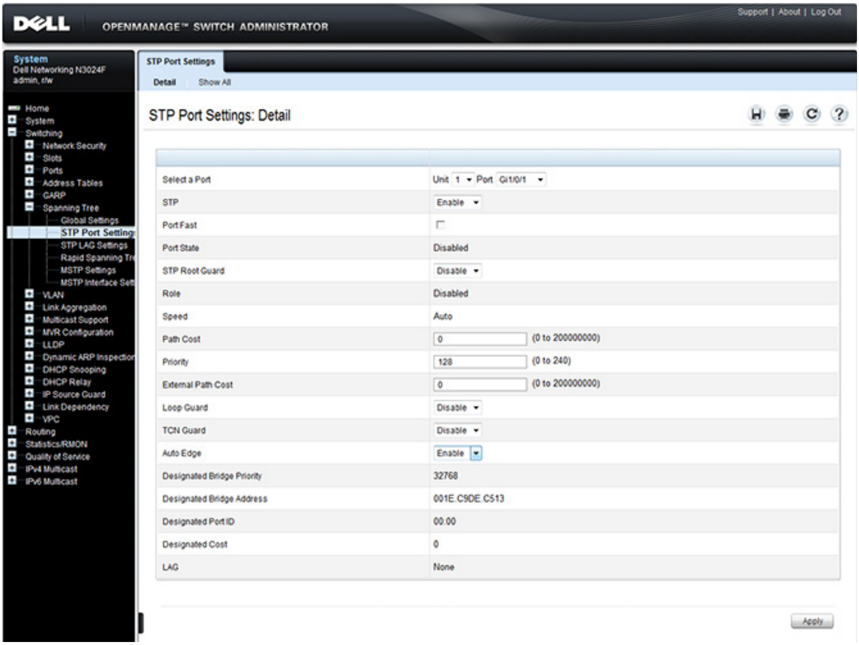


STP Port Settings

Use the **STP Port Settings** page to assign STP properties to individual ports.

To display the **STP Port Settings** page, click **Switching** → **Spanning Tree** → **STP Port Settings** in the navigation panel.

Figure 22-10. STP Port Settings

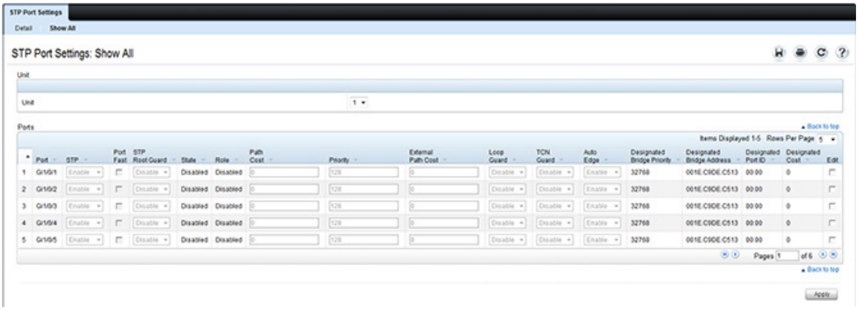


Configuring STP Settings for Multiple Ports

To configure STP settings for multiple ports:

1. Open the **STP Port Settings** page.
2. Click **Show All** to display the **STP Port Table**.

Figure 22-11. Configure STP Port Settings



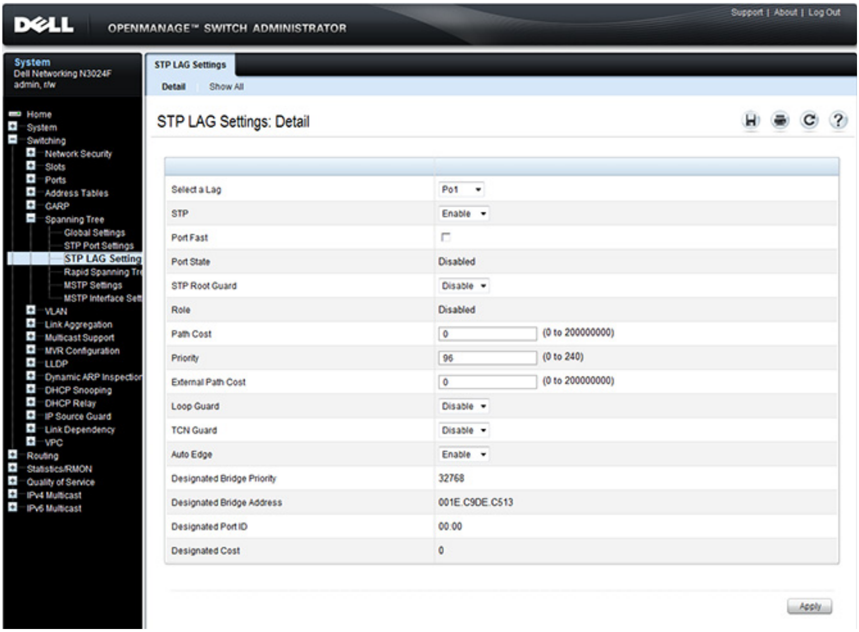
3. For each port to configure, select the check box in the **Edit** column in the row associated with the port.
4. Select the desired settings.
5. Click **Apply**.

STP LAG Settings

Use the **STP LAG Settings** page to assign STP aggregating ports parameters.

To display the **STP LAG Settings** page, click **Switching** → **Spanning Tree** → **STP LAG Settings** in the navigation panel.

Figure 22-12. STP LAG Settings

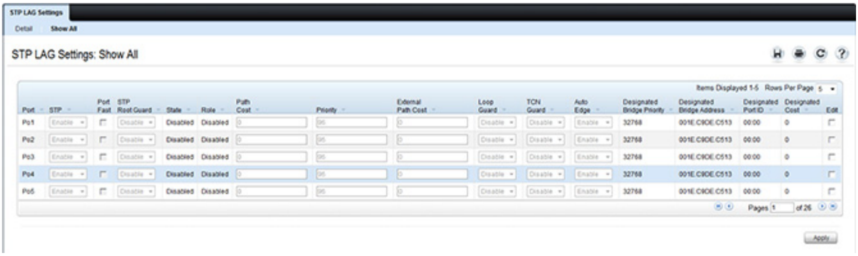


Configuring STP Settings for Multiple LAGs

To configure STP settings on multiple LAGS:

1. Open the **STP LAG Settings** page.
2. Click **Show All** to display the **STP LAG Table**.

Figure 22-13. Configure STP LAG Settings



3 For each LAG to configure, select the check box in the **Edit** column in the row associated with the LAG.

4 Select the desired settings.

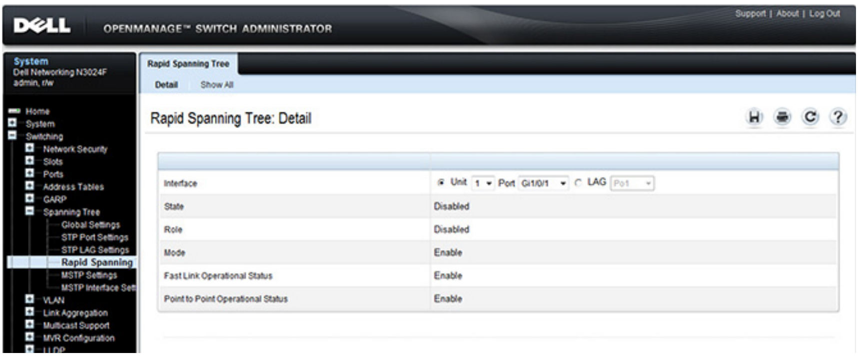
5 Click **Apply**.

Rapid Spanning Tree

Rapid Spanning Tree Protocol (RSTP) detects and uses network topologies that allow a faster convergence of the spanning tree without creating forwarding loops.

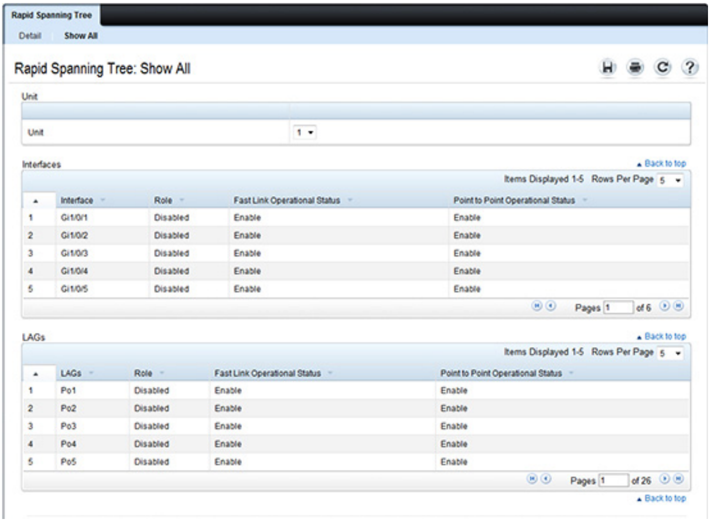
To display the **Rapid Spanning Tree** page, click **Switching** → **Spanning Tree** → **Rapid Spanning Tree** in the navigation panel.

Figure 22-14. Rapid Spanning Tree



To view RSTP Settings for all interfaces, click the **Show All** link. The **Rapid Spanning Tree Table** displays.

Figure 22-15. RSTP Settings

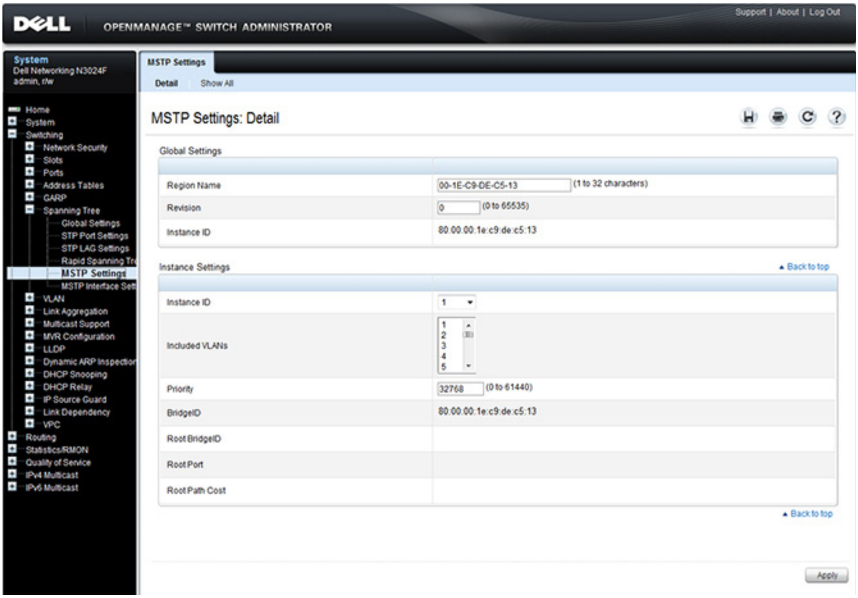


MSTP Settings

The Multiple Spanning Tree Protocol (MSTP) supports multiple instances of Spanning Tree to efficiently channel VLAN traffic over different interfaces. MSTP is compatible with both RSTP and STP; a MSTP bridge can be configured to behave entirely as a RSTP bridge or a STP bridge.

To display the **MSTP Settings** page, click **Switching** → **Spanning Tree** → **MSTP Settings** in the navigation panel.

Figure 22-16. MSTP Settings



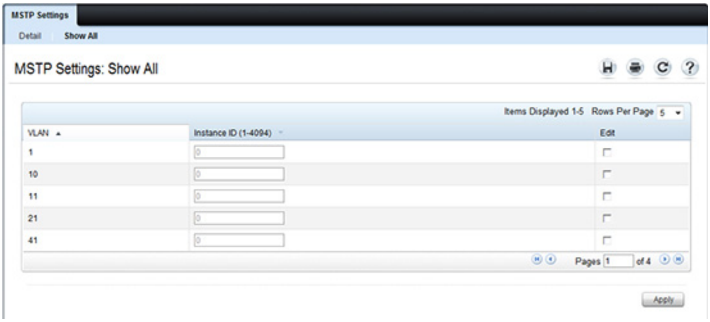
Viewing and Modifying the Instance ID for Multiple VLANs

To configure MSTP settings for multiple VLANS:

1 Open the **MSTP Settings** page.

2 Click **Show All** to display the **MSTP Settings Table**.

Figure 22-17. Configure MSTP Settings



3 For each Instance ID to modify, select the check box in the **Edit** column in the row associated with the VLAN.

4 Update the **Instance ID** settings for the selected VLANs.

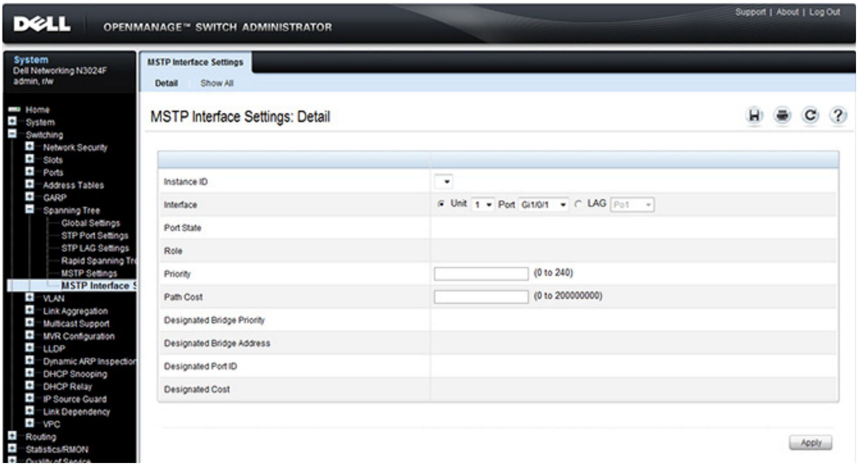
5 Click **Apply**.

MSTP Interface Settings

Use the **MSTP Interface Settings** page to assign MSTP settings to specific interfaces.

To display the **MSTP Interface Settings** page, click **Switching** → **Spanning Tree** → **MSTP Interface Settings** in the navigation panel.

Figure 22-18. MSTP Interface Settings



Configuring MSTP Settings for Multiple Interfaces

To configure MSTP settings for multiple interfaces:

1. Open the **MSTP Interface Settings** page.
2. Click **Show All** to display the **MSTP Interface Table**.
3. For each interface to configure, select the check box in the **Edit** column in the row associated with the interface.
4. Update the desired settings.
5. Click **Apply**.

Configuring Spanning Tree (CLI)

This section provides information about the commands used for configuring STP settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Global STP Bridge Settings

Beginning in Privileged EXEC mode, use the following commands to configure the global STP settings for the switch, such as the priority and timers.

Command	Purpose
configure	Enter global configuration mode.
spanning-tree	Enable spanning tree on the switch.
spanning tree mode {stp | rstp |mst}	Specify which spanning tree mode to use on the switch.
spanning-tree priority *priority*	Specify the priority of the bridge. (Range: 0–61440). The switch with the lowest priority value is elected as the root switch.
spanning-tree max-age *seconds*	Specify the switch maximum age time, which indicates the amount of time in seconds a bridge waits before implementing a topological change. Valid values are from (6 to 40) seconds.
spanning-tree forward-time *seconds*	Specify the switch forward delay time, which indicates the amount of time in seconds a bridge remains in a listening and learning state before forwarding packets. Valid values are from (4 to 30) seconds.
spanning-tree max-hops *hops*	Configure the maximum number of hops for the Spanning tree. Valid values are from (6 to 40).
spanning-tree transmit hold-count [*value*]	Set the maximum number of BPDUs that a bridge is allowed to send within a hello time window (2 seconds). The range for *value* is 1–10.
CTRL + Z	Exit to Privileged EXEC mode.

Command	Purpose
show spanning-tree [**detail**] [**active** | **blockedports**]	View information about spanning tree and the spanning tree configuration on the switch.

Configuring Optional STP Features

Beginning in Privileged EXEC mode, use the following commands to configure the optional STP features on the switch or on specific interfaces.

Command	Purpose
configure	Enter global configuration mode.
spanning-tree bpdu flooding	Allow the flooding of BPDUs received on non-spanning-tree ports to all other non-spanning-tree ports.
spanning-tree portfast	Enable PortFast on all switch ports.
spanning-tree portfast bpdufilter default	Prevent ports configured in PortFast mode from sending BPDUs.
spanning-tree loopguard default	Enable loop guard on all ports.
spanning-tree bpdu-protection	Enable BPDU protection on the switch.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3** or **port-channel 4**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12. The **range** keyword is also valid for LAGs (port-channels).
spanning-tree auto-portfast	Set the port to auto portfast mode. This enables the port to become a portfast port if it does not see any BPDUs for 3 seconds.
spanning-tree guard {**root** | **loop** | **none**}	Enable loop guard or root guard (or disable both) on the interface.

Command	Purpose
spanning-tree tcnguard	Prevent the port from propagating topology change notifications.
CTRL + Z	Exit to Privileged EXEC mode.
show spanning-tree summary	View various spanning tree settings and parameters for the switch.

Configuring STP Interface Settings

Beginning in Privileged EXEC mode, use the following commands to configure the STP settings for a specific interface.

Command	Purpose
configure	Enter global configuration mode.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3** or **port-channel 4**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12. The **range** keyword is also valid for LAGs (port-channels).
spanning-tree disable	Disable spanning tree on the port.
spanning-tree port-priority *priority*	Specify the priority of the port. (Range: 0–240). The priority value is used to determine which ports are put in the forwarding state and which ports are put in the blocking state. A port with a lower priority value is more likely to be put into a forwarding state.
spanning-tree cost *cost*	Specify the spanning tree path cost for the port. (Range: 0–200,000,000). The default cost is 0, which signifies that the cost is automatically calculated based on port speed.
CTRL + Z	Exit to Privileged EXEC mode.
show spanning-tree *interface*	View spanning tree configuration information for the specified port or LAG (port-channel).

Configuring MSTP Switch Settings

Beginning in Privileged EXEC mode, use the following commands to configure MSTP settings for the switch.

Command	Purpose
configure	Enter global configuration mode.
spanning-tree mst configuration	Enable configuring an MST region by entering the multiple spanning tree (MST) mode.
name *string*	Define the MST configuration name. This step is required to establish an MST domain.
revision *version*	Identify the MST configuration revision number.
instance *instance-id* **{add | remove} vlan** *vlan-range*	Map VLANs to an MST instance. • *instance-ID* — ID of the MST instance. (Range: 1-4094) • *vlan-range* — VLANs to be added to the existing MST instance. To specify a range of VLANs, use a hyphen. To specify a series of VLANs, use a comma. (Range: 1-4093)
exit	Return to global configuration mode.
spanning-tree mst *instance-id* **priority** *priority*	Set the switch priority for the specified spanning tree instance. • *instance-id* — ID of the spanning tree instance. (Range: 1-4094) • *priority* — Sets the switch priority for the specified spanning tree instance. This setting affects the likelihood that the switch is selected as the root switch. A lower value increases the probability that the switch is selected as the root switch. (Range: 0-61440)
CTRL + Z	Exit to Privileged EXEC mode.
show spanning-tree mst-configuration	View multiple spanning tree configuration information.
show spanning-tree instance *instance-id*	View information about the specified MSTI.

Configuring MSTP Interface Settings

Beginning in Privileged EXEC mode, use the following commands to configure MSTP settings for the switch.

Command	Purpose
configure	Enter global configuration mode.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3** or **port-channel 4**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12. The **range** keyword is also valid for LAGs (port-channels).
spanning-tree mst 0 external-cost *cost*	Set the external cost for the common spanning tree. (Range: 0–200000000)
spanning-tree mst *instance-id* **cost** *cost*	Configure the path cost for MST calculations. If a loop occurs, the spanning tree considers path cost when selecting an interface to put in the forwarding state. • *instance-ID* — ID of the spanning -tree instance. (Range: 1-4094) • *cost* — The port path cost. (Range: 0–200,000,000)
spanning-tree mst *instance-id* **port-priority** *priority*	Specify the priority of the port. The priority value is used to determine which ports are put in the forwarding state and which ports are put in the blocking state. A port with a lower priority value is more likely to be put into a forwarding state. • *instance-ID* — ID of the spanning tree instance. (Range: 1-4094) • *priority* — The port priority. (Range: 0–240 in multiples of 16)
CTRL + Z	Exit to Privileged EXEC mode.
show spanning-tree *interface* **instance** *instance-id*	View MST configuration information for the specified port or LAG (port-channel) and instance.

STP Configuration Examples

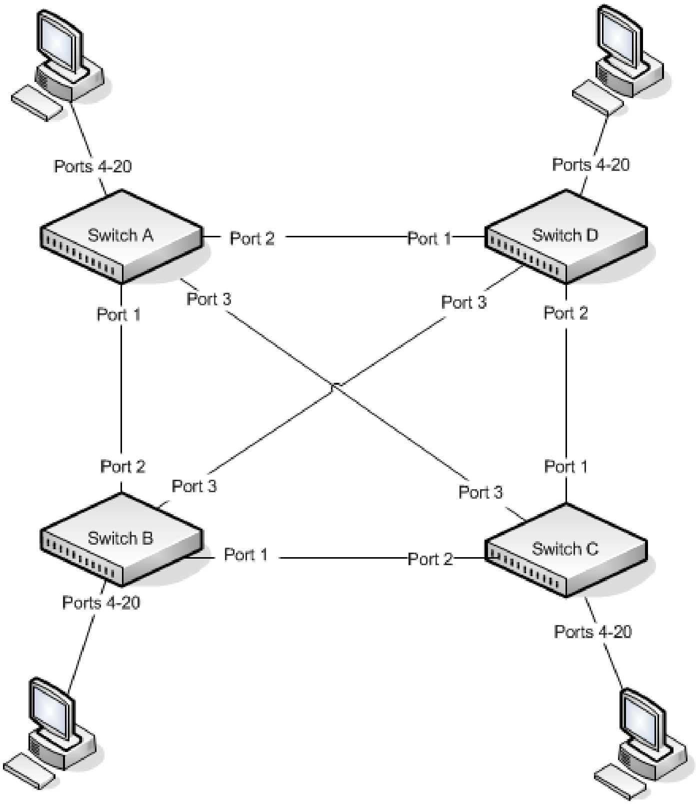
This section contains the following examples:

- STP Configuration Example
- MSTP Configuration Example
- RSTP-PV Access Switch Configuration Example

STP Configuration Example

This example shows a LAN with four switches. On each switch, ports 1, 2, and 3 connect to other switches, and ports 4–20 connect to hosts (in Figure 22-19, each PC represents 17 host systems).

Figure 22-19. STP Example Network Diagram



Of the four switches in Figure 22-19, the administrator decides that Switch A is the most centrally located in the network and is the least likely to be moved or redeployed. For these reasons, the administrator selects it as the root bridge for the spanning tree. The administrator configures Switch A with the highest priority and uses the default priority values for Switch B, Switch C, and Switch D.

For all switches, the administrator also configures ports 4–17 in Port Fast mode because these ports are connected to hosts and can transition directly to the Forwarding state to speed up the connection time between the hosts and the network.

The administrator also configures Port Fast BPDU filtering and Loop Guard to extend STP's capability to prevent network loops. For all other STP settings, the administrator uses the default STP values.

To configure the switch:

1 Connect to Switch A and configure the priority to be higher (a lower value) than the other switches, which use the default value of 32768.

```
console#config
console(config)#spanning-tree priority 8192
```

2 Configure ports 4–20 to be in Port Fast mode.

```
console(config)#interface range gi1/0/4-20
console(config-if)#spanning-tree portfast
console(config-if)#exit
```

3 Enable Loop Guard on ports 1–3 to help prevent network loops that might be caused if a port quits receiving BPDUs.

```
console(config)#interface range gi1/0/1-3
console(config-if)#spanning-tree guard loop
console(config-if)#exit
```

4 Enable Port Fast BPDU Filter. This feature is configured globally, but it affects only Port Fast-enabled access ports.

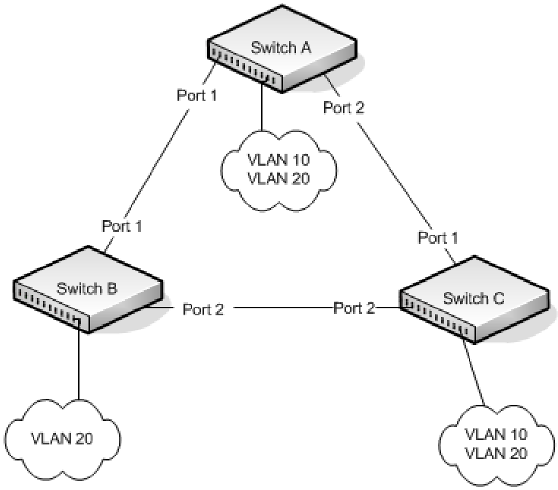
```
console(config)#spanning-tree portfast bpdufilter default
```

5 Repeat step 2 through step 4 on Switch B, Switch C, and Switch D to complete the configuration.

MSTP Configuration Example

This example shows how to configure IEEE 802.1s Multiple Spanning Tree (MST) protocol on the switches shown in Figure 22-20.

Figure 22-20. MSTP Configuration Example



To make multiple switches be part of the same MSTP region, make sure the STP operational mode for all switches is MSTP. Also, make sure the MST region name and revision level are the same for all switches in the region.

To configure the switches:

1 Create VLAN 10 (Switch A and Switch B) and VLAN 20 (all switches).

NOTE: Even Switch B does not have any ports that are members of VLAN 10, this VLAN must be created to allow the formation of MST regions made up of all bridges that exchange the same MST Configuration Identifier. It is only within these MST Regions that multiple instances can exist.

```
console#configure
console(config)#vlan 10,20
console(config-vlan10,20)#exit
console(config-vlan)#exit
```

2 Set the STP operational mode to MSTP.

```
console(config)#spanning-tree mode mst
```

3 Create MST instance 10 and associate it to VLAN 10.

```
console(config)#spanning-tree mst configuration
console(config-mst)#instance 10 add vlan 10
```

4 Create MST instances 20 and associate it to VLAN 20.

```
console(config-mst)#instance 20 add vlan 20
```

5 Change the region name and revision number so that all the bridges that want to be part of the same region can form the region. This step is required for MST to operate properly.

```
console(config-mst)#name dell
console(config-mst)#revision 0
console(config-mst)#exit
```

6 (Switch A only) Configure Switch A to be the root bridge of the spanning tree (CIST Regional Root) by configuring a higher root bridge priority.

```
console(config)#spanning-tree priority 8192
```

7 (Switch A only) Make Switch A the Regional Root for MSTI 1 by configuring a higher priority for MST ID 10.

```
console(config)#spanning-tree mst 10 priority 12288
```

8 (Switch A only) Change the priority of MST ID 20 to ensure Switch C is the Regional Root bridge for this MSTI.

```
console(config)#spanning-tree mst 20 priority 61440
console(config)#spanning-tree priority 8192
```

9 (Switch C only) Change the priority of port 1 to force it to be the root port for MST 20.

```
console(config)#interface gi1/0/1
console(config-if-Gi1/0/1)#spanning-tree mst 20 port-priority
64
console(config-if-Gi1/0/1)#exit
```

RSTP-PV Access Switch Configuration Example

In this configuration, all 1G ports are presumed to be connected to host machines, and the two 10G uplink ports are connected to an aggregation-layer switch with a total layer-2 network diameter of 4. The aggregation-layer switch can be a single switch or multiple switches, running either RSTP-PV or MSTP. For fastest convergence during failover scenarios, it is recommended that the uplink switches be configured in RSTP-PV mode.

Three VLANs are configured in addition to VLAN 1. Te1/0/1 is configured to be the primary uplink port and Te1/0/2 is configured to be the backup uplink.

1 Configure VLANs 2 through 4 and return to Global Config mode.

```
console#configure
console(config)#vlan 2-4
console(config-vlan2-4)#exit
```

2 Enable RSTP-PV.

```
console(config)#spanning-tree mode rapid-pvst
```

3 Configure for a maximum network diameter of 4.

```
console(config)#spanning-tree vlan 1-4 max-age 16
```

4 Configure access and trunk ports

```
console(config)#interface range gi1/0/1-48
console(config-if)#switchport mode access
console(config-if)exit
console(config)#interface range te1/0/1-2
console(config-if)#switchport mode trunk
console(config-if)#exit
```

5 Configure interface te1/0/1 as the preferred uplink.

```
console(config)#interface te1/0/1
console(config-if)#spanning-tree port-priority 112
console(config-if)#exit
```

6 Assign ports to VLANs.

```
console(config)#interface range gi1/0/1-12
console(config-if)#switchport access vlan 1
console(config-if)#exit
console(config)#interface range gi1/0/13-24
console(config-if)#switchport access vlan 2
console(config-if)#exit
console(config)#interface range gi1/0/25-36
console(config-if)#switchport access vlan 3
console(config-if)#exit
```

```
console(config)#interface range gi1/0/37-48
console(config-if)#switchport access vlan 4
console(config-if)#exit
```

RSTP-PV Aggregation-Layer Switch Configuration Example

In this configuration example, two aggregation-layer switches are configured. Ports 1–4 are configured in a LAG connecting the two aggregation-layer switches. Ports 12–24 are configured as down-links to twelve access-layer switches configured as in the previous example. Down-links to the access-layer switches have physical diversity; there is one downlink to each of the twelve access-layer switches from each of the paired aggregation-layer switches.

The uplink ports to the network core are configured as LAGs to provide link redundancy. It is presumed that the core links connect to a router running RSTP-PV. The configuration for the two aggregation-layer switches is identical, except for the diversity configuration noted below.

For forwarding diversity, the even numbered switch is made the root for the even-numbered VLANs. The odd numbered switch is made the root for the odd-numbered VLANs.

1. Create VLANs 2 through 4:

```
console#configure
console(config)#vlan 2-4
console(config-vlan2-4)#exit
```

2. Enable RSTP-PV:

```
console(config)#spanning-tree mode rapid-pvst
```

3. Configure for a max network diameter of 4:

```
console(config)#spanning-tree vlan 1-4 max-age 16
```

4. Configure one downlink trunk port per downlink switch:

```
console(config)#interface range te1/0/12-24
console(config-if-Te1/0/12-24)#switchport mode trunk
exit
```

5. Configure forwarding diversity for the even numbered switches:

```
console(config)#spanning-tree vlan 2,4 root primary
console(config)#spanning-tree vlan 1,3 root secondary
```

6. Configure forwarding diversity for the odd numbered switches:

```
console(config)#spanning-tree vlan 1,3 root primary
console(config)#spanning-tree vlan 2,4 root secondary
```

7. Configure two uplink ports per uplink switch:

```
console(config)#interface range fo1/0/1-2
```

```
console(config-if-fo1/0/1-2)#channel-group 1 mode active
console(config-if-fo1/0/1-2)#exit
```

8 Configure peer switch links:

```
console(config)#interface range te1/0/1-4
console(config-if-te1/0/1-4)#channel-group 2 mode active
console(config-if-te1/0/1-4)#exit
```

9 Configure the uplinks into a port channel:

```
console(config)#interface port-channel 1
console(config-if-port-channel 1)#switchport mode trunk
console(config-if-port-channel 1)#exit
```

10 Configure the peer links into a port channel and prefer to go to the core router or access switches directly, i.e. block the peer link unless it is needed:

```
console(config)#interface port-channel 1
console(config-if-port-channel 1)#switchport mode trunk
console(config-if-port-channel 1)#spanning-tree port-priority
144
console(config-if-port-channel 1)#exit
```

23

Discovering Network Devices

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes the Industry Standard Discovery Protocol (ISDP) feature and the Link Layer Discovery Protocol (LLDP) feature, including LLDP for Media Endpoint Devices (LLDP-MED).

The topics covered in this chapter include:

- Device Discovery Overview
- Default IDSP and LLDP Values
- Configuring ISDP and LLDP (Web)
- Configuring ISDP and LLDP (CLI)
- Device Discovery Configuration Examples

Device Discovery Overview

The switch software includes two different device discovery protocols: IDSP and LLDP. These protocols allow the switch to broadcast information about itself and to learn information about neighboring devices.

What Is ISDP?

The Industry Standard Discovery Protocol (ISDP) is a proprietary layer-2 network protocol that inter-operates with Cisco devices running the Cisco Discovery Protocol (CDP). ISDP is used to share information between neighboring devices. The switch software participates in the CDP protocol and is able to both discover and be discovered by other CDP-supporting devices.

What is LLDP?

LLDP is a standardized discovery protocol defined by IEEE 802.1AB. It allows stations residing on an 802 LAN to advertise major capabilities physical descriptions, and management information to physically adjacent devices allowing a network management system (NMS) to access and display this information.

LLDP is a one-way protocol; there are no request/response sequences. Information is advertised by stations implementing the transmit function, and is received and processed by stations implementing the receive function. The transmit and receive functions can be enabled/disabled separately on each switch port.

What is LLDP-MED?

LLDP-MED is an extension of the LLDP standard. LLDP-MED uses LLDP's organizationally-specific Type- Length-Value (TLV) extensions and defines new TLVs that make it easier for a VoIP deployment in a wired or wireless LAN/MAN environment. It also makes mandatory a few optional TLVs from LLDP and recommends not transmitting some TLVs.

The TLVs only communicate information; these TLVs do not automatically translate into configuration. An external application may query the MED MIB and take management actions in configuring functionality.

Why are Device Discovery Protocols Needed?

The device discovery protocols are used primarily in conjunction with network management tools to provide information about network topology and configuration, and to help troubleshoot problems that occur on the network. The discovery protocols can also facilitate inventory management within a company.

LLDP and the LLDP-MED extension are vendor-neutral discovery protocols that can discover devices made by numerous vendors. LLDP-MED is intended to be used on ports that connect to VoIP phones. Additional applications for LLDP-MED include Power over Ethernet management.

ISDP interoperates with the Cisco-proprietary CDP protocol and is most effective in an environment that contains many Cisco devices.

Default IDSP and LLDP Values

ISDP and LLDP are globally enabled on the switch and enabled on all ports by default. By default, the switch transmits and receives LLDP information on all ports. LLDP-MED is disabled on all ports.

Table 23-1 summarizes the default values for ISDP.

Table 23-1. ISDP Defaults

Parameter	Default Value
ISDP Mode	Enabled (globally and on all ports)
ISDPv2 Mode	Enabled (globally and on all ports)
Message Interval	30 seconds
Hold Time Interval	180 seconds
Device ID	none
Device ID Format Capability	Serial Number, Host Name
Device ID Format	Serial Number

Table 23-2 summarizes the default values for LLDP.

Table 23-2. LLDP Defaults

Parameter	Default Value
Transmit Mode	Enabled on all ports
Receive Mode	Enabled on all ports
Transmit Interval	30 seconds
Hold Multiplier	4
Reinitialization Delay	2 seconds
Notification Interval	5 seconds
Transmit Management Information	Disabled
Notification Mode	Disabled
Included TLVs	None

Table 23-3 summarizes the default values for LLDP-MED.

Table 23-3. LLDP-MED Defaults

Parameter	Default Value
LLDP-MED Mode	Disabled on all ports
Config Notification Mode	Disabled on all ports
Transmit TVLs	MED Capabilities Network Policy

Configuring ISDP and LLDP (Web)

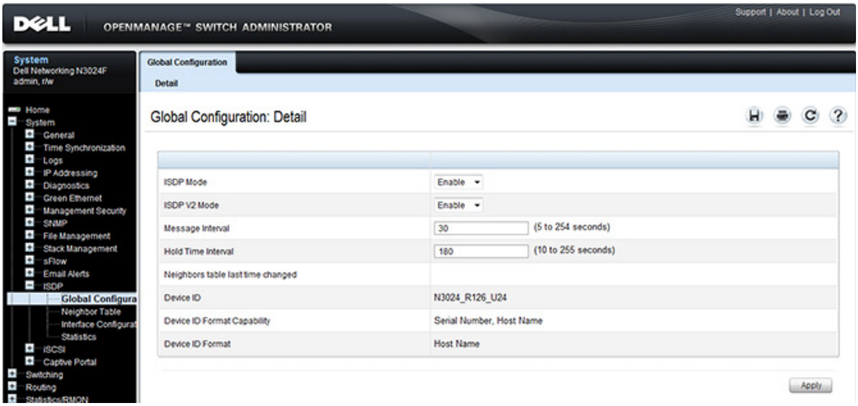
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring IDSP and LLDP/LLDP-MED on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

ISDP Global Configuration

The **ISDP Global Configuration** page enables configuring the ISDP settings for the switch, such as the administrative mode.

To access the **ISDP Global Configuration** page, click **System** → **ISDP** → **Global Configuration** in the navigation panel.

Figure 23-1. ISDP Global Configuration

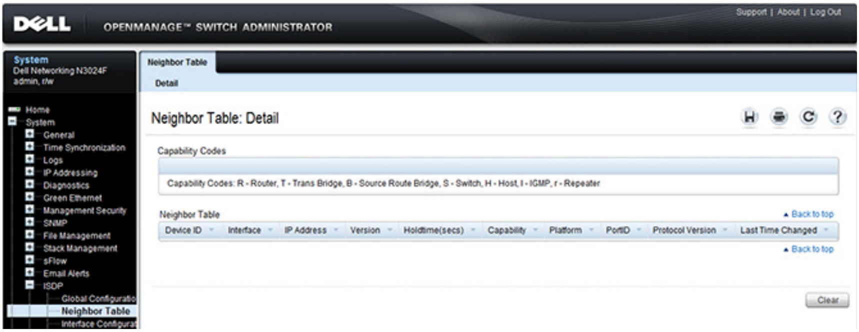


ISDP Cache Table

The **ISDP Neighbor Table** page enables viewing information about other devices the switch has discovered through the ISDP.

To access the **ISDP Neighbor Table** page, click **System** → **ISDP** → **Neighbor Table** in the navigation panel.

Figure 23-2. ISDP Neighbor Table



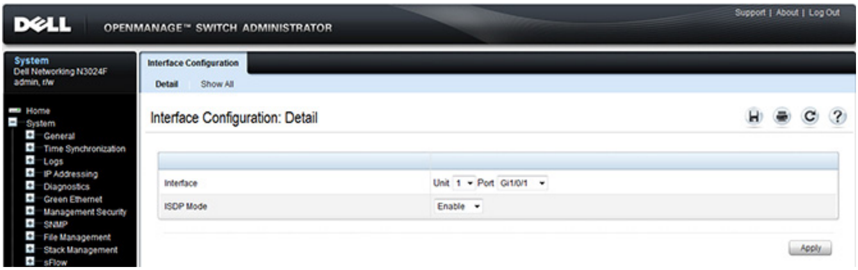
ISDP Interface Configuration

The **ISDP Interface Configuration** page enables configuring the ISDP settings for each interface.

If ISDP is enabled on an interface, it must also be enabled globally in order for the interface to transmit ISDP packets. If the ISDP mode on the ISDP Global Configuration page is disabled, the interface will not transmit ISDP packets, regardless of the mode configured on the interface.

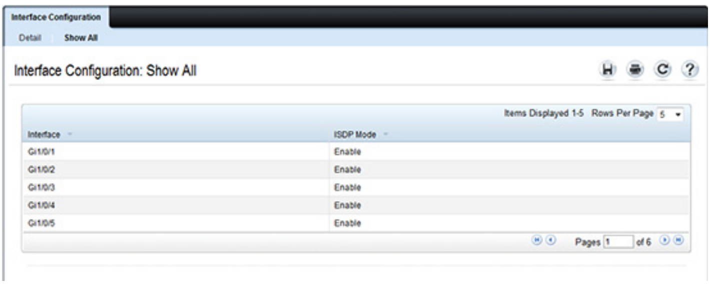
To access the **ISDP Interface Configuration** page, click **System** → **ISDP** → **Interface Configuration** in the navigation panel.

Figure 23-3. ISDP Interface Configuration



To view view the ISDP mode for multiple interfaces, click **Show All**.

Figure 23-4. ISDP Interface Summary

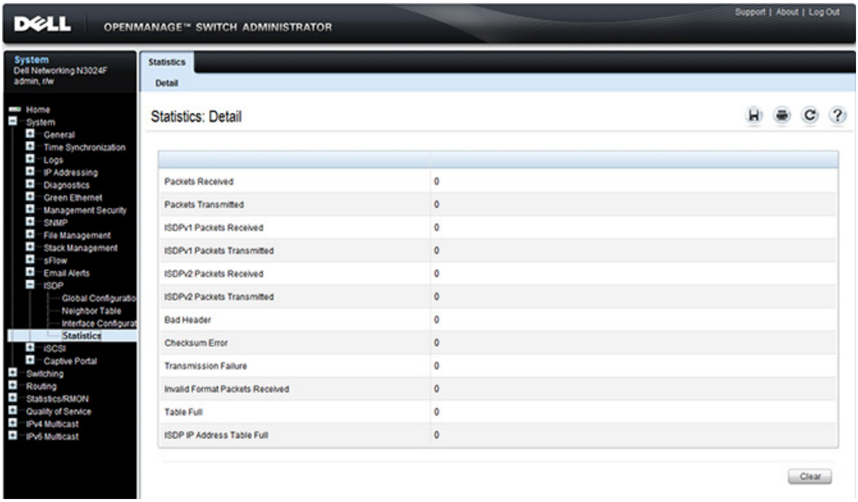


ISDP Statistics

The **ISDP Statistics** page enables viewing information about the ISDP packets sent and received by the switch.

To access the **ISDP Statistics** page, click **System** → **ISDP** → **Statistics** in the navigation panel.

Figure 23-5. ISDP Statistics

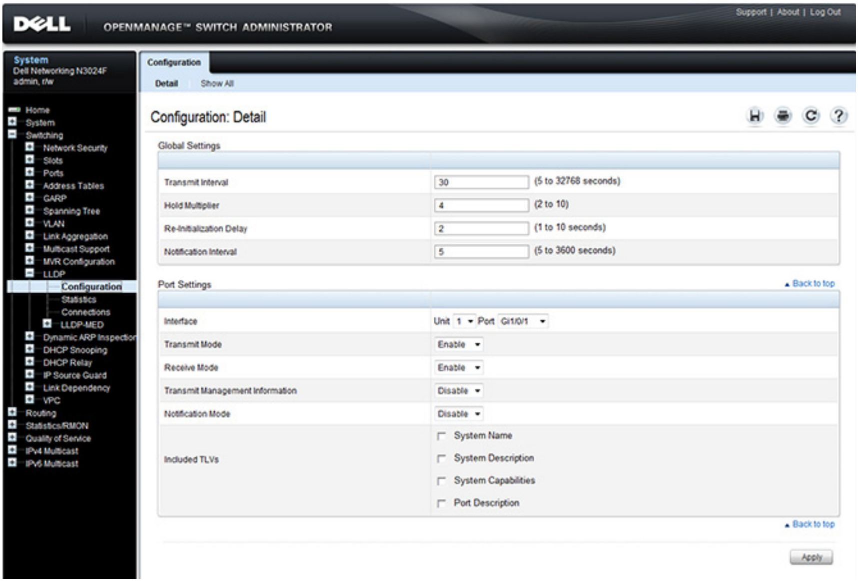


LLDP Configuration

Use the **LLDP Configuration** page to specify LLDP parameters. Parameters that affect the entire system as well as those for a specific interface can be specified here.

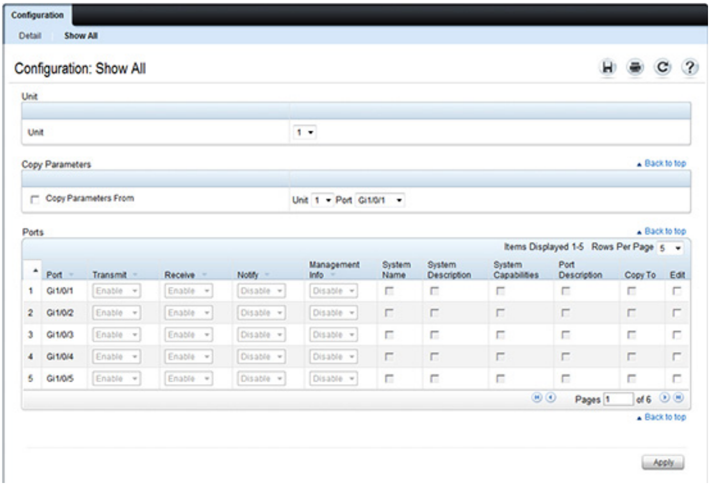
To display the **LLDP Configuration** page, click **Switching** → **LLDP** → **Configuration** in the navigation panel.

Figure 23-6. LLDP Configuration



To view the **LLDP Interface Settings Table**, click **Show All**. The **LLDP Interface Settings Table** page enables viewing and editing information about the LLDP settings for multiple interfaces.

Figure 23-7. LLDP Interface Settings Table



LLDP Statistics

Use the LLDP Statistics page to view LLPD-related statistics.

To display the **LLDP Statistics** page, click **Switching → LLDP → Statistics** in the navigation panel.

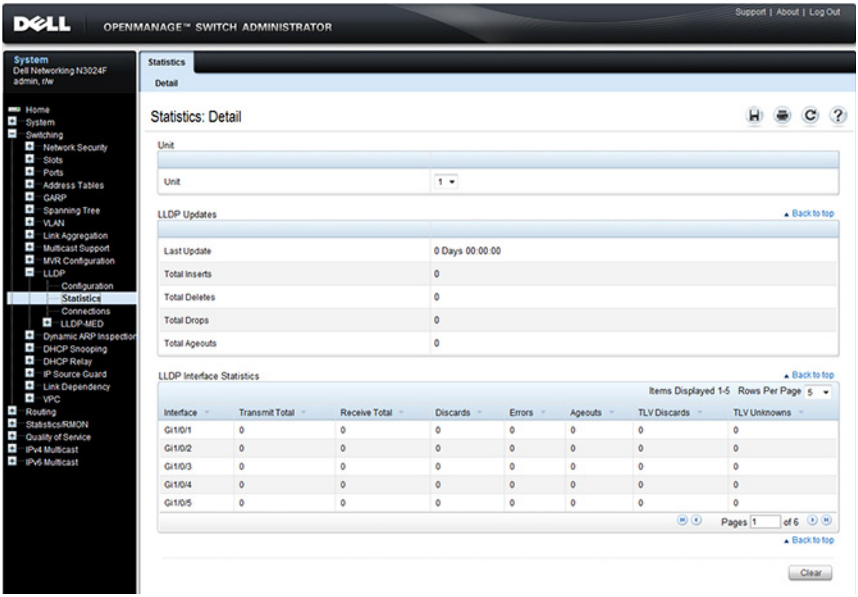


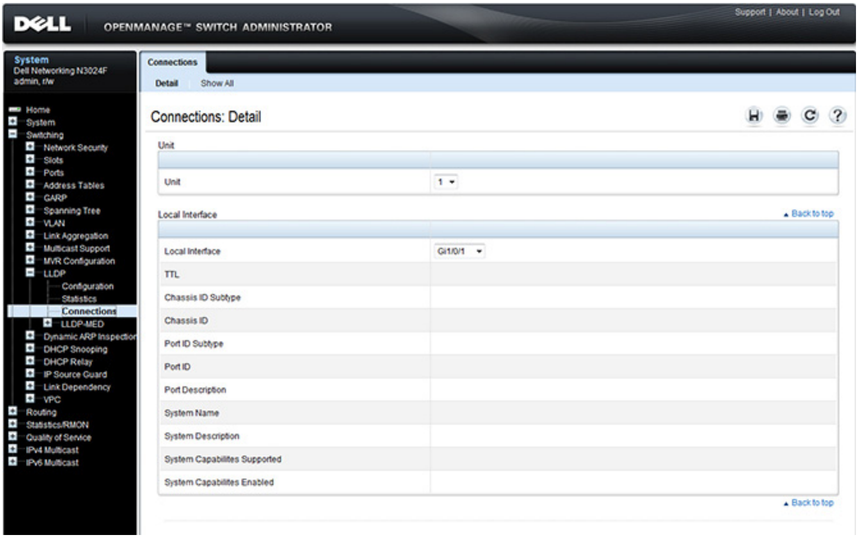
Figure 23-8. LLDP Statistics

LLDP Connections

Use the **LLDP Connections** page to view the list of ports with LLDP enabled. Basic connection details are displayed.

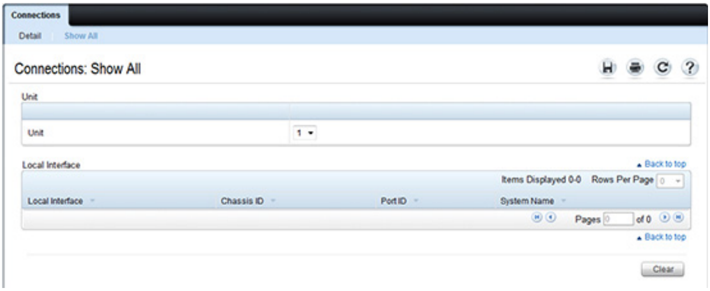
To display the **LLDP Connections** page, click **Switching** → **LLDP** → **Connections** in the navigation panel.

Figure 23-9. LLDP Connections



To view additional information about a device connected to a port that has been discovered through LLDP, click the port number in the Local Interface table (it is a hyperlink), or click **Details** and select the port with the connected device.

Figure 23-10. LLDP Connection Detail

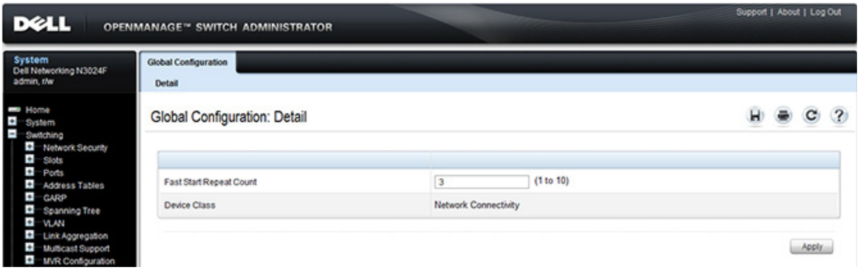


LLDP-MED Global Configuration

Use the **LLDP-MED Global Configuration** page to change or view the LLDP-MED parameters that affect the entire system.

To display the **LLDP-MED Global Configuration** page, click **Switching LLDP** → **LLDP-MED** → **Global Configuration** in the navigation panel.

Figure 23-11. LLDP-MED Global Configuration

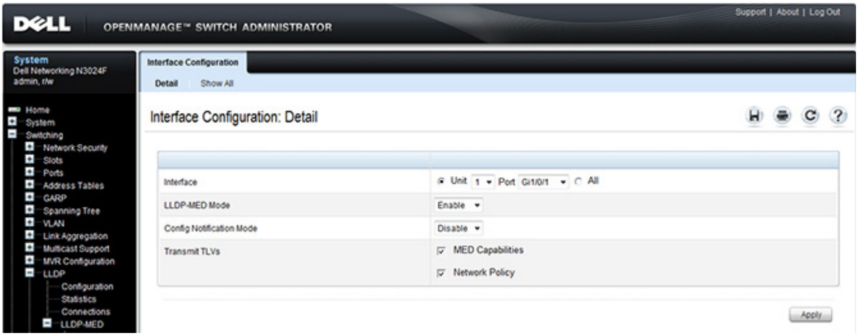


LLDP-MED Interface Configuration

Use the **LLDP-MED Interface Configuration** page to specify LLDP-MED parameters that affect a specific interface.

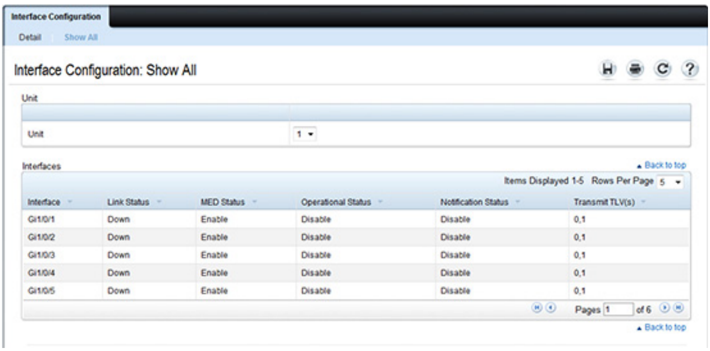
To display the **LLDP-MED Interface Configuration** page, click **Switching** → **LLDP** → **LLDP-MED** → **Interface Configuration** in the navigation panel.

Figure 23-12. LLDP-MED Interface Configuration



To view the **LLDP-MED Interface Summary** table, click **Show All**.

Figure 23-13. LLDP-MED Interface Summary

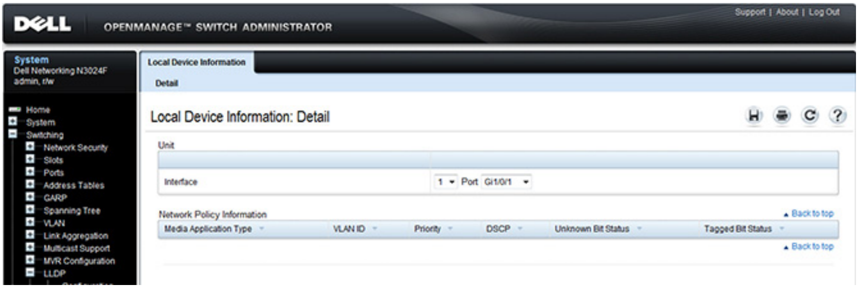


LLDP-MED Local Device Information

Use the **LLDP-MED Local Device Information** page to view the advertised LLDP local data for each port.

To display the **LLDP-MED Local Device Information** page, click **Switching LLDP LLDP-MED Local Device Information** in the navigation panel.

Figure 23-14. LLDP-MED Local Device Information

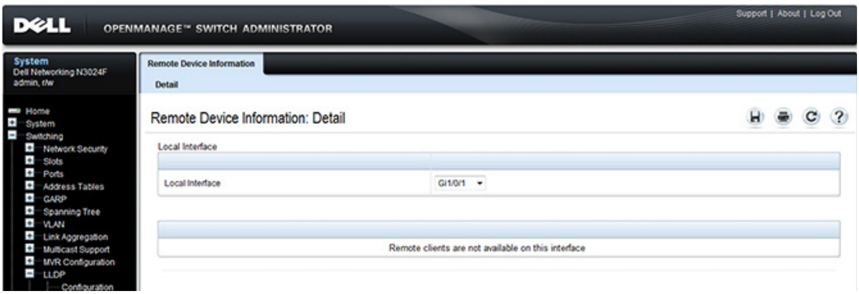


LLDP-MED Remote Device Information

Use the **LLDP-MED Remote Device Information** page to view the advertised LLDP data advertised by remote devices.

To display the **LLDP-MED Remote Device Information** page, click **Switching LLDP LLDP-MED Remote Device Information** in the the navigation panel.

Figure 23-15. LLDP-MED Remote Device Information



Configuring ISDP and LLDP (CLI)

This section provides information about the commands you use to manage and view the device discovery protocol features on the switch. For more information about these commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Global ISDP Settings

Beginning in Privileged EXEC mode, use the following commands to configure ISDP settings that affect the entire switch.

Command	Purpose
configure	Enter Global Configuration mode.
isdp enable	Administratively enable ISDP on the switch.
isdp advertise-v2	Allow the switch to send ISDPv2 packets.
isdp holdtime *time*	Specify the number of seconds the device that receives ISDP packets from the switch should store information sent in the ISDP packet before discarding it.
isdp timer *time*	Specify the number of seconds to wait between sending new ISDP packets.
exit	Exit to Privileged EXEC mode.
show isdp	View global ISDP settings.

Enabling ISDP on a Port

Beginning in Privileged EXEC mode, use the following commands to enable ISDP on a port.

Command	Purpose
configure	Enter Global Configuration mode.
interface *interface*	Enter interface configuration mode for the specified interface.
isdp enable	Administratively enable ISDP on the switch.
exit	Exit to Global Config mode.
exit	Exit to Privileged Exec mode.
show isdp interface all	View the ISDP mode on all interfaces.

Viewing and Clearing ISDP Information

Beginning in Privileged EXEC mode, use the following commands to view and clear the contents of the ISDP table and to view and clear ISDP statistics.

Command	Purpose
show isdp entry {**all** | *deviceid*}	View information about all entries or a specific entry in the ISDP table.
show isdp neighbors	View the neighboring devices discovered through ISDP.
clear isdp table	Clear all entries, including discovered neighbors, from the ISDP table.
show isdp traffic	View ISDP statistics.
clear isdp counters	Reset all ISDP statistics to zero.

Configuring Global LLDP Settings

Beginning in Privileged EXEC mode, use the following commands to configure LLDP settings that affect the entire switch.

Command	Purpose
configure	Enter Global Configuration mode.
lldp notification-interval *interval*	Specify how often, in seconds, the switch should send remote data change notifications.
lldp timers [**interval** *transmit-interval*] [**hold** *hold-value*] [**reinit** *reinit-delay*]	Configure the timing for local data transmission on ports enabled for LLDP. • *transmit-interval* — The interval in seconds at which to transmit local data LLDP PDUs. (Range: 5–32768 seconds) • *hold-value* — Multiplier on the transmit interval used to set the TTL in local data LLDP PDUs. (Range: 2–10) • *reinit-delay* — The delay in seconds before re-initialization. (Range: 1–10 seconds)
exit	Exit to Privileged EXEC mode.
show lldp	View global LLDP settings.

Configuring Port-based LLDP Settings

Beginning in Privileged EXEC mode, use the following commands to configure per-port LLDP settings.

Command	Purpose
configure	Enter Global Configuration mode.
interface *interface*	Enter interface configuration mode for the specified Ethernet interface.
lldp transmit	Enable the LLDP advertise (transmit) capability.
lldp receive	Enable the LLDP receive capability so that the switch can receive LLDP Protocol Data Units (LLDP PDUs) from other devices.
lldp transmit-mgmt	Include the transmission of local system management address information in the LLDP PDUs.

Command	Purpose
lldp notification	Enable remote data change notifications on the interface.
lldp transmit-tlv [**sys-desc**] [**sys-name**] [**sys-cap**] [**port-desc**]	Specify which optional type-length-value settings (TLVs) in the 802.1AB basic management set will be transmitted in the LLDP PDUs. • sys-name — Transmits the system name TLV • sys-desc — Transmits the system description TLV • sys-cap — Transmits the system capabilities TLV • port desc — Transmits the port description TLV
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show lldp interface all	View LLDP settings for all interfaces.

Viewing and Clearing LLDP Information

Beginning in Privileged EXEC mode, use the following commands to view transmitted and received LLDP information and to view and clear LLDP statistics.

Command	Purpose
show lldp local-device {**all** | *interface* | **detail** *interface*}	View LLDP information advertised by all ports or the specified port. Include the keyword detail to see additional information.
show lldp remote-device {**all** | *interface* | **detail** *interface*}	View LLDP information received by all ports or by the specified port. Include the keyword detail to see additional information.
clear lldp remote-data	Delete all LLDP information from the remote data table.
show lldp statistics	View LLDP traffic statistics.
clear lldp statistics	Reset the LLDP statistics counters to zero.

Configuring LLDP-MED Settings

Beginning in Privileged EXEC mode, use the following commands to configure LLDP-MED settings that affect the entire switch.

Command	Purpose
configure	Enter Global Configuration mode.
lldp med faststartrepeatcount *count*	Specifies the number of LLDP PDUs that will be transmitted when the protocol is enabled.
interface *interface*	Enter interface configuration mode for the specified Ethernet interface.
lldp med	Enable LLDP-MED on the interface.
lldp med confignotification	Allow the port to send topology change notifications.
lldp med transmit-tlv [**capabilities**] [**network-policy**]	Specify which optional TLVs in the LLDP MED set are transmitted in the LLDP PDUs.
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show lldp med	View global LLDP-MED settings.
show lldp med interface {**all** | *interface*}	View LLDP-MED settings for all ports or for the specified port.

Viewing LLDP-MED Information

Beginning in Privileged EXEC mode, use the following commands to view information about the LLDP-MED Protocol Data Units (PDUs) that are sent and have been received.

Command	Purpose
show lldp med local-device detail *interface*	View LLDP information advertised by the specified port.
show lldp remote-device {**all** | *interface* | **detail** *interface*}	View LLDP-MED information received by all ports or by the specified port. Include the keyword detail to see additional information.

Device Discovery Configuration Examples

This section contains the following examples:

- Configuring ISDP
- Configuring LLDP

Configuring ISDP

This example shows how to configure ISDP settings on the switch.

To configure the switch:

1 Specify the number of seconds that a remote device should keep the ISDP information sent by the switch before discarding it.

```
console#configure
console(config)#isdp holdtime 60
```

2 Specify how often, in seconds, the ISDP-enabled ports should transmit information.

```
console(config)#isdp timer 45
```

3 Enable ISDP on interface 1/0/3.

```
console(config)#interface tengigabitEthernet1/0/3
console(config-if-Te1/0/3)#isdp enable
```

4 Exit to Privileged EXEC mode and view the LLDP settings for the switch and for interface 1/0/3.

```
console(config-if-Te1/0/3)# <CTRL + Z>
console#show isdp
```

```
Timer...................................45
Hold Time...............................60
Version 2 Advertisements................Enabled
Neighbors table time since last change...00 days 00:00:00
Device ID...............................none
Device ID format capability.............Serial Number, Host Name
Device ID format........................Serial Number

console#show isdp interface te1/0/3

Interface        Mode
---------------  ----------
Te1/0/3          Enabled
```

Configuring LLDP

This example shows how to configure LLDP settings for the switch and to allow 10-Gigabit Ethernet port 1/0/3 to transmit all LLDP information available.

To configure the switch:

1. Configure the transmission interval, hold multiplier, and reinitialization delay for LLDP PDUs sent from the switch.

```
console#configure
console(config)#lldp timers interval 60 hold 5 reinit 3
```

2. Enable port 1/0/3 to transmit and receive LLDP PDUs.

```
console(config)#interface TengigabitEthernet1/0/3
console(config-if-Te1/0/3)#lldp transmit
console(config-if-Te1/0/3)#lldp receive
```

3. Enable port 1/0/3 to transmit management address information in the LLDP PDUs and to send topology change notifications if a device is added or removed from the port.

```
console(config-if-Te1/0/3)#lldp transmit-mgmt
console(config-if-Te1/0/3)#lldp notification
```

4. Specify the TLV information to be included in the LLDP PDUs transmitted from port 1/0/3.

```
console(config-if-Te1/0/3)#lldp transmit-tlv sys-name sys-desc
sys-cap port-desc
```

5. Set the port description to be transmitted in LLDP PDUs.

```
console(config-if-Te1/0/3)#description "Test Lab Port"
```

6. Exit to Privileged EXEC mode.

```
console(config-if-Te1/0/3)# <CTRL + Z>
```

7 View global LLDP settings on the switch.

```
console#show lldp

LLDP Global Configuration

Transmit Interval..................... 60 seconds
Transmit Hold Multiplier............. 5
Reinit Delay......................... 3 seconds
Notification Interval................ 5 seconds
```

8 View summary information about the LLDP configuration on port 1/0/3.

```
console#show lldp interface te1/0/3

LLDP Interface Configuration

Interface  Link    Transmit  Receive   Notify    TLVs     Mgmt
---------  ------  --------  --------  --------  -------  ----
Te1/0/3    Down    Enabled   Enabled   Enabled   0,1,2,3  Y

TLV Codes: 0- Port Description,   1- System Name
           2- System Description, 3- System Capabilities
```

9 View detailed information about the LLDP configuration on port 1/0/3.

```
console#show lldp local-device detail te1/0/3

LLDP Local Device Detail

Interface: Te1/0/3

Chassis ID Subtype: MAC Address
Chassis ID: 00:1E:C9:AA:AA:07
Port ID Subtype: Interface Name
Port ID: Te1/0/3
System Name: console
System Description: Dell Networking N3048
Port Description: Test Lab Port
System Capabilities Supported: bridge, router
System Capabilities Enabled: bridge
Management Address:
    Type: IPv4
    Address: 192.168.2.1
```

24

Port-Based Traffic Control

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure features that provide traffic control through filtering the type of traffic or limiting the speed or amount of traffic on a per-port basis. The features this section describes includes flow control, storm control, protected ports, and Link Local Protocol Filtering (LLPF), which is also known as Cisco Protocol Filtering.

The topics covered in this chapter include:

- Port-Based Traffic Control Overview
- Default Port-Based Traffic Control Values
- Configuring Port-Based Traffic Control (Web)
- Configuring Port-Based Traffic Control (CLI)
- Port-Based Traffic Control Configuration Example

Port-Based Traffic Control Overview

Table 24-1 provides a summary of the features this chapter describes.

Table 24-1. Port-Based Traffic Control Features

Feature	Description
Flow control	Allows traffic transmission between a switch port and another Ethernet device to be paused for a specified period of time when congestion occurs.
Storm control	Limits the amount of broadcast, unknown layer-2 unicast, and multicast frames accepted and forwarded by the switch.
Protected ports	Prevents traffic from flowing between members of the same protected port group.
Error recovery	Automatically brings up interfaces that have been diagnostically disabled.
Loop protection	Disables ports that are looped.

Table 24-1. Port-Based Traffic Control Features

Feature	Description
LLPF	Filters proprietary protocols that should not normally be relayed by a bridge.

The Priority Flow Control (PFC) feature, which is available on the Dell Networking N4000 Series switches only, provides a way to distinguish which traffic on a physical link is paused when congestion occurs based on the priority of the traffic. For more information, see "Data Center Bridging Features " on page 1051.

What is Flow Control?

IEEE 802.3 Annex 31B flow control allows nodes that transmit at slower speeds to communicate with higher speed switches by requesting that the higher speed switch refrain from sending packets. Transmissions are temporarily halted to prevent buffer overflows. Enabling the flow control feature allows Dell Networking N-Series switches to process pause frames received from connected devices. Dell Networking N-Series switches do not transmit pause frames.

Flow control is supported only on ports that are configured for full-duplex mode of operation.

What is Storm Control?

A LAN storm is the result of an excessive number of broadcast, multicast, or unknown unicast messages simultaneously transmitted across a network by a single port. Forwarded message responses can overload network resources and cause network congestion.

The storm control feature allows the switch to measure the incoming broadcast, multicast, and/or unknown unicast packet rate per port and discard packets when the rate exceeds the defined threshold. Optionally, the system can issue a log message and a trap, or it can shut down (diagnostically disable) the port. Storm control is enabled per interface, by defining the packet type and the rate at which the packets are transmitted. For each type of traffic (broadcast, multicast, or unknown unicast) a threshold level can be configured, which is expressed as a percentage of the total available bandwidth on the port. If the ingress rate of that type of packet is greater than the configured threshold level the port drops the excess traffic until the ingress rate for the packet type falls below the threshold.

When configuring the limit in terms of link bandwidth, the actual rate of ingress traffic required to activate storm-control is based on the size of incoming packets, and a hard-coded average packet size of 512 bytes is used to calculate a packet-per-second (pps) rate, as the forwarding-plane requires PPS versus an absolute rate in Kbps. For example, if the configured limit is 10% on a 1 Gbps link, this is converted to ~25000 PPS, and this PPS limit is set in the hardware.

What are Protected Ports?

The switch supports up to three separate groups of protected ports. Traffic can flow between protected ports belonging to different groups, but not within the same group.

A port can belong to only one protected port group. You must remove an interface from one group before adding it to another group.

Port protection occurs within a single switch. Protected port configuration does not affect traffic between ports on two different switches. No traffic forwarding is possible between two protected ports.

What is Error Recovery?

The error recovery feature enables the administrator to configure the switch to automatically bring interfaces that have been diagnostically disabled back into service automatically. The administrator may configure the causes for which error recovery will bring an interface back into service and may also configure the time interval over which error recovery is attempted. If an interface brought back into service by the error recovery mechanism has a subsequent failure, it will again be diagnostically disabled.

What is Link Local Protocol Filtering?

The Link Local Protocol Filtering (LLPF) feature can help troubleshoot network problems that occur when a network includes proprietary protocols running on standards-based switches. LLPF allows Dell Networking N1500, N2000, N3000, and N4000 Series switches to filter out various Cisco proprietary protocol data units (PDUs) and/or ISDP packets if problems occur with these protocols running on standards-based switches. If certain protocol PDUs cause unexpected results, LLPF can be enabled to prevent those PDUs from being processed by the switch.

The LLPF feature can be configured per-port to block any combination (or all) of the following PDUs:

- Industry Standard Discovery Protocol (ISDP)
- VLAN Trunking Protocol (VTP)
- Dynamic Trunking Protocol (DTP)
- UniDirectional Link Detection (UDLD)
- Port Aggregation Protocol (PAgP)
- Shared Spanning Tree Protocol (SSTP)

Access Control Lists (ACLs) and LLPF can exist on the same interface. However, the ACL rules override the LLPF rules when there is a conflict. Similarly, DiffServ and LLPF can both be enabled on an interface, but DiffServ rules override LLPF rules when there is a conflict.

If Industry Standard Discovery Protocol (ISDP) is enabled on an interface, and the LLPF feature on an interface is enabled and configured to drop ISDP PDUs, the ISDP configuration overrides the LLPF configuration, and the ISDP PDUs are allowed on the interface.

What is Loop Protection?

Dell Networking implements a subset of the Configuration Testing Protocol (CTP) for the detection of network loops. The Configuration Testing Protocol is part of the original Ethernet specification. It does not appear in the IEEE 802 standard.

The Dell implementation of the Loop Protocol unicasts a CTP reply packet with the following field settings:

Source MAC Address:	switch L2 MAC address
Destination MAC Address:	switch L2 MAC address
Ether Type:	0x9000 (LOOP)
Skip Count:	0
Functions:	Reply
Receipt Number:	0
Data:	0

If any interface receives CTP packets with the switch's MAC address as the source, and the number of such packets received is in excess of the configured limit, the interface is error-disabled with a Loop Protection cause. The default limit is three packets received. Since all switch ports share the same MAC address, multiple ports may be disabled by a network loop. Disabled ports may be configured to be brought back into service by the Error Recovery feature.

The switch never sends a response to received CTP packets. The switch may flood the first few CTP packets it receives until a MAC address entry is placed in the CAM.

The CTP protocol operates on physical Ethernet interfaces only. It does not operate over Link Aggregation Groups.

The CTP protocol does not operate over the out-of-band interface.

The CTP protocol is enabled on all physical Ethernet interfaces by default. CTP packet reception may blocked by spanning tree. With the default settings, spanning tree will detect a loop in the network prior to CTP and set the affected interfaces into the spanning tree blocked state.

Default Port-Based Traffic Control Values

Table 24-2 lists the default values for the port-based traffic control features that this chapter describes.

Table 24-2. Default Port-Based Traffic Control Values

Feature	Default
Flow control	Enabled
Storm control	Disabled
Protected ports	None
LLPF	No protocols are blocked

Configuring Port-Based Traffic Control (Web)

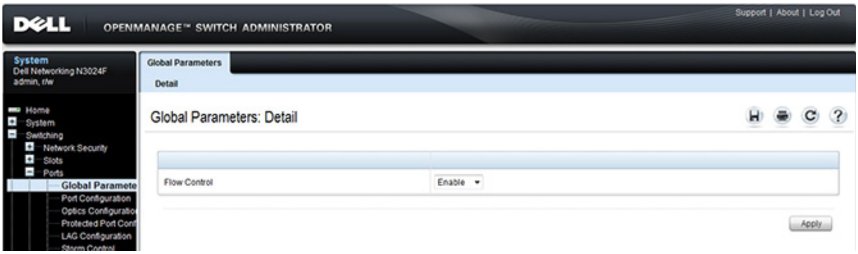
This section provides information about the OpenManage Switch Administrator pages to use to control port-based traffic on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click  at the top of the page.

Flow Control (Global Port Parameters)

Use the **Global Parameters** page for ports to enable or disable flow control support on the switch.

To display the **Global Parameters** page, click **Switching → Ports → Global Parameters** in the navigation menu.

Figure 24-1. Global Port Parameters

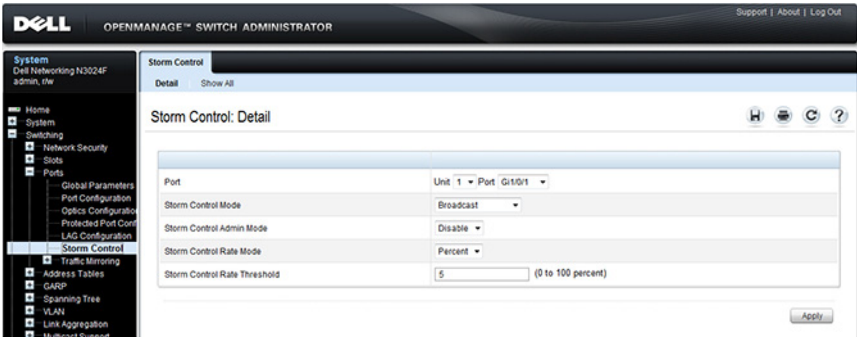


Storm Control

Use the **Storm Control** page to enable and configure the storm control feature.

To display the **Storm Control** interface, click **Switching** → **Ports** → **Storm Control** in the navigation menu.

Figure 24-2. Storm Control



Configuring Storm Control Settings on Multiple Ports

To configure storm control on multiple ports:

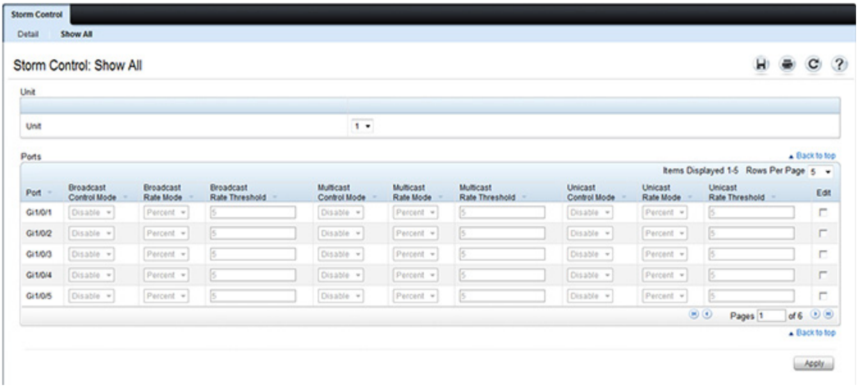
1 Open the **Storm Control** page.

2 Click **Show All** to display the **Storm Control Settings Table**.

3 In the Ports list, select the check box in the **Edit** column for the port to configure.

4 Select the desired storm control settings.

Figure 24-3. Storm Control



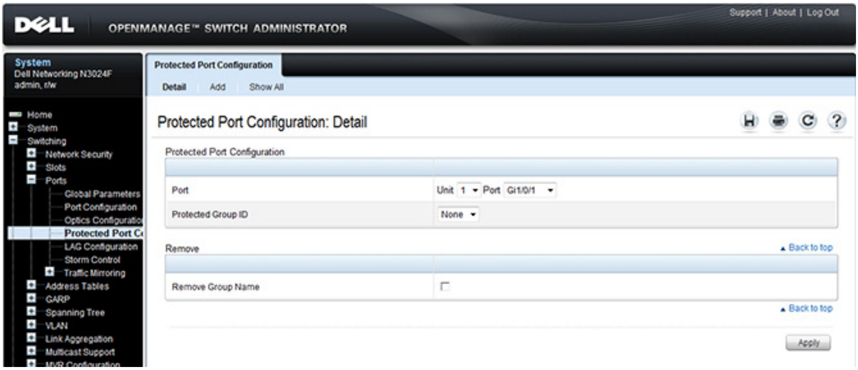
5 Click **Apply**.

Protected Port Configuration

Use the **Protected Port Configuration** page to prevent ports in the same protected ports group from being able to see each other’s traffic.

To display the **Protected Port Configuration** page, click **Switching** → **Ports** → **Protected Port Configuration** in the navigation menu.

Figure 24-4. Protected Port Configuration

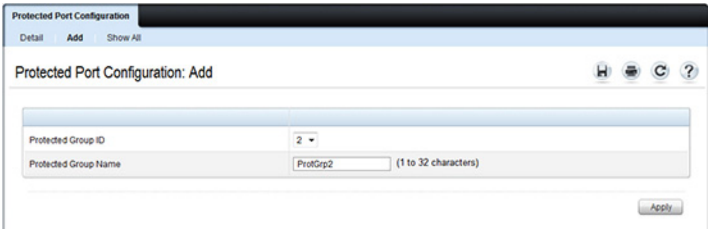


Configuring Protected Ports

To configure protected ports:

1. Open the **Protected Ports** page.
2. Click **Add** to display the **Add Protected Group** page.
3. Select a group (0–2).
4. Specify a name for the group.

Figure 24-5. Add Protected Ports Group



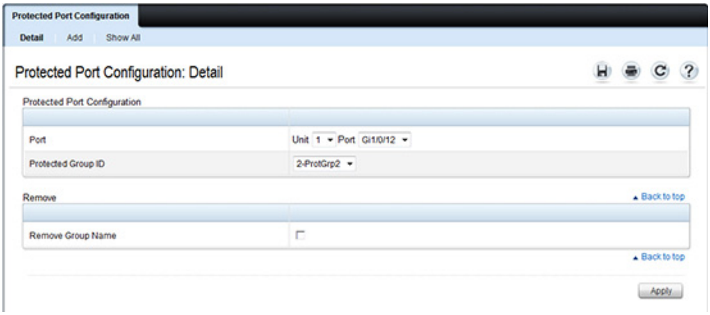
5 Click **Apply**.

6 Click **Protected Port Configuration** to return to the main page.

7 Select the port to add to the group.

8 Select the protected port group ID.

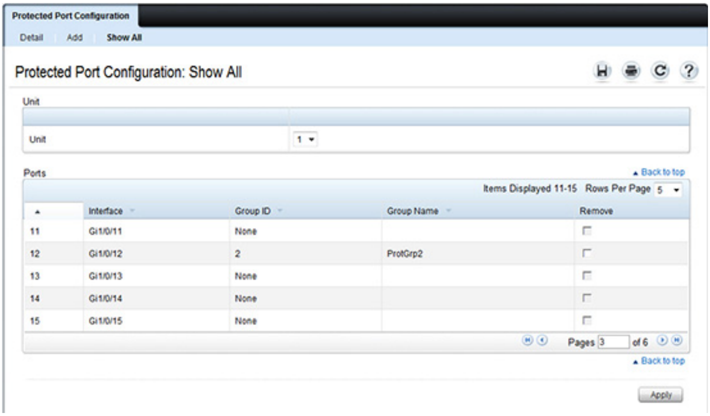
Figure 24-6. Add Protected Ports



9 Click **Apply**.

10 To view protected port group membership information, click **Show All**.

Figure 24-7. View Protected Port Information



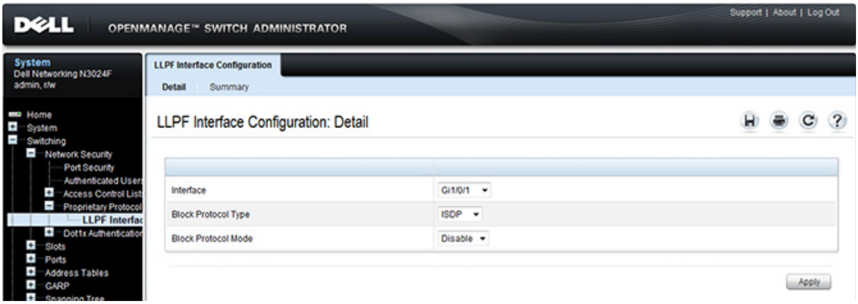
11 To remove a port from a protected port group, select the **Remove** check box associated with the port and click **Apply**.

LLPF Configuration

Use the **LLPF Interface Configuration** page to filter out various proprietary protocol data units (PDUs) and/or ISDP if problems occur with these protocols running on standards-based switches.

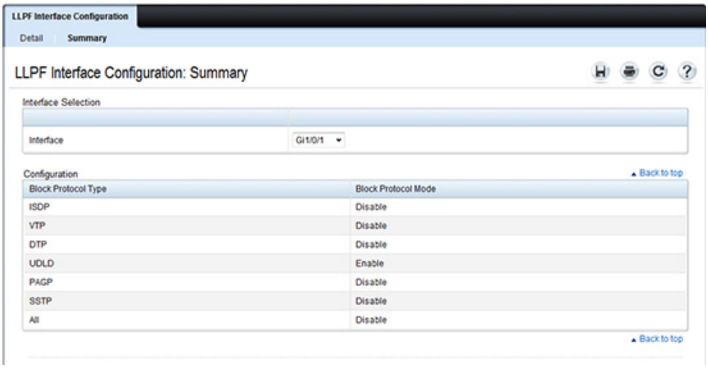
To display the **LLPF Interface Configuration** page, click **Switching** → **Network Security** → **Proprietary Protocol Filtering** → **LLPF Interface Configuration** the navigation menu.

Figure 24-8. LLPF Interface Configuration



To view the protocol types that have been blocked for an interface, click **Show All**.

Figure 24-9. LLPF Filtering Summary



Configuring Port-Based Traffic Control (CLI)

This section provides information about the commands used for configuring port-based traffic control settings. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Flow Control and Storm Control

Beginning in Privileged EXEC mode, use the following commands to configure the flow control and storm control features.

Command	Purpose
configure	Enter global configuration mode.
flowcontrol	Globally enable flow control.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
storm-control broadcast [**level** *rate*]	Enable broadcast storm recovery mode on the interface and (optionally) set the threshold. *rate* — threshold as percentage of port speed. The percentage is converted to a PacketsPerSecond value based on a 512 byte average packet size.
storm-control multicast [**level** *rate*]	Enable multicast storm recovery mode on the interface and (optionally) set the threshold. *rate* — threshold as percentage of port speed. The percentage is converted to a PacketsPerSecond value based on a 512 byte average packet size.
storm-control unicast [**level** *rate*]	Enable unknown unicast storm recovery mode on the interface and (optionally) set the threshold. *rate* — threshold as percentage of port speed. The percentage is converted to a PacketsPerSecond value based on a 512 byte average packet size.

Command	Purpose
CTRL + Z	Exit to Privileged EXEC mode.
show interfaces detail *interface*	Display detailed information about the specified interface, including the flow control status.
show storm-control	View whether 802.3x flow control is enabled on the switch.
show storm-control [*interface* | **all**]	View storm control settings for all interfaces or the specified interface.

Configuring Protected Ports

Beginning in Privileged EXEC mode, use the following commands to add a name to a protected port group and add ports to the group.

Command	Purpose
configure	Enter global configuration mode.
switchport protected *groupid* **name** *name*	Specify a name for one of the three protected port groups. • *groupid* — Identifies which group the port is to be protected in. (Range: 0-2) • *name* — Name of the group. (Range: 0-32 characters)
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**.
switchport protected *groupid*	Add the interface to the specified protected port group.
CTRL + Z	Exit to Privileged EXEC mode.
show switchport protected	View protected group and port information.

Configuring LLPF

Beginning in Privileged EXEC mode, use the following commands to configure LLPF settings.

Command	Purpose
configure	Enter global configuration mode.
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
service-acl input {blockcdp | blockvtp | blockdtp | blockudld | blockpagp | blocksstp | blockall}	Use the appropriate keyword, or combination of keywords to block any (or all) of the following PDUs on the interface: • VTP • DTP • UDLD • PAgP • SSTP • All
CTRL + Z	Exit to Privileged EXEC mode.
show service-acl interface {*interface* | **all**}	View information about the blocked PDUs on the specified interface or all interfaces.

Port-Based Traffic Control Configuration Example

The commands in this example configure storm control, LLPF, and protected port settings for various interfaces on the switch.

The storm control configuration in this example sets thresholds on the switch so that if broadcast traffic occupies more than 10% on the bandwidth on any physical port, the interface blocks the broadcast traffic until the measured amount of this traffic drops below the threshold.

The LLPF configuration in this example disables all PAgP and VTP PDUs from being forwarded on any switch port or LAG.

The protected port configuration in this example prevents the clients connected to ports 3, 4, and 9 from being able to communicate with each other.

To configure the switch:

1 Configure storm control for broadcast traffic on all physical interfaces.

```
console(config)#interface range te1/0/1-24
console(config-if)#storm-control broadcast level 10
```

2 Configure LLPF to block PAgP and VTP PDUs on all physical interfaces.

```
console(config-if)#service-acl blockpagp blockvtp
console(config-if)#exit
```

3 Specify a name for protected port group 0.

```
console(config)#protected 0 name clients
```

4 Add the ports to the protected port group.

```
console(config)#interface te1/0/3
console(config-if-Te1/0/3)#switchport protected 0
console(config-if-Te1/0/3)#exit
console(config)#interface te1/0/4
console(config-if-Te1/0/4)#switchport protected 0
console(config-if-Te1/0/4)#exit
console(config)#interface te1/0/9
console(config-if-Te1/0/9)#switchport protected 0
console(config-if-Te1/0/9)#exit
console(config)#exit
```

5 Verify the configuration.

```
console#show storm-control te1/0/1

          Bcast    Bcast    Mcast    Mcast    Ucast    Ucast
```

```
 Intf   Mode    Level   Mode    Level   Mode    Level
------ ------- ------- ------- ------- ------- -------
Te1/0/1 Enable    10    Enable    5     Disable   5

console#show service-acl interface te1/0/1

Protocol         Mode
---------------  ----------
CDP              Disabled
VTP              Enabled
DTP              Disabled
UDLD             Disabled
PAGP             Enabled
SSTP             Disabled
ALL              Disabled

console#show switchport protected 0
Name......................................... "clients"


Member Ports: Te1/0/1, Te1/0/2, Te1/0/3, Te1/0/4, Te1/0/9
```

25

Layer-2 Multicast Features

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes the layer-2 (L2) multicast features on the Dell Networking N-Series switches. The features this chapter describes include bridge multicast flooding and forwarding, Internet Group Management Protocol (IGMP) snooping, Multicast Listener Discovery (MLD) snooping, and Multicast VLAN Registration (MVR).

The topics covered in this chapter include:

- L2 Multicast Overview
- Snooping Switch Restrictions
- Default L2 Multicast Values
- Configuring L2 Multicast Features (Web)
- Configuring L2 Multicast Features (CLI)
- Case Study on a Real-World Network Topology

L2 Multicast Overview

Multicast traffic is traffic from one source that has multiple destinations. The L2 multicast features on the switch help control network flooding of Ethernet multicast and IP multicast traffic by keeping track of multicast group membership. It is essential that a multicast router be connected to a Dell Networking layer-2 multicast switch for IGMP/MLD snooping to operate properly. The presence of a multicast router allows the snooping switch to relay IGMP reports to the router and to forward multicast data sources to the multicast router as well as restrict flooding of multicast sources in a VLAN.

Multicast Flooding and Forwarding

Flooding behavior is to send incoming multicast packets to all ports in the VLAN other than the ingress port. Forwarding behavior is to send incoming multicast packets to selected ports in the VLAN. Forwarding behavior is desirable as it reduces the network load by sending packets only to other hosts/switches/routers that have indicated an interest in receiving the multicast.

If L2 snooping is not enabled, multicast packets are flooded in the ingress VLAN.

What Are the Multicast Bridging Features?

The Dell Networking N-Series switches support multicast forwarding and multicast flooding. For multicast traffic, the switch uses a database called the layer-2 Multicast Forwarding Database (MFDB) to make forwarding decisions for packets that arrive with a multicast destination MAC address. By limiting multicasts to only certain ports in the switch, traffic is prevented from going to parts of the network where that traffic is unnecessary.

When a packet enters the switch, the destination MAC address is combined with the VLAN ID, and a search is performed in the layer-2 MFDB. If no match is found, then the packet is flooded. If a match is found, then the packet is forwarded only to the ports that are members of that multicast group within the VLAN.

Multicast traffic destined to well-known (reserved) multicast IP addresses (control plane traffic) is always flooded to all ports in the VLAN. The well-known IP multicast addresses are 224.0.0.x for IPv4 and FF0x:: for IPv6.

By default IGMP/MLD snooping is enabled and multicast data traffic is flooded to all ports in the VLAN if no multicast router ports have been identified. Once a multicast router port is identified, multicast data traffic is forwarded to the multicast router ports. The MFDB is populated by snooping the membership reports sent to the multicast routers. This causes multicast data traffic to be forwarded to any hosts joining the multicast group. Enabling multicast routing on the switch internally enables an mrouter port, and snooping will forward multicast to hosts joining the group instead of flooding it in the VLAN.

It is possible to statically define an mrouter port. This causes IGMP/MLD snooping to forward multicast data traffic to hosts from which it has received membership reports. This behavior exists even if the mrouter port is not enabled.

What Is L2 Multicast Traffic?

L3 IP multicast traffic is traffic that is destined to a host group. Host groups are identified by class D IPv4 addresses, which range from 224.0.1.0 to 239.255.255.255, or by FF0x:: or FF3x:: IPv6 addresses. In contrast to L3 multicast traffic, layer-2 multicast traffic is identified by the MAC address, i.e., the range 01:00:5e:00:00:00 to 01:00:5e:7f:ff:ff:ff for IPv4 multicast traffic or 33:33:xx:xx:xx:xx for IPv6 multicast traffic.

When a packet with a broadcast or multicast destination MAC address is received, the switch will flood a copy into each of the remaining network segments in accordance with the IEEE MAC Bridge standard. Eventually, the packet is made accessible to all nodes connected to the network.

This approach works well for broadcast packets that are intended to be seen or processed by all connected nodes. In the case of multicast packets, however, this approach could lead to less efficient use of network bandwidth, particularly when the packet is intended for only a small number of nodes. Packets will be flooded into network segments where no node has any interest in receiving the packet.

What Is IGMP Snooping?

IGMP snooping allows the switch to snoop on IGMP exchanges between hosts and multicast routers and perform multicast forwarding within a VLAN. The IGMP snooping feature complies with RFC 4541. When a switch "sees" an IGMP report from a host for a given multicast address, the switch adds the host's interface/VLAN to the L2 multicast group forwarding table and floods the report to all ports in the VLAN. When the switch sees a leave message for the group, it removes the host interface/VLAN from the L2 multicast group forwarding table.

IGMP snooping learns about multicast routers by listening for the following messages:

- An IGMP query packet.
- PIMv1 (IGMP type 0x14) packets with destination IP address 224.0.0.13.
- DVMRP (IGMP type 0x13) packets with destination IP address 224.0.0.4.
- PIMv2 (IP protocol type 0x67) packets with destination IP address 224.0.0.13.

Group addresses that fall into the range 224.0.0.x are never pruned by IGMP snooping—they are always flooded to all ports in the VLAN. Note that this flooding is based on the IP address, not the corresponding 01-00-5e-00-00-xx MAC address.

When a multicast router is discovered (or locally configured on the switch), its interface is added to the interface distribution list for all multicast groups in the VLAN. If a switch is connected to a multicast source and no client, the switch filters the traffic from that group to all interfaces in the VLAN. If the switch sees an IGMP join from a host in the same VLAN, then it forwards the traffic to the host. Likewise, if the switch sees a multicast router in the VLAN, it forwards the group to the multicast router and does not flood in the VLAN. If snooping is disabled, the switch always floods multicast data and control plane packets in the VLAN.

A multicast router can also be statically configured, either by configuring a port as a static L2 mrouter port in the VLAN, or by enabling L3 multicast routing in the switch. If a port is configured as a static L2 mrouter port, IGMP snooping forwards multicast data plane packets in the VLAN regardless of the interface state of the port.

By default, dynamically discovered multicast routers are aged out every five minutes. The user can control whether or not multicast routers age out. If all multicast routers age out, the switch floods the VLAN with any received multicast groups.

Multicast routers send an IGMP query every 60 seconds. This query is intercepted by the switch and forwarded to all ports in the VLAN. All hosts that are members of the group answer that query. The switch intercepts the replies and forwards only one report per group from all of the received responses.

In summary:

- IGMP snooping controls the flooding/forwarding behavior for multicast groups. Multicast data is flooded in the VLAN until a multicast router port is identified.
- IGMP snooping is enabled by default
- IGMP snooping forwards multicast sources to multicast routers by default
- Reserved multicast IP addresses (224.0.0.x) are always flooded to all ports in the VLAN

- Unregistered multicast traffic may be flooded in the VLAN by a user configuration option.

NOTE: It is strongly recommended that operators enable MLD snooping if IGMP snooping is enabled and vice-versa. This is because both IGMP snooping and MLD snooping utilize the same forwarding table. Not enabling both may cause unwanted pruning of protocol packets utilized by other protocols, e.g. OSPFv3.

NOTE: IGMP snooping (and IGMP querier) validates IGMP packets. As part of the validation, IGMP checks for the router alert option. If other devices in the network do not send IGMP packets with the router alert option, IGMP snooping (and snooping querier) will discard the packet. Use the **no ip igmp snooping router-alert-check** command to disable checking for the router alert option.

IGMP Snooping Querier

When PIM and IGMP are enabled in a network with IP multicast routing, the IP multicast router acts as the IGMP querier. However, if the IP-multicast traffic in a VLAN needs to be layer-2 switched only, an IP-multicast router is not required. The IGMP snooping querier can perform the role of generating IGMP queries that would normally be performed by the multicast router.

NOTE: Without an IP-multicast router on a VLAN, you must configure another switch as the IGMP querier so that it can send queries.

When IGMP snooping querier is enabled, the querier switch sends out periodic IGMP queries that trigger IGMP report messages from the hosts that want to receive IP multicast traffic. The IGMP snooping feature listens to these IGMP reports to identify multicast router ports. If there is another querier in the network and the local querier is in election mode, then the querier with the lower IP address is elected and the other querier stops querying. If the local querier is not in election mode and another querier is detected, the local querier stops querying.

What Is MLD Snooping?

In IPv4, layer-2 switches use IGMP snooping to limit the flooding of multicast traffic by dynamically configuring the multicast forwarding database so that multicast data traffic is forwarded to only those ports associated with a multicast router or host that has indicated an interest in receiving a particular multicast group. In IPv6, MLD snooping performs a similar function.

With MLD snooping, IPv6 multicast data is selectively forwarded to a list of ports that want to receive the data instead of being flooded to all ports in a VLAN. This list is constructed in the MFDB by snooping IPv6 multicast control packets. MLD snooping floods multicast data packets until a multicast router port has been identified. MLD snooping forwards unregistered multicast data packets to IPv6 multicast routers. MLD snooping discovers multicast routers by listening for MLD queries and populates the MFDB.

MLD Snooping learns of multicast routers by listening for the following packets:

- MLD query packets
- PIMv2 hello packets with destination IP address as FF02::D

Dynamically learned multicast routers are timed out after an administrator-configurable period of time.

MLD is a protocol used by IPv6 multicast routers to discover the presence of multicast listeners (nodes wishing to receive IPv6 multicast packets) on its directly-attached links and to discover which multicast packets are of interest to neighboring nodes. MLD is derived from IGMP; MLD version 1 (MLDv1) is equivalent to IGMPv2, and MLD version 2 (MLDv2) is equivalent to IGMPv3. MLD is a subprotocol of Internet Control Message Protocol version 6 (ICMPv6), and MLD messages are a subset of ICMPv6 messages.

A multicast router can also be statically configured, either by configuring a port as a static L2 mrouter port in the VLAN, or by enabling L3 multicast routing in the switch. If a port is configured as a static L2 mrouter port, IGMP/MLD snooping forwards packets regardless of the interface state of the port.

The switch snoops both MLDv1 and MLDv2 protocol packets and forwards IPv6 multicast data based on destination IPv6 multicast MAC addresses (33:33::). The switch floods multicast control plane traffic addressed to the permanently assigned (well-known) multicast address FF0x::/8 to all ports in the VLAN, except for MLD packets, which are handled according the MLD snooping rules.

NOTE: It is strongly recommended that users enable IGMP snooping if MLD snooping is enabled and vice-versa. This is because both IGMP snooping and MLD snooping utilize the same forwarding table, and not enabling both may cause unwanted pruning of protocol packets utilized by other protocols, e.g. OSPFv2.

What Is Multicast VLAN Registration?

IGMP snooping helps limit multicast traffic when member ports are in the same VLAN; however, when ports belong to different VLANs, a copy of the multicast stream is sent to each VLAN that has member ports in the multicast group. MVR eliminates the need to duplicate the multicast traffic when multicast group member ports belong to different VLANs.

MVR uses a dedicated multicast VLAN to forward multicast traffic over the L2 network. Only one MVLAN can be configured per switch, and it is used only for certain multicast traffic, such as traffic from an IPTV application, to avoid duplication of multicast streams for clients in different VLANs. Clients can dynamically join or leave the mutlicast VLAN without interfering with their membership in other VLANs.

MVR, like IGMP snooping, allows a layer-2 switch to listen to IGMP messages to learn about multicast group membership.

There are two types of MVR ports: source and receiver.

- Source port is the port where multicast traffic is flowing to. It has to be the member of so called multicast VLAN.
- Receiver port is the port where listening host is connected to the switch. It can be the member of any VLAN, except multicast VLAN.

There are two configured learning modes of the MVR operation: dynamic and compatible.

- In the dynamic mode MVR learns existent multicast groups by parsing the IGMP queries from router on source ports and forwarding the IGMP joins from the hosts to the router.
- In the compatible mode MVR does not learn multicast groups, but they have to be configured by administrator and protocol does not forward joins from the hosts to the router. To work in this mode the IGMP router has to be configured to transmit required multicast streams to the network with the MVR switch.

Enabling MVR and IGMP on the Same Interface

MVR and IGMP snooping operate independently and can both be enabled on an interface. When both MVR and IGMP snooping are enabled, MVR listens to the IGMP join and report messages for static multicast group information, and IGMP snooping manages dynamic multicast groups.

When Are Layer-3 Multicast Features Required?

In addition to L2 multicast features, the switch suports IPv4 and IPv6 multicast features. You configure the IPv4/IPv6 multicast features if the switch functions as a multicast router that can route multicast traffic between VLAN routing interfaces. In this case, you must enable a multicast routing protocol on the switch, such as PIM-SM. For information about layer-3 multicast features, see "IPv4 and IPv6 Multicast " on page 1509.

If the switch functions as a multicast router, it is possible to enable IGMP so that IGMP forwards multicast traffic for directly connected hosts between VLANs. It is recommended that IGMP snooping and MLD snooping be enabled in L3 multicast routed networks, as this allows the switch to limit multicast flooding in multi-access routed VLANs based as controlled by IGMP snooping.

NOTE: If MVR is enabled, IP Multicast should be disabled. Multicast routing and MVR cannot coexist on a switch.

For information about configuring Dell Networking N1500, N2000, N3000, and N4000 Series switches as a multicast router that also performs IGMP snooping, see "Configuring Multicast VLAN Routing With IGMP and PIM-SM " on page 1590.

What Are GARP and GMRP?

Generic Attribute Registration Protocol (GARP) is a general-purpose protocol that registers any network connectivity or membership-style information. GARP defines a set of switches interested in a given network attribute, such as VLAN ID or multicast address.

Dell Networking N-Series switches can use GARP functionality for two applications:

- GARP VLAN Registration Protocol (GVRP) to help dynamically manage VLAN memberships on trunk ports.

- GARP Multicast Registration Protocol (GMRP) to help control the flooding of multicast traffic by keeping track of group membership information.

GVRP and GMRP use the same set of GARP Timers to specify the amount of time to wait before transmitting various GARP messages.

GMRP is similar to IGMP snooping in its purpose, but IGMP snooping is more widely used. GMRP must be running on both the host and the switch to function properly and IGMP/MLD snooping must be disabled on the switch, as IGMP snooping and GMRP cannot simultaneously operate within the same VLAN.

Snooping Switch Restrictions

MAC Address-Based Multicast Group

The L2 multicast forwarding table consists of the Multicast group MAC address filtering entries. For IPv4 multicast groups, 16 IP multicast group addresses map to the same multicast MAC address. For example, 224.1.1.1 and 225.1.1.1 map to the MAC address 01:00:5E:01:01:01, and IP addresses in the range [224-239].3.3.3 map to 01:00:5E:03:03:03. As a result, if a host requests 225.1.1.1, then it might receive multicast traffic of group 226.1.1.1 as well.

Topologies Where the Multicast Source Is Not Directly Connected to the Querier

If the multicast source is not directly connected to a multicast querier, the multicast stream is forwarded to any router ports on the switch (within the VLAN). Because multicast router queries are flooded to all ports in the VLAN, intermediate IGMP snooping switches will receive the multicast stream from the multicast source and forward it to the multicast router.

Using Static Multicast MAC Configuration

If configuring static multicast MAC group addresses on a port in a VLAN, it is necessary to configure all ports in the VLAN over which it is desired that the group traffic flow (both host and router) on all switches. IGMP snooping does not dynamically add ports to a VLAN for a multicast group when a static entry is configured for that group in the VLAN. This restriction applies to both multicast router-connected ports and host-connected ports.

IGMP Snooping and GMRP

IGMP snooping and GMRP are not compatible. Only one of IGMP snooping or GMRP should be configured to filter multicast groups for any VLAN. Simultaneous operation of GMRP and IMGP snooping is not supported and will lead to undesirable results, such as flooding in the VLAN due to the inability to identify multicast router ports.

Default L2 Multicast Values

Details about the L2 multicast are in Table 25-1.

Table 25-1. L2 Multicast Defaults

Parameter	Default Value
IGMP Snooping mode	Enabled
MLD Snooping mode	Enabled
Bridge multicast group	None configured
IGMP/MLD snooping	Enabled on all VLANs
IGMP/MLD snooping auto-learn	Disabled
IGMP/MLD snooping host timeout	260 seconds
IGMP/MLD snooping multicast router timeout	300 seconds
IGMP/MLD snooping leave timeout	10 seconds
IGMP snooping querier	Disabled
IGMP version	v2
MLD version	v1
IGMP/MLD snooping querier query interval	60 seconds
IGMP/MLD snooping querier expiry interval	60 seconds
IGMP/MLD snooping VLAN querier	Disabled
VLAN querier election participate mode	Disabled
Snooping Querier VLAN Address	0.0.0.0
MVR running	Disabled
MVR multicast VLAN	1
MVR max multicast groups	256
MVR Global query response time	5 tenths of a second
MVR Mode	Compatible
GARP Leave Timer	60 centiseconds
GARP Leave All Timer	1000 centiseconds
GARP Join Timer	20 centiseconds

Table 25-1. L2 Multicast Defaults (Continued)

Parameter	Default Value
GMRP	Disabled globally and per-interface

Configuring L2 Multicast Features (Web)

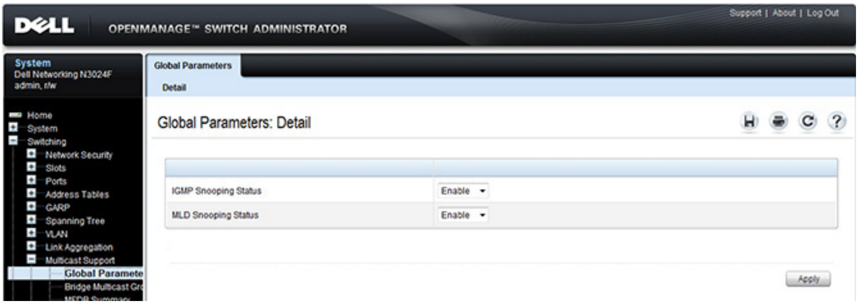
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring L2 multicast features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

Multicast Global Parameters

Use the **Multicast Global Parameters** page to enable or disable IGMP snooping, or MLD snooping on the switch.

To display the **Multicast Global Parameters** page, click **Switching** → **Multicast Support** → **Global Parameters** in the navigation menu.

Figure 25-1. Multicast Global Parameters



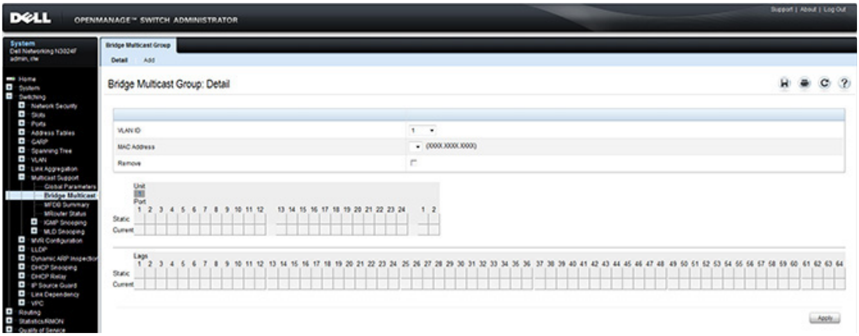
NOTE: It is strongly recommended that users enable IGMP snooping if MLD snooping is enabled and vice-versa. This is because both IGMP snooping and MLD snooping utilize the same forwarding table, and not enabling both may cause unwanted pruning of protocol packets utilized by other protocols, e.g. OSPFv2.

Bridge Multicast Group

Use the **Bridge Multicast Group** page to create new multicast service groups or to modify ports and LAGs assigned to existing multicast service groups. Attached interfaces display in the Port and LAG tables and reflect the manner in which each is joined to the Multicast group.

To display the **Bridge Multicast Group** page, click **Switching** → **Multicast Support** → **Bridge Multicast Group** in the navigation menu.

Figure 25-2. Bridge Multicast Group



Understanding the Port and LAG Member Tables

The **Bridge Multicast Group** tables display which Ports and LAGs are members of the multicast group, and whether they're static (S), dynamic (D), or forbidden (F). The tables have two rows: **Static** and **Current**. Only the **Static** row is accessible from this page. The **Current** row is updated when the **Static** row is changed and **Apply** is clicked.

The **Bridge Multicast Group** page contains two editable tables:

- **Unit and Ports** — Displays and assigns multicast group membership to ports. To assign membership, click in **Static** for a specific port. Each click toggles between S, F, and blank. See Table 25-2 for definitions.
- **LAGs** — Displays and assigns multicast group membership to LAGs. To assign membership, click in **Static** for a specific LAG. Each click toggles between S, F, and blank. See Table 25-2 for definitions.

Table 25-2 contains definitions for port/LAG IGMP management settings.

Table 25-2. Port/LAG IGMP Management Settings

Port Control	Definition
D	Dynamic: Indicates that the port/LAG was dynamically joined to the Multicast group (displays in the *Current* row).
S	Static: Attaches the port to the Multicast group as a static member in the *Static* row. Displays in the *Current* row once **Apply** is clicked.
F	Forbidden: Indicates that the port/LAG is forbidden entry into the Multicast group in the *Static* row. Displays in the *Current* row once **Apply** is clicked.
Blank	Blank: Indicates that the port is not attached to a Multicast group.

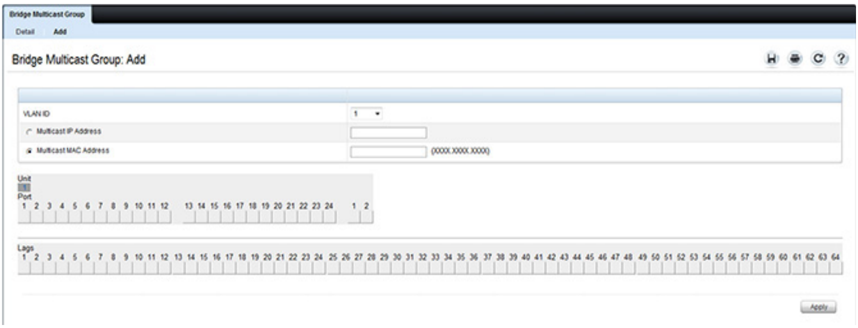
Adding and Configuring Bridge Multicast Address Groups

To configure a bridge multicast group:

1 From the **Bridge Multicast Group** page, click **Add**.

The **Add Bridge Multicast Group** page displays.

Figure 25-3. Add Bridge Multicast Group



2 Select the ID of the VLAN to add to the multicast group or to modify membership for an existing group.

3 For a new group, specify the multicast group IP or MAC address associated with the selected VLAN.

4 In the **Bridge Multicast Group** tables, assign a setting by clicking in the **Static** row for a specific port/LAG. Each click toggles between S, F, and blank. (not a member).

5 Click **Apply**.

The bridge multicast address is assigned to the multicast group, ports/LAGs are assigned to the group (with the **Current** rows being updated with the **Static** settings), and the switch is updated.

Removing a Bridge Multicast Group

To delete a bridge multicast group:

1 Open the **Bridge Multicast Group** page.

2 Select the **VLAN ID** associated with the bridge multicast group to be removed from the drop-down menu.

The **Bridge Multicast Address** and the assigned ports/LAGs display.

3 Check the **Remove** check box.

4 Click **Apply**.

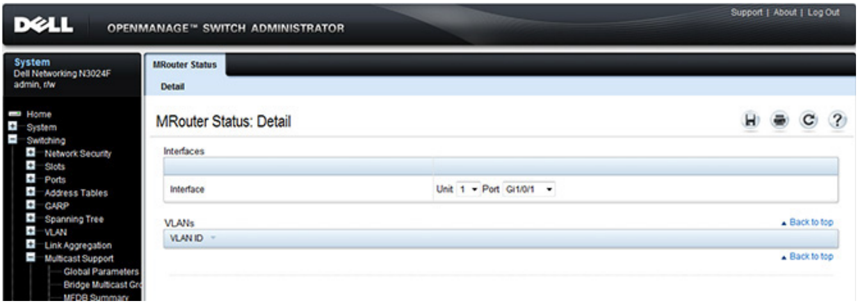
The selected bridge multicast group is removed, and the device is updated.

MRouter Status

Use the **MRouter Status** page to display the status of dynamically learned multicast router interfaces.

To access this page, click **Switching** → **Multicast Support** → **MRouter Status** in the navigation panel.

Figure 25-4. MRouter Status

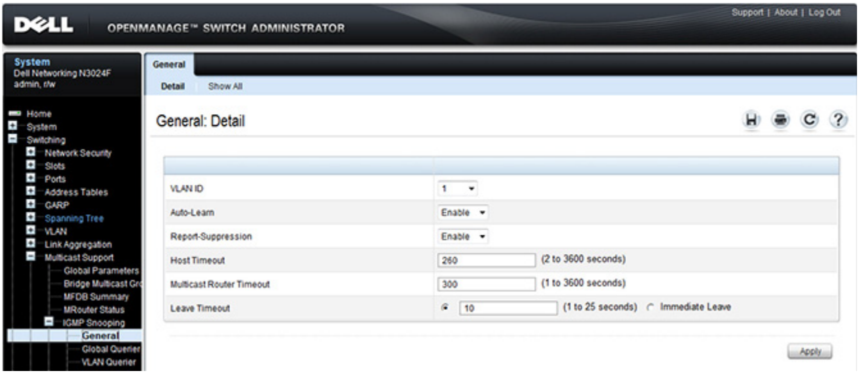


General IGMP Snooping

Use the **General** IGMP snooping page to configure IGMP snooping settings on specific ports and LAGs.

To display the **General** IGMP snooping page, click **Switching** → **Multicast Support** → **IGMP Snooping** → **General** in the navigation menu.

Figure 25-5. General IGMP Snooping



Modifying IGMP Snooping Settings for Multiple Ports, LAGs, or VLANs

To modify the IGMP snooping settings:

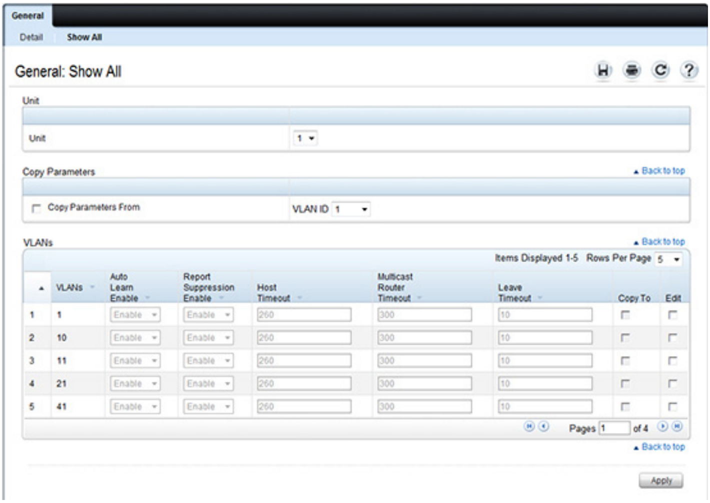
1. From the **General** IGMP snooping page, click **Show All**.

   The **IGMP Snooping Table** displays.

2. Select the **Edit** checkbox for each Port, LAG, or VLAN to modify.

   In Figure 25-6, 2 and 3 are to be modified.

Figure 25-6. Edit IGMP Snooping Settings



3 Edit the IGMP snooping fields as needed.

4 Click **Apply**.

The IGMP snooping settings are modified, and the device is updated.

Copying IGMP Snooping Settings to Multiple Ports, LAGs, or VLANs

To copy IGMP snooping settings:

1 From the **General** IGMP snooping page, click **Show All**.

The **IGMP Snooping Table** displays.

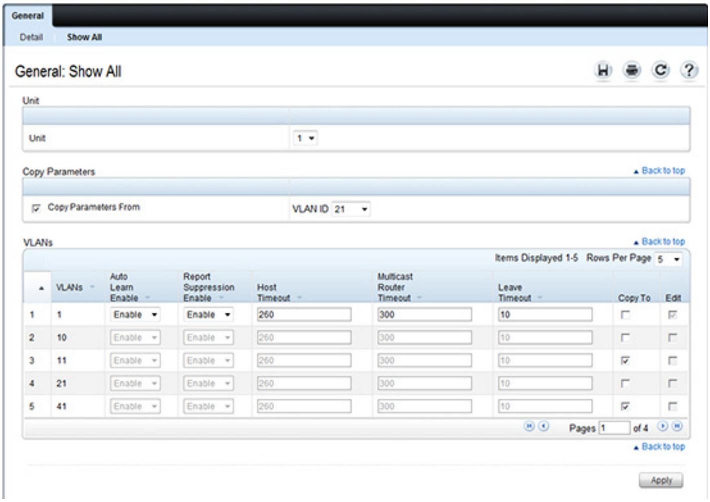
2 Select the **Copy Parameters From** checkbox.

3 Select a Unit/Port, LAG, or VLAN to use as the source of the desired parameters.

4 Select the **Copy To** checkbox for the Unit/Ports, LAGs, or VLANs that these parameters will be copied to.

In Figure 25-7, the settings for VLAN 21 will be copied to ports 3 and 5.

Figure 25-7. Copy IGMP Snooping Settings



5 Click **Apply**.

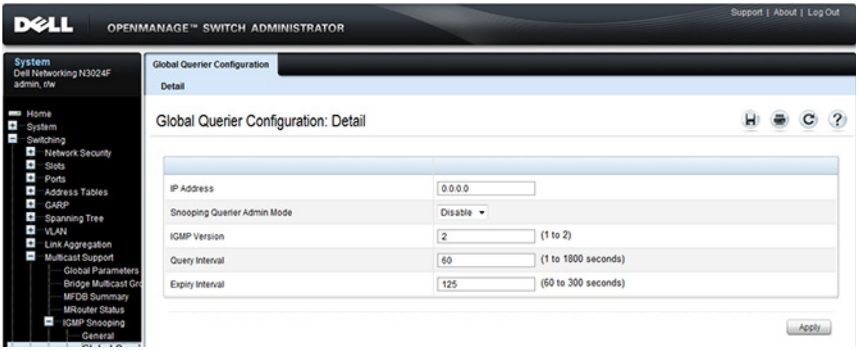
The IGMP snooping settings are modified, and the device is updated.

Global Querier Configuration

Use the **Global Querier Configuration** page to configure IGMP snooping querier settings, such as the IP address to use as the source in periodic IGMP queries when no source address has been configured on the VLAN.

To display the **Global Querier Configuration** page, click **Switching** → **Multicast Support** → **IGMP Snooping** → **Global Querier Configuration** in the navigation menu.

Figure 25-8. Global Querier Configuration

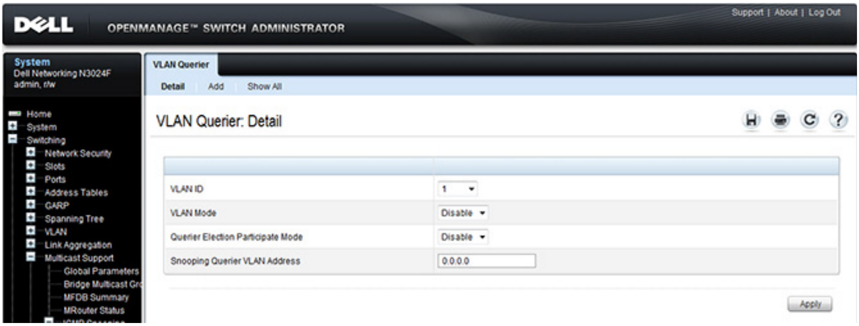


VLAN Querier

Use the **VLAN Querier** page to specify the IGMP snooping querier settings for individual VLANs.

To display the **VLAN Querier** page, click **Switching** → **Multicast Support** → **IGMP Snooping** → **VLAN Querier** in the navigation menu.

Figure 25-9. VLAN Querier



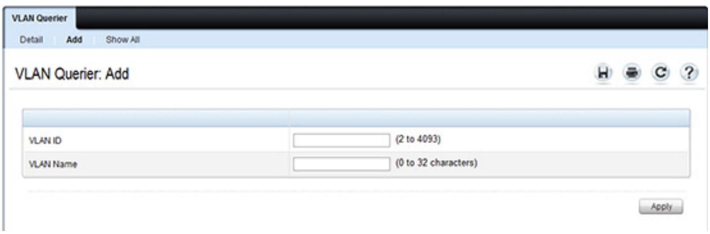
Adding a New VLAN and Configuring its VLAN Querier Settings

To configure a VLAN querier:

1 From the **VLAN Querier** page, click **Add**.

The page refreshes, and the **Add VLAN** page displays.

Figure 25-10. Add VLAN Querier



2 Enter the VLAN ID and, if desired, an optional VLAN name.

3 Return to the **VLAN Querier** page and select the new VLAN from the **VLAN ID** menu.

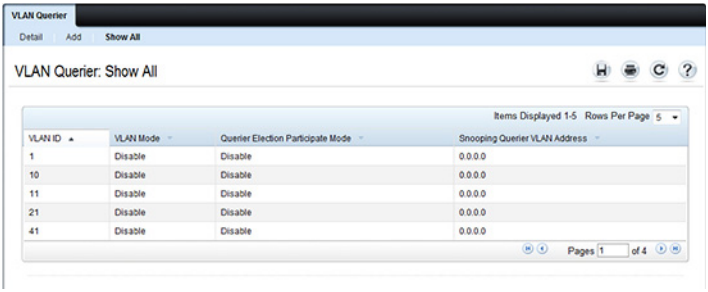
4 Specify the VLAN querier settings.

5 Click **Apply**.

The VLAN Querier settings are modified, and the device is updated.

To view a summary of the IGMP snooping VLAN querier settings for all VLANs on the switch, click **Show All**.

Figure 25-11. Add VLAN Querier

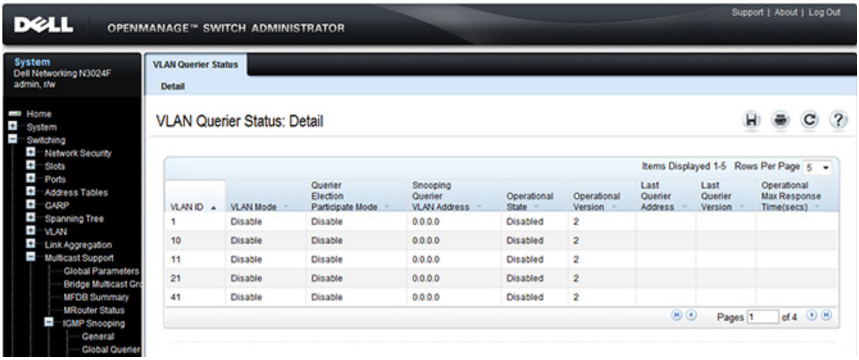


VLAN Querier Status

Use the **VLAN Querier Status** page to view the IGMP snooping querier settings for individual VLANs.

To display the **VLAN Querier Status** page, click **Switching** → **Multicast Support** → **IGMP Snooping** → **VLAN Querier Status** in the navigation menu.

Figure 25-12. IGMP Snooping VLAN Querier Status

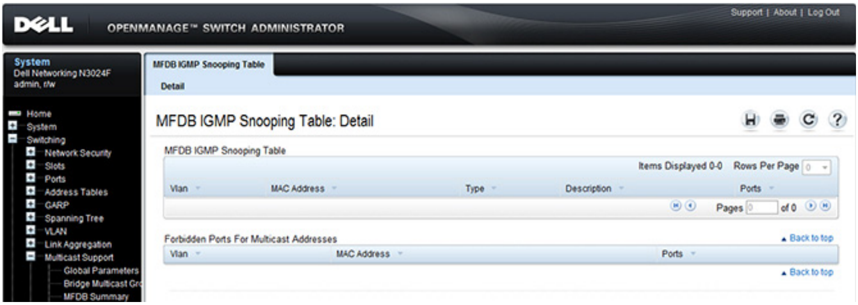


MFDB IGMP Snooping Table

Use the **MFDB IGMP Snooping Table** page to view the multicast forwarding database (MFDB) IGMP Snooping Table and Forbidden Ports settings for individual VLANs.

To display the **MFDB IGMP Snooping Table** page, click **Switching** → **Multicast Support** → **IGMP Snooping** → **MFDB IGMP Snooping Table** in the navigation menu.

Figure 25-13. MFDB IGMP Snooping Table

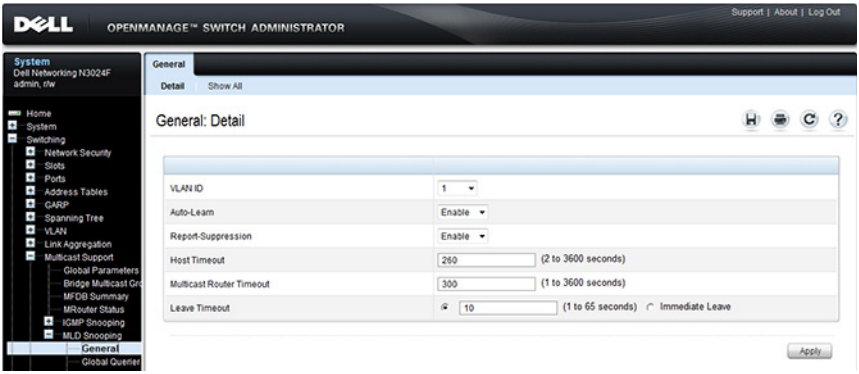


MLD Snooping General

Use the MLD Snooping **General** page to add MLD members.

To access this page, click **Switching** → **Multicast Support** → **MLD Snooping** → **General** in the navigation panel.

Figure 25-14. MLD Snooping General



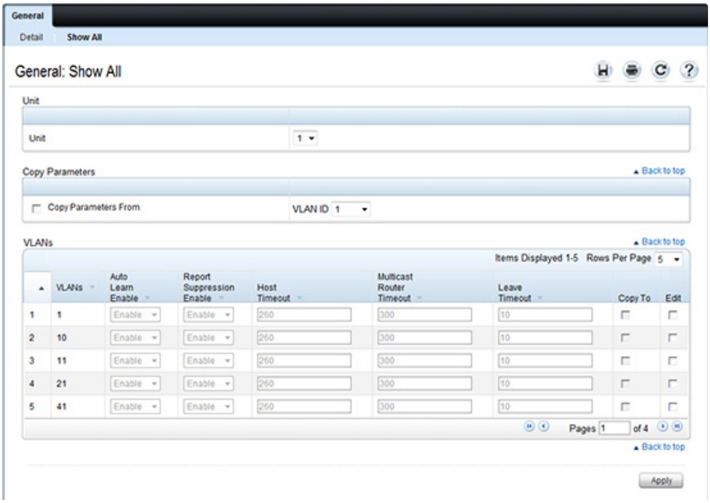
Modifying MLD Snooping Settings for VLANs

To configure MLD snooping:

1 From the **General** MLD snooping page, click **Show All**.

The **MLD Snooping Table** displays.

Figure 25-15. MLD Snooping Table



2 Select the **Edit** checkbox for each VLAN to modify.

3 Edit the MLD snooping fields as needed.

4 Click **Apply**.

The MLD snooping settings are modified, and the device is updated.

Copying MLD Snooping Settings to VLANs

To copy MLD snooping settings:

1 From the **General** MLD snooping page, click **Show All**.

  The **MLD Snooping Table** displays.

2 Select the **Copy Parameters From** checkbox.

3 Select a VLAN to use as the source of the desired parameters.

4 Select the **Copy To** checkbox for the VLANs that these parameters will be copied to.

5 Click **Apply**.

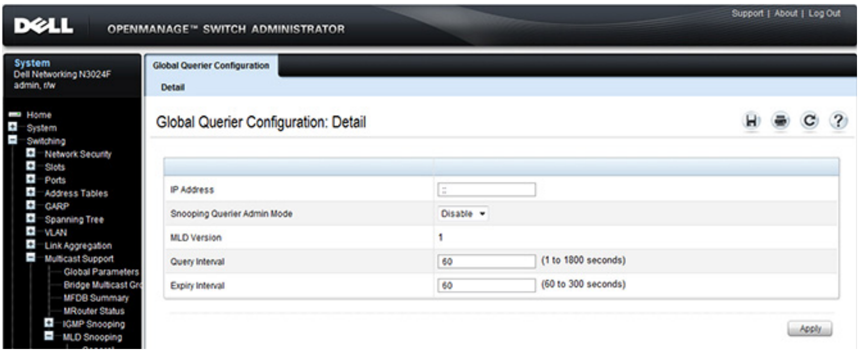
  The MLD snooping settings are modified, and the device is updated.

MLD Snooping Global Querier Configuration

Use the MLD Snooping **Global Querier Configuration** page to configure the parameters for the MLD snooping querier.

To display the **Global Querier Configuration** page, click **Switching** → **Multicast Support** → **MLD Snooping** → **Global Querier Configuration** in the navigation menu.

Figure 25-16. MLD Snooping Global Querier Configuration

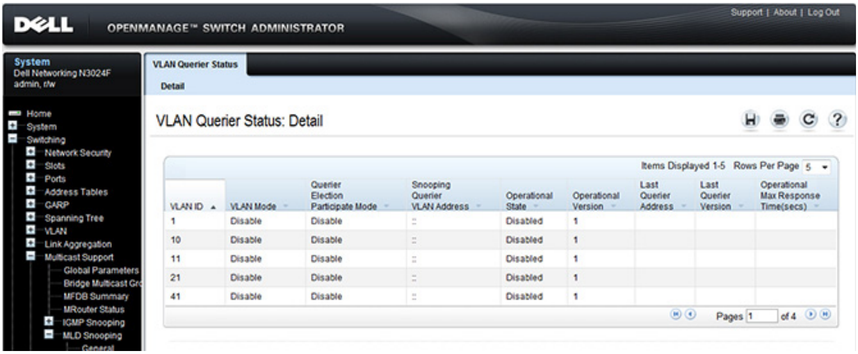


MLD Snooping VLAN Querier

Use the MLD Snooping **VLAN Querier** page to specify the MLD snooping querier settings for individual VLANs.

To display the MLD Snooping **VLAN Querier** page, click **Switching** → **Multicast Support** → **MLD Snooping** → **VLAN Querier** in the navigation menu.

Figure 25-17. MLD Snooping VLAN Querier



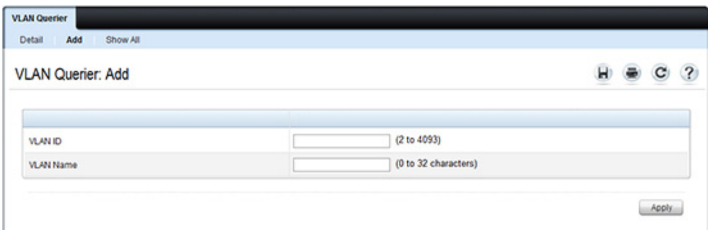
Adding a New VLAN and Configuring its MLD Snooping VLAN Querier Settings

To configure an MLD snooping VLAN querier:

1 From the **VLAN Querier** page, click **Add**.

The page refreshes, and the **Add VLAN** page displays.

Figure 25-18. Add MLD Snooping VLAN Querier



2 Enter the VLAN ID and, if desired, an optional VLAN name.

3 Return to the **VLAN Querier** page and select the new VLAN from the **VLAN ID** menu.

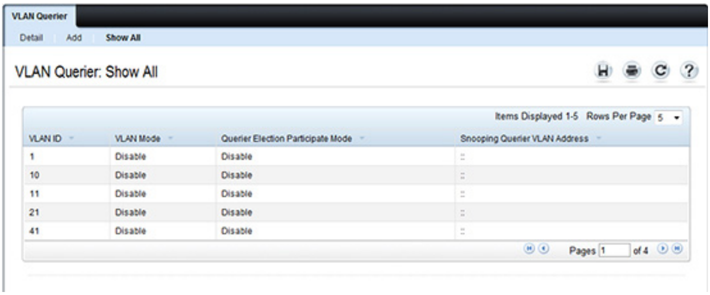
4 Specify the VLAN querier settings.

5 Click **Apply**.

The VLAN Querier settings are modified, and the device is updated.

To view a summary of the IGMP snooping VLAN querier settings for all VLANs on the switch, click **Show All**.

Figure 25-19. Add VLAN Querier

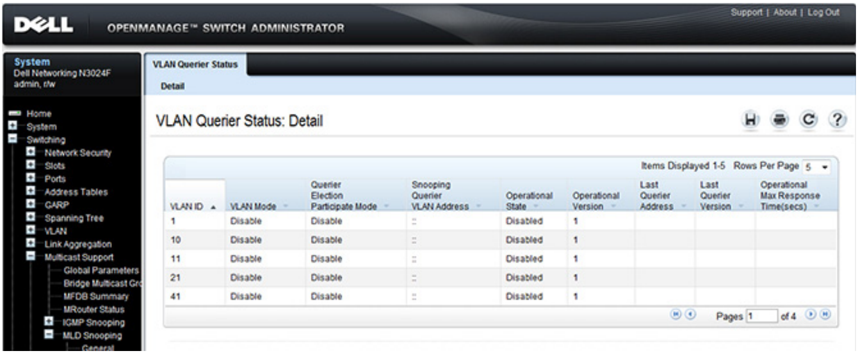


MLD Snooping VLAN Querier Status

Use the **VLAN Querier Status** page to view the MLD snooping querier settings for individual VLANs.

To display the **VLAN Querier Status** page, click **Switching** → **Multicast Support** → **MLD Snooping** → **VLAN Querier Status** in the navigation menu.

Figure 25-20. MLD Snooping VLAN Querier Status

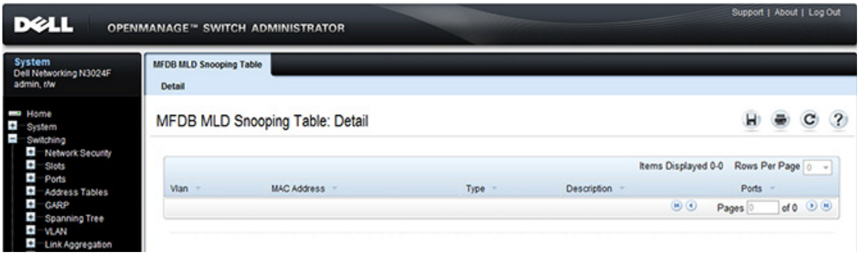


MFDB MLD Snooping Table

Use the **MFDB MLD Snooping Table** page to view the MFDB MLD snooping table settings for individual VLANs.

To display the **MFDB MLD Snooping Table** page, click **Switching** → **Multicast Support** → **MLD Snooping** → **MFDB MLD Snooping Table** in the navigation menu.

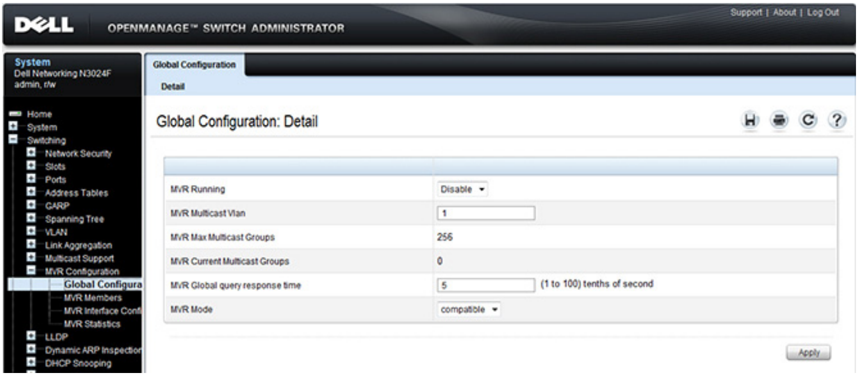
Figure 25-21. MFDB MLD Snooping Table



MVR Global Configuration

Use the **MVR Global Configuration** page to enable the MVR feature and configure global parameters. To display the **MVR Global Configuration** page, click **Switching** → **MVR Configuration** → **Global Configuration** in the navigation panel.

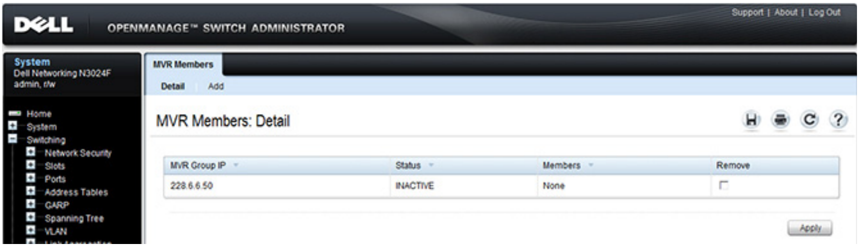
Figure 25-22. MVR Global Configuration



MVR Members

Use the MVR Members page to view and configure MVR group members. To display the MVR Members page, click **Switching** → **MVR Configuration** → **MVR Members** in the navigation panel.

Figure 25-23. MVR Members



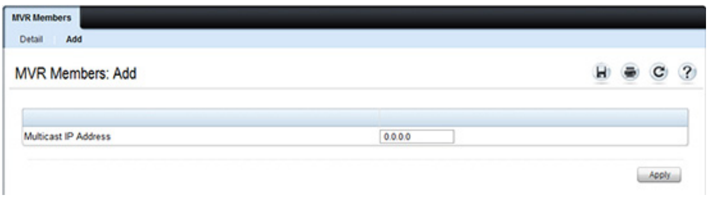
Adding an MVR Membership Group

To add an MVR membership group:

1. From the **MVR Membership** page, click **Add**.

   The MVR **Add Group** page displays.

Figure 25-24. MVR Member Group

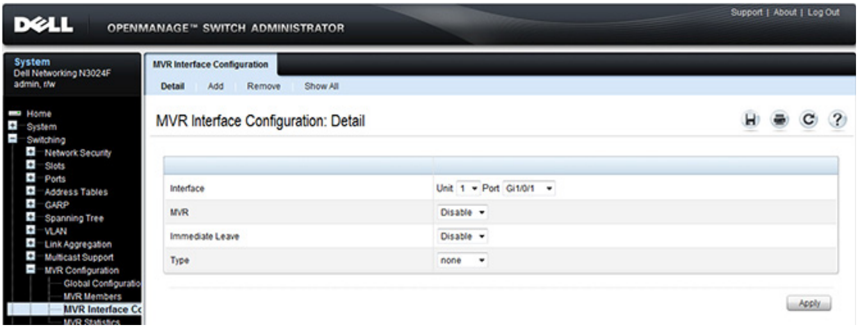


2. Specify the MVR group IP multicast address.
3. Click **Apply**.

MVR Interface Configuration

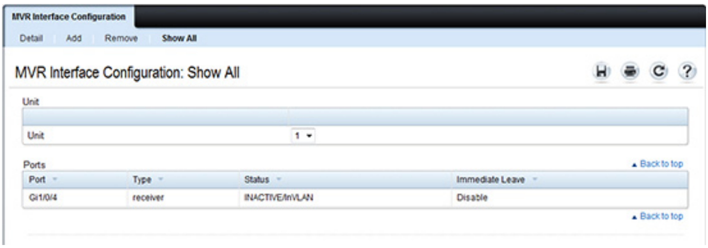
Use the MVR Interface Configuration page to enable MVR on a port, configure its MVR settings, and add the port to an MVR group. To display the MVR Interface Configuration page, click **Switching MVR Configuration → MVR Interface Configuration** in the navigation panel.

Figure 25-25. MVR Interface Configuration



To view a summary of the MVR interface configuration, click **Show All**.

Figure 25-26. MVR Interface Summary

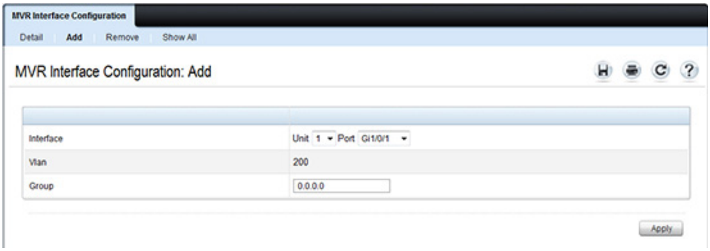


Adding an Interface to an MVR Group

To add an interface to an MVR group:

1 From the **MVR Interface** page, click **Add**.

Figure 25-27. MVR - Add to Group



2 Select the interface to add to the MVR group.

3 Specify the MVR group IP multicast address.

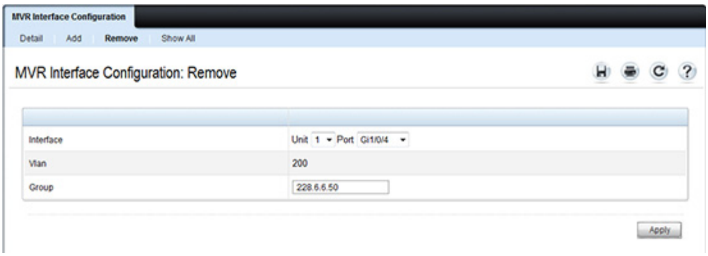
4 Click **Apply**.

Removing an Interface from an MVR Group

To remove an interface from an MVR group:

1 From the **MVR Interface** page, click **Remove**.

Figure 25-28. MVR - Remove from Group



2 Select the interface to remove from an MVR group.

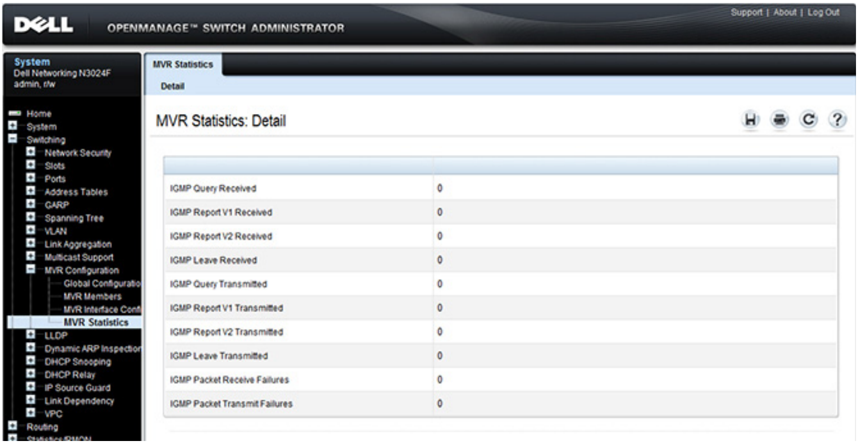
3 Specify the IP multicast address of the MVR group.

4 Click **Apply**.

MVR Statistics

Use the **MVR Statistics** page to view MVR statistics on the switch. To display the **MVR Statistics** page, click **Switching MVR Configuration** → **MVR Statistics** in the navigation panel.

Figure 25-29. MVR Statistics

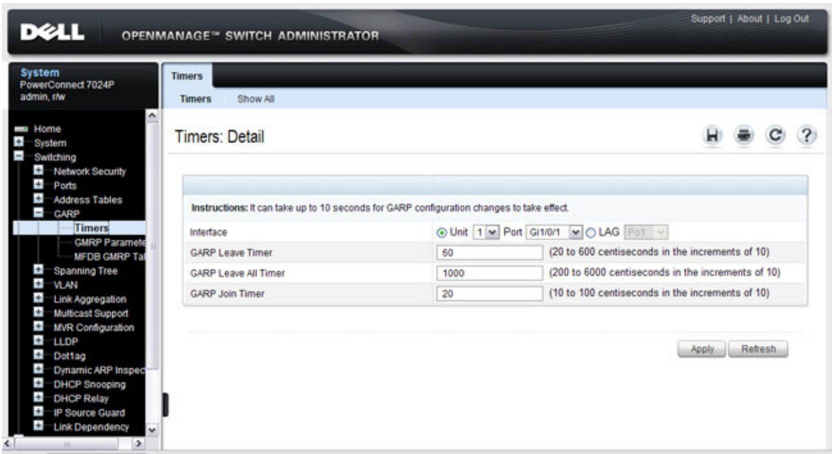


GARP Timers

The **Timers** page contains fields for setting the GARP timers used by GVRP and GMRP on the switch.

To display the **Timers** page, click **Switching** → **GARP** → **Timers** in the navigation panel.

Figure 25-30. GARP Timers



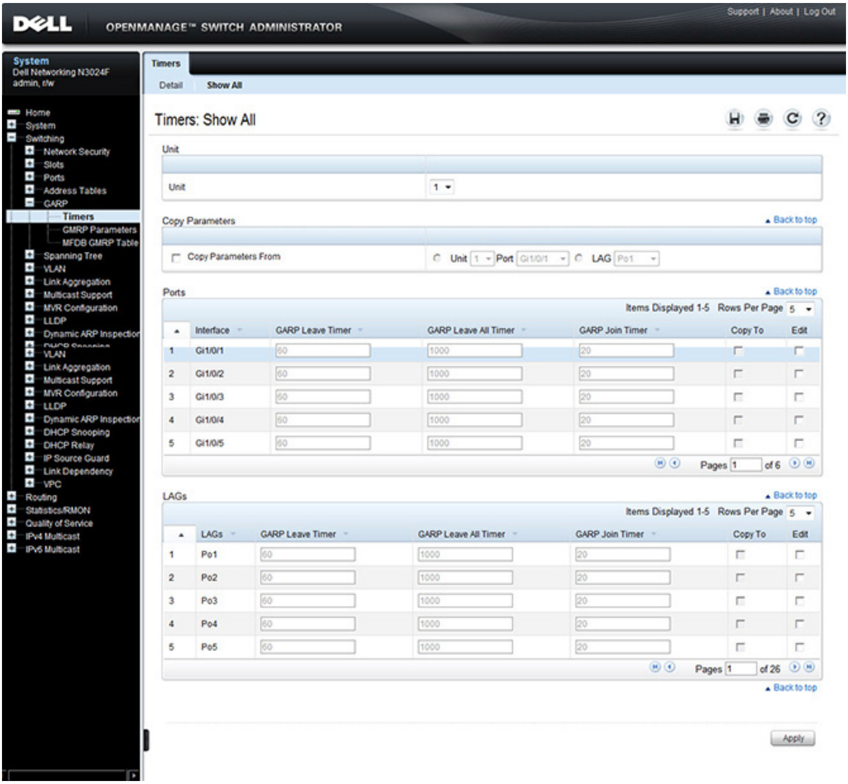
Configuring GARP Timer Settings for Multiple Ports

To configure GARP timers on multiple ports:

1 Open the **Timers** page.

2 Click **Show All** to display the **GARP Timers Table**.

Figure 25-31. Garp Timers Table



3 For each port or LAG to configure, select the check box in the **Edit** column in the row associated with the port.

4 Specify the desired timer values.

5 Click **Apply**.

Copying GARP Timer Settings From One Port to Others

To copy GARP timer settings:

1 Select the **Copy Parameters From** check box, and select the port or LAG with the settings to apply to other ports or LAGs.

2 In the Ports or LAGs list, select the check box(es) in the **Copy To** column that will have the same settings as the port selected in the **Copy Parameters From** field.

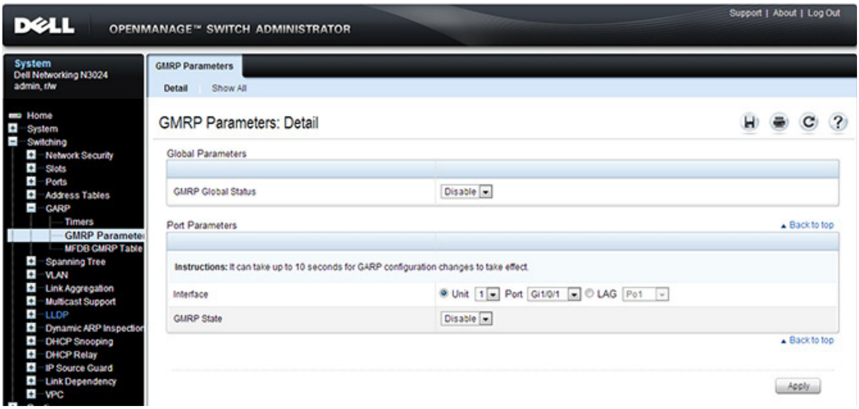
3 Click **Apply** to copy the settings.

GMRP Parameters

Use the **GMRP Parameters** page to configure the administrative mode of GMRP on the switch and on each port or LAG.

To display the **GMRP Parameters** page, click **Switching → GARP → GMRP Parameters** in the navigation panel.

Figure 25-32. GMRP Parameters



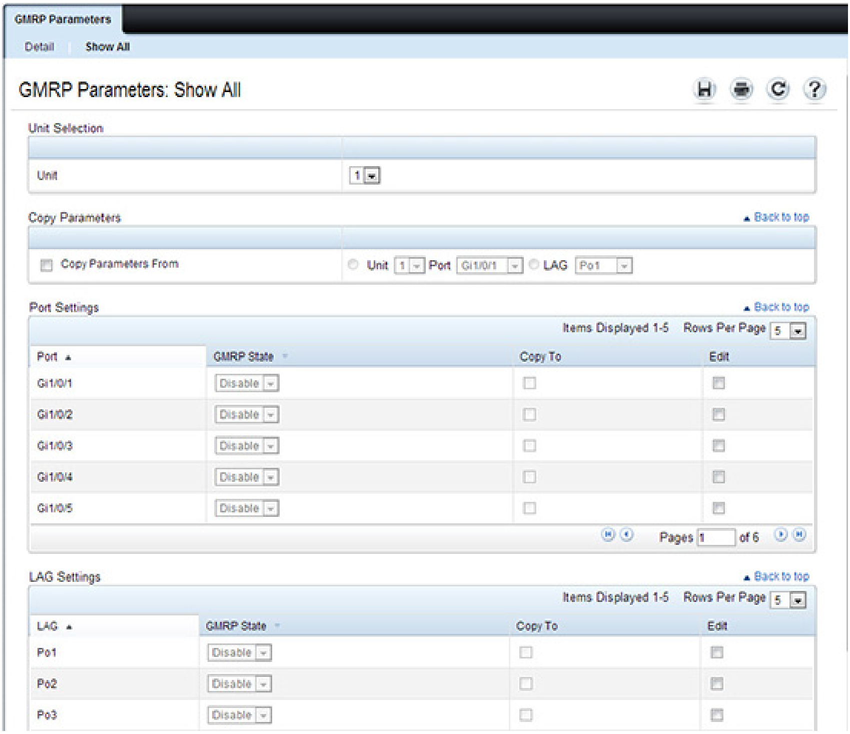
Configuring GMRP Parameters on Multiple Ports

To configure GMRP settings:

1 Open the **GMRP Parameters** page.

2 Click **Show All** to display the **GMRP Port Configuration Table**.

Figure 25-33. GMRP Port Configuration Table



3 For each port or LAG to configure, select the check box in the **Edit** column in the row associated with the port.

4 Specify the desired timer values.

5 Click **Apply**.

Copying Settings From One Port or LAG to Others

To copy GMRP settings:

1 Select the **Copy Parameters From** check box, and select the port or LAG with the settings to apply to other ports or LAGs.

2 In the Ports or LAGs list, select the check box(es) in the **Copy To** column that will have the same settings as the port selected in the **Copy Parameters From** field.

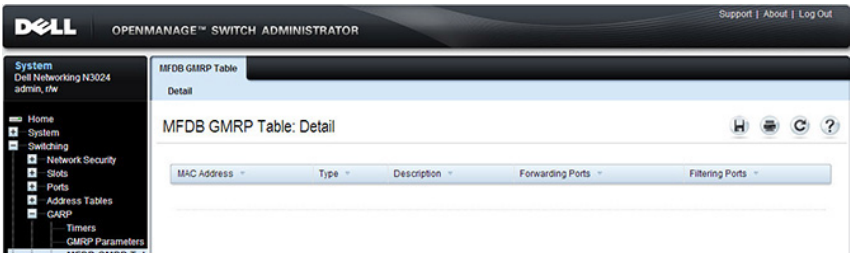
3 Click **Apply** to copy the settings.

MFDB GMRP Table

Use the **MFDB GMRP Table** page to view all of the entries in the Multicast Forwarding Database that were created for the GMRP

To display the **MFDB GMRP Table** page, click **Switching** → **GARP** → **MFDB GMRP Table** in the navigation panel.

Figure 25-34. MFDB GMRP Table



Configuring L2 Multicast Features (CLI)

This section provides information about the commands used for configuring L2 multicast settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Layer-2 Multicasting

Beginning in Privileged EXEC mode, use the following commands to configure MAC address table features.

Command	Purpose
configure	Enter global configuration mode.
mac address-table static *mac-multicast-address* **vlan** *vlan-id* **interface** *interface-id*	Register a MAC-layer Multicast address in the bridge table. • *mac-multicast-address* — MAC multicast address in the format xxxx.xxxx.xxxx or xx:xx:xx:xx:xx:xx. • *interface-id* — A physical interface or port-channel.
mac address-table multicast forbidden address vlan *vlan-id* {*mac-multicast-address* | *ip-multicast-address*} {**add** | **remove**} **interface** *interface-list*	Forbid adding a specific Multicast address to specific ports. • *mac-multicast-address* — MAC multicast address in the format xxxx.xxxx.xxxx. • *ip- multicast-address* — IP multicast address. • **add** — Adds ports to the group. If no option is specified, this is the default option. • **remove** — Removes ports from the group. • *interface-list* — Specifies the interface type (**port-channel**, **gigabitethernet**, **tengigabitethernet**) and number. Separate nonconsecutive interfaces with a comma and no spaces; use a hyphen to designate a range of ports.
exit	Exit to Privileged EXEC mode.

Command	Purpose
show mac address-table multicast [**vlan** *vlan-id*] [**address** *mac-multicast-address* | *ip-multicast-address*] [**format ip** | **mac**]]	View entries in the multicast MAC address table. The **show mac address-table multicast** command shows only multicast addresses. Multicast address are shown along with unicast addresses if the multicast keyword is not used.

Configuring IGMP Snooping on VLANs

Beginning in Privileged EXEC mode, use the following commands to configure IGMP snooping settings on VLANs.

Command	Purpose
configure	Enter global configuration mode.
ip igmp snooping vlan *vlan-id*	Enable IGMP snooping on the specified VLAN.
ip igmp snooping vlan *vlan-id* **groupmembership-interval** *seconds*	Specify the host time-out value for the specified VLAN. If an IGMP report for a multicast group is not received in the number of seconds specified by the *seconds* value, this port is deleted from the VLAN member list of that multicast group. This command also enables IGMP snooping on the VLAN.
ip igmp snooping vlan *vlan-id* **last-member-query-interval** *seconds*	Specify the leave time-out value for the VLAN. If an IGMP report for a multicast group is not received within the number of seconds configured with this command after an IGMP leave was received from a specific interface, the current port is deleted from the VLAN member list of that multicast group.
ip igmp snooping vlan *vlan-id* **immediate-leave**	Enables IGMP snooping immediate-leave mode on the specified VLAN. Enabling immediate-leave allows the switch to immediately remove the layer-2 LAN interface from its forwarding table entry upon receiving an IGMP leave message for that multicast group without first sending out MAC-based general queries to the interface.

Command	Purpose
ip igmp snooping vlan *vlan-id* **mcrtexpiretime** *seconds*	Specify the multicast router time-out value for to associate with a VLAN. This command sets the number of seconds to wait to age out an automatically-learned multicast router port.
CTRL + Z	Exit to Privileged EXEC mode.
show ip igmp snooping groups	Shows IGMP snooping configuration on all VLANs.
show ip igmp snooping vlan *vlan-id*	View the IGMP snooping settings on the VLAN.

Configuring IGMP Snooping Querier

Beginning in Privileged EXEC mode, use the following commands to configure IGMP snooping querier settings on the switch and on VLANs.

Command	Purpose
configure	Enter global configuration mode.
ip igmp snooping querier [**vlan** *vlan-id*] [**address** *ip-address*]	Enable the IGMP snooping querier on the switch or on the VLAN specified with the *vlan-id* parameter. Use the optional *ip-address* parameter to specify the IP address that the snooping querier switch should use as the source address when generating periodic queries.
ip igmp snooping querier query-interval *interval-count*	Set the IGMP snooping querier query interval time, which is the amount of time in seconds that the switch waits before sending another periodic query. The range is 1–1800 seconds.
ip igmp snooping querier timer expiry *seconds*	Set the IGMP snooping querier timer expiration period. This is the time period, in seconds, that the switch remains in non-querier mode after it has discovered that there is a multicast querier in the network.
ip igmp snooping querier version *version*	Set the IGMP version of the query that the switch sends periodically. The *version* range is 1–2.
ip igmp snooping querier *vlan-id*	Enable the IGMP snooping querier on the specified VLAN.

Command	Purpose
ip igmp snooping querier election participate *vlan-id*	Allow the IGMP snooping querier to participate in the querier election process when it discovers the presence of another querier in the VLAN. When this mode is enabled, if the snooping querier finds that the other querier source address is more than the snooping querier address, it stops sending periodic queries. If the snooping querier wins the election, then it continues sending periodic queries and the other querier ceases sending queries. Use of election mode is not recommended when multicast routers are present in the network.
CTRL + Z	Exit to Privileged EXEC mode.
show ip igmp snooping querier [detail |vlan *vlan-id*]	View IGMP snooping querier settings configured on the switch, on all VLANs, or on the specified VLAN.

Configuring MLD Snooping on VLANs

Beginning in Privileged EXEC mode, use the following commands to configure MLD snooping settings on VLANs.

Command	Purpose
configure	Enter global configuration mode.
ipv6 mld snooping vlan *vlan-id*	Enable MLD snooping on the specified VLAN.
ipv6 mld snooping vlan *vlan-id* **groupmembership-interval** *seconds*	Specify the host time-out value for the specified VLAN. If an MLD report for a multicast group is not received in the number of seconds specified by the *seconds* value, this VLAN is deleted from the member list of that multicast group.
ipv6 mld snooping *vlan-id* **last-listener-query-interval** *seconds*	Specify the leave time-out value for the VLAN. If an MLD report for a multicast group is not received within the number of seconds configured with this command after an MLD leave was received from a specific interface, the current port is deleted from the VLAN member list of that multicast group.

Command	Purpose
ipv6 mld snooping vlan *vlan-id* **immediate-leave**	Enables MLD snooping immediate-leave mode on the specified VLAN. Enabling immediate-leave allows the switch to immediately remove the layer-2 LAN interface from its forwarding table entry upon receiving an MLD leave message for that multicast group without first sending out MAC-based general queries to the interface.
ipv6 mld snooping vlan *vlan-id* **mcrtexpiretime** *seconds*	Specify the multicast router time-out value for to associate with a VLAN. This command sets the number of seconds to wait to age out an automatically-learned multicast router port.
CTRL + Z	Exit to Privileged EXEC mode.
show ipv6 mld snooping vlan *vlan-id*	View the MLD snooping settings on the VLAN.

Configuring MLD Snooping Querier

Beginning in Privileged EXEC mode, use the following commands to configure MLD snooping querier settings on the switch and on VLANs.

Command	Purpose
configure	Enter global configuration mode.
ipv6 mld snooping querier	Enable the MLD snooping querier on the switch.
ipv6 mld snooping querier vlan *vlan-id* [**address** *ipv6-address*]	Enable the MLD snooping querier on VLAN specified with the *vlan-id* parameter. Use the optional *ip-address* parameter to specify the IP address that the snooping querier switch should use as the source address when generating periodic queries.

Command	Purpose
ipv6 mld snooping querier election participate *vlan-id*	Allow the MLD snooping querier to participate in the querier election process when it discovers the presence of another querier in the VLAN. When this mode is enabled, if the snooping querier finds that the other querier source address is more than the snooping querier address, it stops sending periodic queries. If the snooping querier wins the election, then it continues sending periodic queries. Use of election mode is not recommended when multicast routers are present in the network.
exit	Exit to Global Configuration mode.
ipv6 mld snooping querier address *ipv6-address*	Specify the IP address that the snooping querier switch should use as the source address when generating periodic queries.
ipv6 mld snooping querier query-interval *interval-count*	Set the MLD snooping querier query interval time, which is the amount of time in seconds that the switch waits before sending another periodic query. The range is 1–1800 seconds.
ipv6 mld snooping querier timer expiry *seconds*	Set the MLD snooping querier timer expiration period. This is the time period, in seconds, that the switch remains in non-querier mode after it has discovered that there is a multicast querier in the network.
exit	Exit to Privileged EXEC mode.
show ipv6 mld snooping querier [detail |vlan *vlan-id*]	View MLD snooping querier settings configured on the switch, on all VLANs, or on the specified VLAN.

Configuring MVR

Beginning in Privileged EXEC mode, use the following commands to configure MVR features on the switch.

Command	Purpose
configure	Enter global configuration mode.
mvr	Enable MVR on the switch.
mvr vlan *vlan-id*	Set the VLAN to use as the multicast VLAN for MVR.

Command	Purpose
mvr querytime *time*	Set the MVR query response time. The value for *time* is in units of tenths of a second. This is the time to wait for a response to the query sent after receiving a leave message and before removing the port from the group.
mvr mode {**compatible** | **dynamic**}	Specify the MVR mode of operation.
mvr group *mcast-address* [*groups*]	Add an MVR membership group. • *mcast-address*—The group IP multicast address • *group*—Specifies the number of contiguous groups
interface *interface*	Enter interface configuration mode for the specified port. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of ports can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
mvr	Enable MVR on the port.
mvr immediate	Enable MVR immediate leave mode on the port.
mvr type {**source** | **receiver**}	Specify the MVR port type.
mvr vlan *vlan-id* **group** *mcast-address*	Allow the port to participate in the specified MVR group. The *vlan-id* parameter is the ID of the MVR multicast VLAN.
CTRL + Z	Exit to Privileged EXEC mode.
show ip dhcp snooping [**interfaces**]	View the DHCP snooping global and per port configuration.
show ip dhcp snooping binding [{**static** | **dynamic**}] [**interface** *port*] [**vlan** *vlan-id*]	View the entries in the DHCP snooping bindings database.
show mvr	View information about the administrative mode of MVR.
show mvr members	View information about MVR groups and their members.

Command	Purpose
show mvr interface *interface*	View information about the MVR configuration for a specific port.
show mvr traffic	View information about IGMP traffic in the MVR table.

Configuring GARP Timers and GMRP

Beginning in Privileged EXEC mode, use the following commands to configure the GARP timers and to control the administrative mode GMRP on the switch and per-interface.

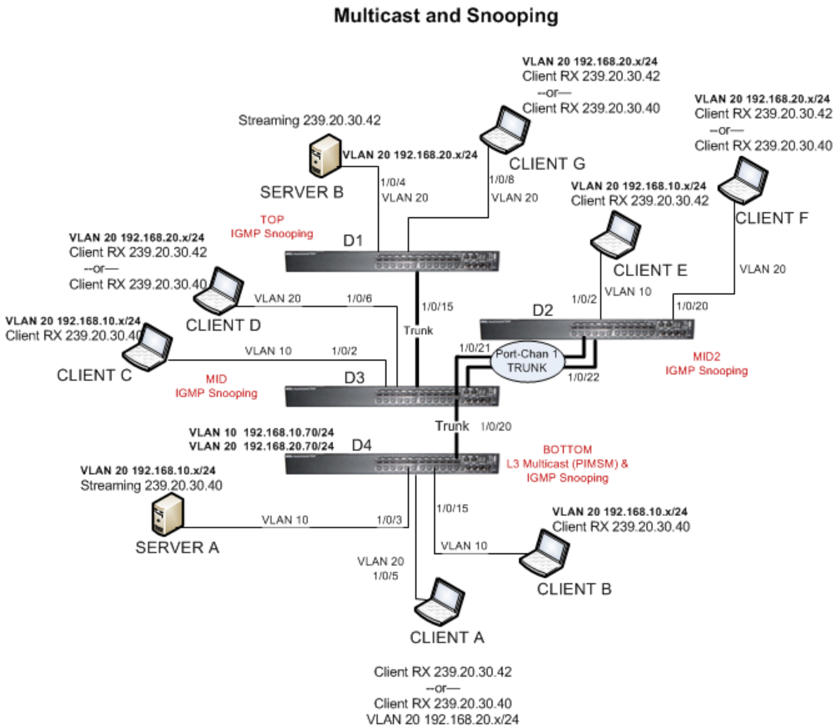
Command	Purpose
configure	Enter global configuration mode.
garp timer {join | leave | leaveall} *timer_value*	Adjust the GARP application join, leave, and leaveall GARP timer values The *timer_value* variable is in centiseconds. The range is 10-100 for **join**, 20-600 for **leave**, and 200-6000 for **leaveall**.
gmrp enable	Enable GMRP globally on the switch.
interface *interface*	Enter interface configuration mode for the specified port or LAG. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. For a LAG, the interface type is **port-channel**. A range of ports can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
gmrp enable	Enable GMRP on the interface or range of interfaces.
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show gmrp configuration	View the administrative status of GMRP on the switch and all interfaces.

Case Study on a Real-World Network Topology

Multicast Snooping Case Study

Figure 25-35 shows the topology that the scenarios in this case study use.

Figure 25-35. Case Study Topology



The topology in Figure 25-35 includes the following elements:

- Snooping Switches: D1, D2, D3 with IGMP snooping enabled on VLANs 10, 20
- Multicast Router: D4 with PIM-SM enabled on VLANs 10, 20
- Multicast Listeners: Client A-G
- Multicast Sources: Server A – 239.20.30.40, Server B – 239.20.30.42

- Subnets: VLAN 10 – 192.168.10.x, VLAN 20 – 192.168.20.x
- Mrouter ports: D3 – 1/0/20, D2 – PortChannel1, D1 – 1/0/15

Snooping Within a Subnet

In the example network topology, the multicast source and listeners are in the same subnet VLAN 20 – 192.168.20.x/24. D4 sends periodic queries on VLAN 10 and 20, and these queries are forwarded to D1, D2, and D3 via trunk links. Snooping switches D1, D2, and D3 flood these queries in VLANs 10 and 20 to clients G, F, and D, respectively.

Multicast Source and Listener directly connected to a snooping switch: Server B → Client G

1. Client G sends a report for 239.20.30.42.
2. The report is forwarded to multicast router D4 via D1 – 1/0/15 and D3 – 1/0/20.
3. A forwarding entry is created by D1 for VLAN 20, 239.20.30.42 – 1/0/8, 1/0/15.
4. Client G receives the multicast stream from Server B.
5. D3 receives the multicast stream and it is forwarded to D4 because D4 is a multicast router.
6. Client D sends a report for 239.20.30.42.
7. The report is forwarded to multicast router D4 via D3 – 1/0/20.
8. A forwarding entry is created by D3 for VLAN 20, 239.20.30.42 – 1/0/6, 1/0/20.
9. Client D receives the multicast stream from Server B.
10. Client F does not receive the multicast stream because it did not respond to queries from D4.

Multicast Source and Listener connected by intermediate snooping switches: Server B → Client D

1. Client D sends a report for 239.20.30.42.
2. The report is forwarded to multicast router D4 via D3 – 1/0/20.
3. A forwarding entry is created by D3 for VLAN20, 239.20.30.42 – 1/0/6, 1/0/20.

4 Client D will receive the multicast stream from Server B because it is forwarded by D1 to D3 and then to D4 because D4 is a multicast router. Because the multicast stream is present on D3, a L2 forwarding entry is created on D3, where 239.20.30.42 is not a registered group.

5 Client F does not receive the multicast stream because it did not respond to queries from D4.

Snooping Switch Interaction with a Multicast Router

In the example network topology, consider Client B and Server A. Both are in the same subnet VLAN10 – 192.168.10.70/24. Server A is a source for multicast stream 239.20.30.40. D4 sends periodic queries on VLAN 10 and VLAN 20, and these queries reach D1, D2, and D3 via trunk links, which in turn forward them in VLAN 10 and VLAN 20 to reach their respective attached clients. PIM-SM is enabled on router D4, and IGMP snooping is enabled on D1, D2, and D3.

Multicast Source and Listener directly connected to Multicast Router on the same routing VLAN: Server A → Client B

1 Because multicast routing is enabled on D4 VLAN 10, an IP multicast table entry is created to include D4 – 1/0/15, D4 – 1/0/20 as part of the L2 forwarding list members.

2 Client B sends a report for 239.20.30.40.

3 The IP multicast table entry is modified to include only D4 – 1/0/15 as the L2 forwarding list member. IGMP snooping creates an L2 forwarding entry for Client B.

4 Client B receives multicast data.

5 The multicast stream is not forwarded to D3 on trunk link 1/0/20 because no other clients requested this data.

Multicast Source directly connected to Multicast Router, and Listener connected to a different routing VLAN via intermediate snooping switches: Server A → Client F

Clients A, D and F are in the same subnet VLAN20 - 192.168.20.70/24. Server A is in a different subnet VLAN10 – 192.168.10.70/24.

1 Client F sends a report for 239.20.30.40.

2. A multicast forwarding entry is created on D2 VLAN20, 239.20.30.40 – 1/0/20, PortChannel1.
3. The Client F report message is forwarded to D3-PortChannel1 (multicast router attached port).
4. A multicast forwarding entry is created on D3 VLAN 20, 239.20.30.40 – PortChannel1, 1/0/20.
5. The Client F report message is forwarded to D4 via D3 – 1/0/20 (multicast router attached port).
6. An IP multicast routing entry is created on D4 VLAN 10 – VLAN 20 with the layer-3 outgoing port list as VLAN 20 – 1/0/20.
7. The multicast stream is routed to D3.
8. The multicast stream is forwarded to listener Client F using forwarding entries created on D3 and D2.
9. Clients A and D do not receive the Server A multicast stream because they did not send a report.

Multicast Source connected to Multicast Router via intermediate snooping switches, and Listener directly connected to multicast router in a different routing interface: Server B → Client B

Server A and Clients B, C, and E are on the same subnet VLAN10 – 192.168.10.70/24. Server B is in a different subnet VLAN20 – 192.168.20.70/24.

1. Client B sends a report for 239.20.30.42.
2. Multicast Router D4 learns group 239.20.30.42.
3. The administrator creates a static multicast forwarding entry on D1 VLAN 20, 239.20.30.42 – 1/0/15 and on D3 VLAN 20, 239.20.30.42 – 1/0/20.
4. The multicast stream from Server B reaches D4 via trunk links because it is a statically registered group on D1 and D3. D4 is a multicast router.
5. An IP multicast routing entry is created on D4 VLAN 20 – VLAN 10 with the layer-3 outgoing port list as VLAN 10 – 1/0/15.
6. Client B receives multicast data from Server B.
7. Server A and Clients C and E do not receive Server B data because no report messages were sent requesting Server B traffic.

Multicast Source and Listener connected to Multicast Router via intermediate snooping switches and are part of different routing VLANs: Server B → Client E

Clients E, B, and C are on the same subnet VLAN10 – 192.168.10.70/24. Server B is in a different subnet VLAN20 – 192.168.20.70/24.

1. Client E sends a report for 239.20.30.42.
2. A multicast forwarding entry is created on D2 VLAN10, 239.20.30.42 – 1/0/2, PortChannel 1.
3. The report from Client E is forwarded to D3 via D2 – PortChannel 1.
4. A multicast forwarding entry is created on D3 VLAN10, 239.20.30.42 – PortChannel 1, 1/0/20.
5. The report from Client E is forwarded to D4 via D3 – 1/0/20.
6. Multicast Router D4 learns group 239.20.30.42.
7. The multicast stream from Server B reaches D4 via trunk links because it is a multicast router.
8. An IP multicast routing entry is created on D4 VLAN 20 – VLAN 10 with the layer-3 outgoing port list as VLAN 10 – 1/0/20.
9. Client E receives multicast data from Server B.
10. Clients B and C do not receive Server B data because no report messages were sent requesting Server B traffic.

26

Connectivity Fault Management

Dell Networking N4000 Series Switches

NOTE: This feature is supported only on the Dell Networking N4000 Series switches.

This chapter describes how to configure the Connectivity Fault Management feature, which is specified in IEEE 802.1ag (*IEEE Standard for Local and Metropolitan Area Networks Virtual Bridged Local Area Networks Amendment 5: Connectivity Fault Management*). This protocol, also known as Dot1ag, enables the detection and isolation of connectivity faults at the service level for traffic that is bridged over a metropolitan Ethernet LAN.

The topics covered in this chapter include:

- Dot1ag Overview
- Default Dot1ag Values
- Configuring Dot1ag (Web)
- Configuring Dot1ag (CLI)
- Dot1ag Configuration Example

Dot1ag Overview

With the emergence of Ethernet as a Metropolitan and Wide-Area Networking technology, different operators often work together to provide end-to-end services to enterprise customers. This has driven the need of a new set of OAM (Operations, Administration, and Maintenance) Protocols.

Service-Level Connectivity Fault Management (CFM) is the OAM protocol provision for end-to-end service-layer instances in carrier networks. CFM provides mechanisms to support the administrator in performing connectivity checks, fault detection, fault verification and isolation, and fault notification per service in the network domain of interest. Unlike Ethernet OAM (IEEE 802.3ah), where the faults are detected and notified on a single point-to-point

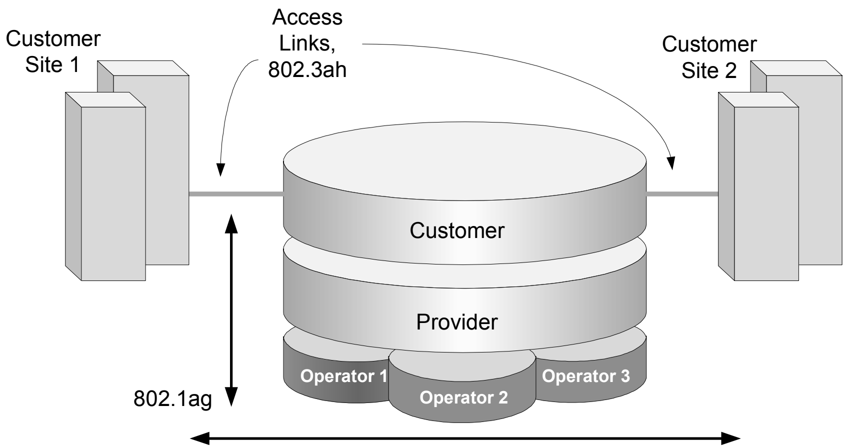
IEEE Std. 802.3 LAN, Dot1ag addresses fault diagnosis at the service layer across networks comprising multiple LANs, including LANs other than 802.3 media.

How Does Dot1ag Work Across a Carrier Network?

A typical metropolitan area network comprises operator, service provider, and customer networks. To suit this business model, CFM relies on a functional model of hierarchical maintenance domains (MDs). These domains are assigned a unique MD level. There is a maximum of 8 levels, which can be nested but cannot overlap. Each organization can have its own maintenance domain. The MD level limits administrator access to the appropriate domain.

Figure 26-1 depicts three domains: the customer subscribes to the services of a provider, who, in turn, subscribes to the services of two operators. This scenario is a likely one, since no operator has complete coverage of a large region. A service instance would span the provider network covering one or more operators. Every domain has its own network management system. Dot1ag defines OAM services that operate across these domains (the vertical arrow) and within them (the horizontal arrow)

Figure 26-1. Organization of Domains



Entities at different levels have different responsibilities. For example, the lower level (operator) overlooks a subset of the network in detail and provides information about its status to its higher levels such as the provider level).

Higher levels have a broader, but less detailed, view of the network. As a result, a provider could include multiple operators, provided that the domains never intersect. The operator transparently passes frames from the customer and provider, and the customer does not see the operator frames. Multiple levels within a domain (say, operator) are supported for flexibility.

What Entities Make Up a Maintenance Domain?

Dot1ag defines three primary entities that make up the maintenance domain: Maintenance End Points (MEPs), Maintenance Intermediate Points (MIPs), and Maintenance Associations (MAs).

MEPs, and MIPs

MEPs and MIPs are software or sometimes hardware per-service entities where CFM functionalities are present.

- MEPs define the boundary of a maintenance domain. They initiate and respond to CFM messages. MEPs prevent the leaking of CFM messages between domains (for example, among operators or between operators and customers). Each MEP has a configurable unique identifier (MEPID) in a maintenance domain.

  MEPs periodically issue Continuity Check Messages (CCM) to discover each other and issue SNMP traps to report connectivity losses or malformed or incorrect CCMs.

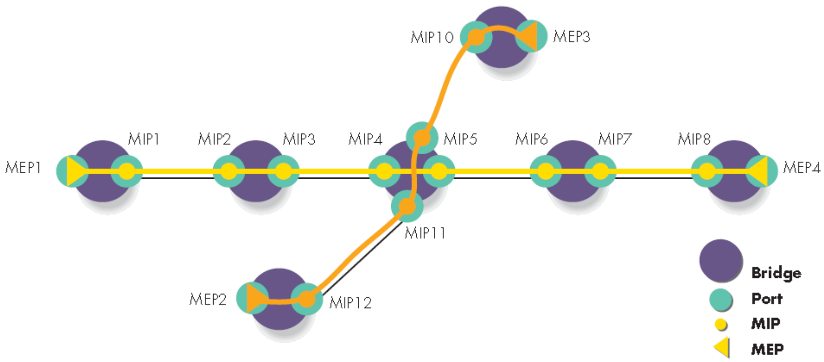
  A MEP can be defined as “down MEP” or an “up MEP”. A down MEPs reside in a bridge that transmits CFM PDUs towards, and receives them from, the direction of the LAN. An up MEP resides in a bridge that transmits CFM PDUs towards, and receives them from, the direction of the Bridge Relay Entity.

  NOTE: An entity at the boundary of maintenance domain that offers connectivity and other services to systems outside the domain is referred to as a Domain Service Access Point (DoSAP). A MEP is a type of DoSAP whose services relate to connectivity fault management.

- MIPs are entities within a domain that enable the outer domain to achieve end-to-end connectivity checks. MIPs passively receive CFM messages and respond back to the originating MEP.

Figure 26-2 depicts two MEPs and the MIPs that connect them in a maintenance domain.

Figure 26-2. Maintenance Endpoints and Intermediate Points



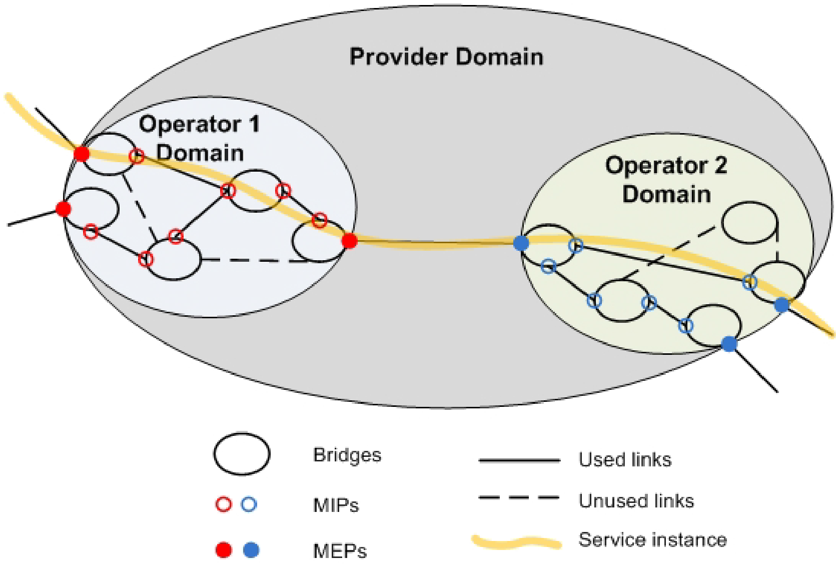
Maintenance Associations

An MA is a logical connection between one or more MEPs that enables monitoring a particular service instance. Each MA is associated with a unique SVLAN ID. An MA is identified by a maintenance association ID. All MEPs in the MA are assigned the maintenance identifier (MAID) for the association.

An MD consists of one or more MAs at the same domain level.

Figure 26-3 depicts one provider-level domain and two operator-level domains. Dot1ag operation for a service instance is indicated by the path that traverses the different domains to provide the end-to-end connectivity fault management for the service.

Figure 26-3. Provider View for Service Level OAM



What is the Administrator's Role?

On the switch, the administrator configures the customer-level maintenance domains, associations, and endpoints used to participate in Dot1ag services with other switches connected through the provider network. The Administrator can also use utilities to troubleshoot connectivity faults when reported via SNMP traps. All the domains within the customer domain should use different domain levels.

Configuration Tasks

The administrator defines the maintenance domains by configuring the domain level (from 0–7) and a name. For each domain, the administrator defines maintenance associations that are specified by a SVLAN ID and an MA name. Then the administrator defines the switch ports that serve as MEPs for a service instance and as MIPs within a domain.

Troubleshooting Tasks

In the event of a connectivity loss between MEPs, the administrator can perform path discovery, similar to traceroute, from one MEP to any MEP or MIP in a maintenance domain using Link Trace Messages (LTMs). The connectivity loss is narrowed down using path discovery and is verified using Loop-back Messages (LBMs), which are similar to ping operations in IP networks.

Default Dot1ag Values

Dot1ag service are disabled by default and no maintenance domains, associations, or endpoints are configured by default.

Table 26-1 shows the global default values for Dot1ag.

Table 26-1. Dot1ag Global Defaults

Parameter	Default Value
CFM Admin Mode	Disabled
Archive Hold Time	600 seconds

When you configure an association between a VLAN and a maintenance domain, the following default value applies:

Table 26-2. MA Configuration Defaults

Parameter	Default Value
Continuity Check Message (CCM) Interval	1 second

When you associate endpoints with SVLAN IDs, the following default values apply and are configurable:

Table 26-3. MEP Configuration Defaults

Parameter	Default Value
MEP Active	False
Continuity Check Interval (CCI) Enabled	True

Configuring Dot1ag (Web)

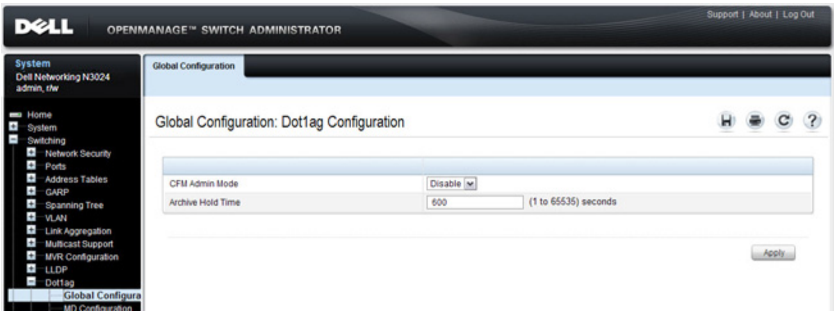
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring Dot1ag features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

Dot1ag Global Configuration

Use the **Global Configuration** page to enable and disable the Dot1ag admin mode and to configure the time after which inactive RMEP messages are removed from the MEP database.

To display the page, click **Switching** → **Dot1ag** → **Global Configuration** in the tree view.

Figure 26-4. Dot1ag Global Configuration

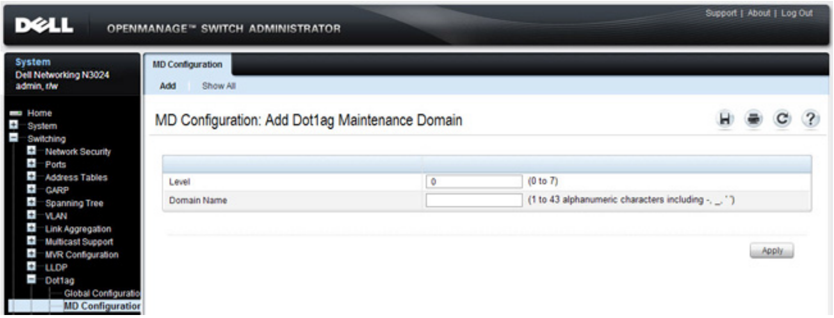


Dot1ag MD Configuration

Use the **MD Configuration** page to configure maintenance domain levels and names.

To display the page, click **Switching** → **Dot1ag** → **MD Configuration** in the tree view.

Figure 26-5. Dot1ag MD Configuration

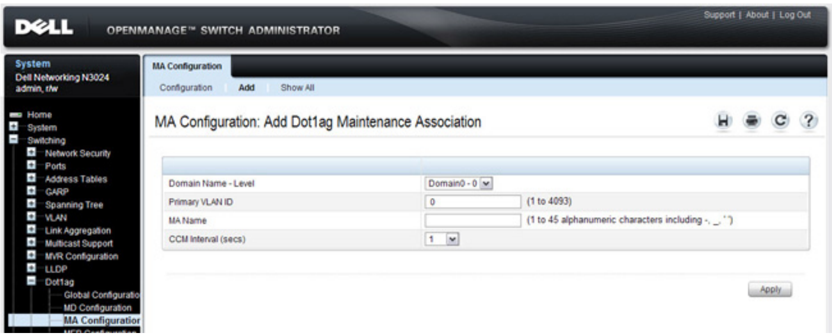


Dot1ag MA Configuration

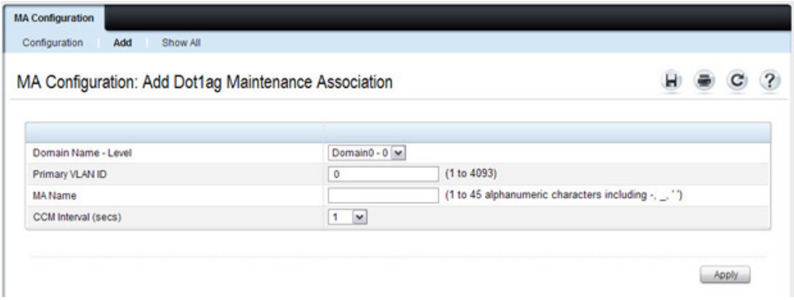
Use the **MA Configuration** page to associate a maintenance domain level with one or more VLAN ID, provide a name for each maintenance association (MA), and to set the interval between continuity check messages sent by MEPs for the MA.

To display the page, click **Switching** → **Dot1ag** → **MA Configuration** in the tree view.

Figure 26-6. Dot1ag MA Configuration



To add an MA, click the **Add** link at the top of the page.

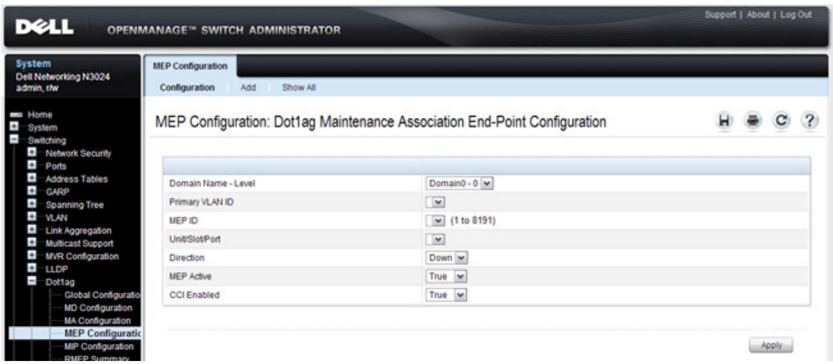


Dot1ag MEP Configuration

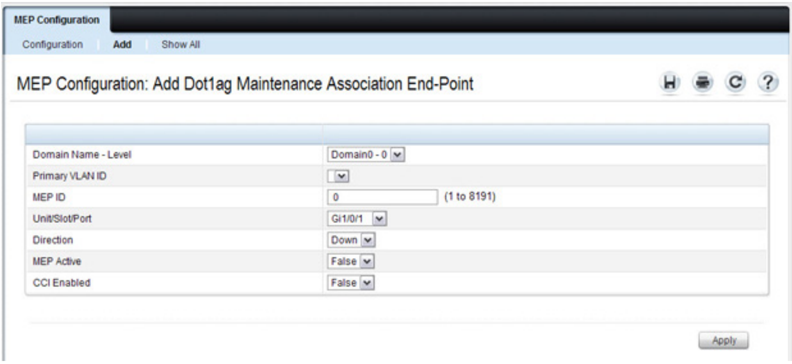
Use the **MEP Configuration** page to define switch ports as Management End Points. MEPs are configured per domain and per VLAN.

To display the page, click **Switching** → **Dot1ag** → **MEP Configuration** in the tree view.

Figure 26-7. Dot1ag MEP Configuration



To add a MEP, click the **Add** link at the top of the page. A VLAN must be associated with the selected domain before you configure a MEP to be used within an MA (see the **MA Configuration** page).

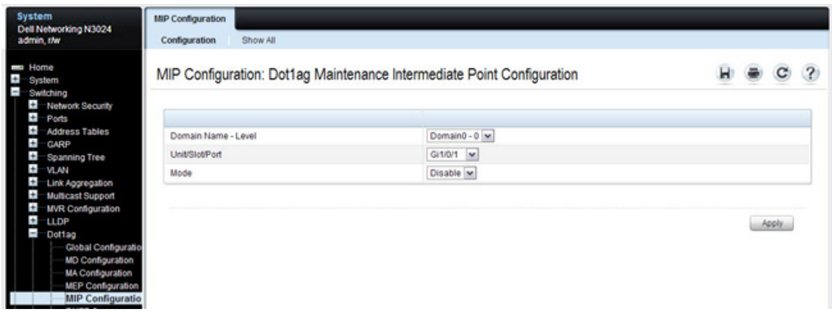


Dot1ag MIP Configuration

Use the **MIP Configuration** page to define a switch port as an intermediate bridge for a selected domain.

To display the page, click **Switching** → **Dot1ag** → **MIP Configuration** in the tree view.

Figure 26-8. Dot1ag MIP Configuration

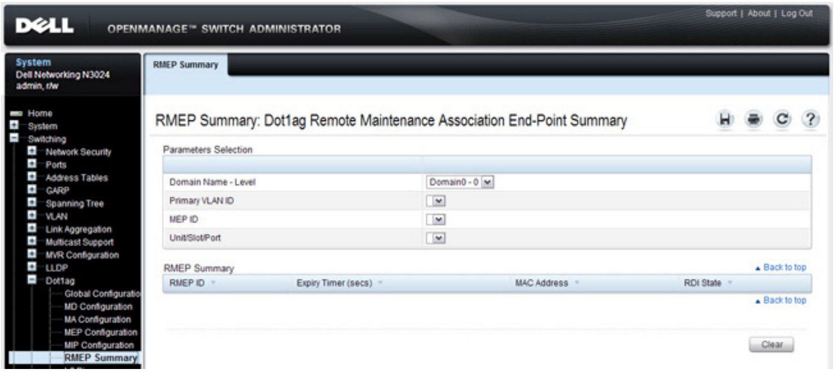


Dot1ag RMEP Summary

Use the **RMEP Summary** page to view information on remote MEPs that the switch has learned through CFM PDU exchanges with MEPs on the switch.

To display the page, click **Switching → Dot1ag → RMEP Summary** in the tree view.

Figure 26-9. Dot1ag RMEP Summary

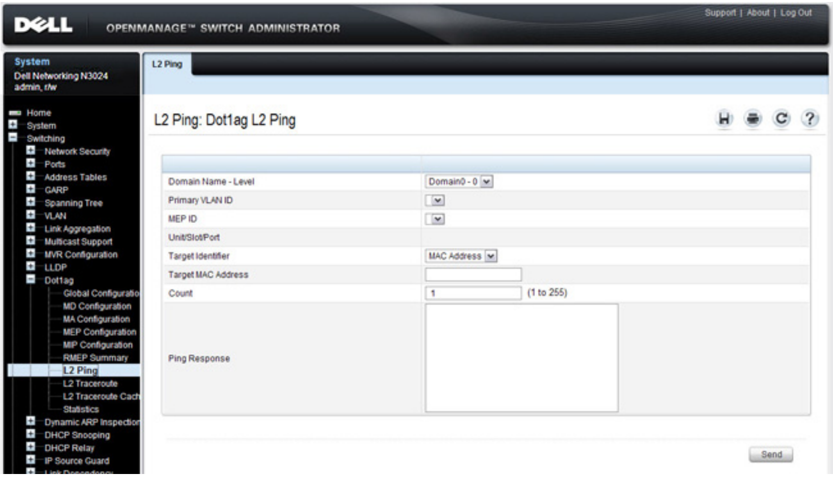


Dot1ag L2 Ping

Use the **L2 Ping** page to generate a loopback message from a specified MEP. The MEP can be identified by the MEP ID or by its MAC address.

To display the page, click **Switching** → **Dot1ag** → **L2 Ping** in the tree view.

Figure 26-10. Dot1ag L2 Ping

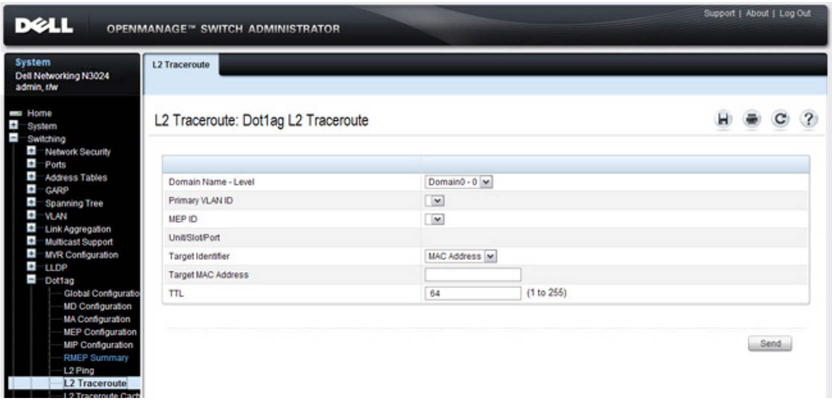


Dot1ag L2 Traceroute

Use the **L2 Traceroute** page to generate a Link Trace message from a specified MEP. The MEP can be specified by the MAC address, or by the remote MEP ID.

To display the page, click **Switching** → **Dot1ag** → **L2 Traceroute** in the tree view.

Figure 26-11. Dot1ag L2 Traceroute

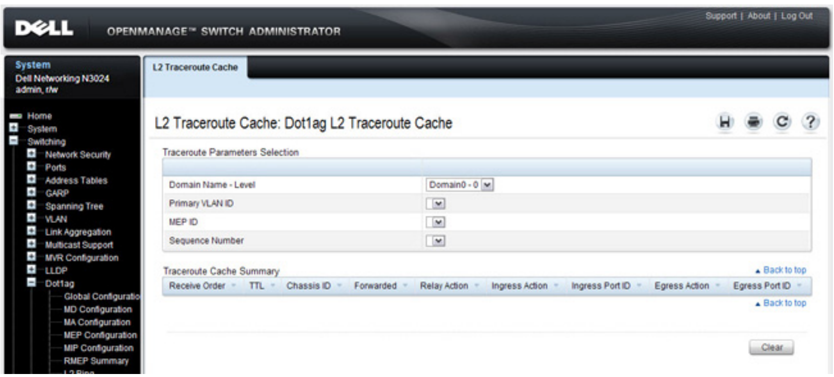


Dot1ag L2 Traceroute Cache

Use the **L2 Traceroute Cache** page to view link traces retained in the link trace database.

To display the page, click **Switching** → **Dot1ag** → **L2 Traceroute Cache** in the tree view.

Figure 26-12. Dot1ag L2 Traceroute Cache

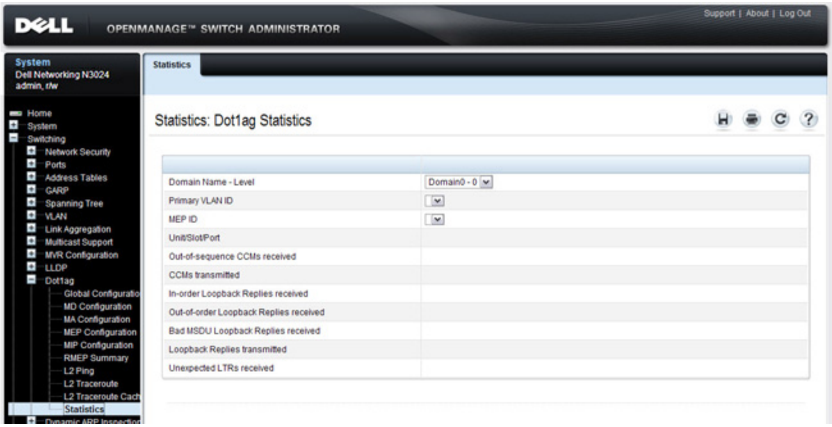


Dot1ag Statistics

Use the **Statistics** page to view Dot1ag information for a selected domain and VLAN ID.

To display the page, click **Switching** → **Dot1ag** → **Statistics** in the tree view.

Figure 26-13. Dot1ag Statistics



Configuring Dot1ag (CLI)

This section provides information about the commands used for configuring Dot1ag settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Dot1ag Global Settings and Creating Domains

Beginning in Privileged Exec mode, use the following commands to configure CFM settings and to view global status and domain information.

CLI Command	Description
configure	Enter global configuration mode.
ethernet cfm enable	Enables connectivity fault management services.
ethernet cfm mep archive-hold-time *time*	Set the time interval (range: 1–65535 seconds) after which inactive RMEPs are removed.
ethernet cfm cc level *level* **vlan** *vlan-id* **interval** {1 | 10 | 60 | 600}	Configure the Continuity Check Message (CCM) transmit interval for the specified VLAN.
ethernet cfm domain *name* **level** *level*	Create a maintenance domain (MD) by assigning a name and level (0–7), and enter Maintenance Domain Config mode for that MD.
service *name* **vlan** *vlan-id*	Create a maintenance association (MA) within the current MD by associating it with a VLAN and naming the association (as a service instance).
exit	Exit to privileged Exec Mode
show ethernet cfm domain brief	Display the configured parameters in the Maintenance Domain.

Configuring MEP Information

Beginning in Privileged Exec mode, use the following commands to configure the mode and view related settings.

CLI Command	Description
configure	Enter global configuration mode.
interface *interface*	Enter Interface Config mode for the specified interface, where *interface* is replaced by **gigabitethernet** *unit/slot/port*, or **tengigabitethernet** *unit/slot/port*.
ethernet cfm mep enable level *level* **vlan** *vlan-id* **mpid** *mep-id*	Define the port as a maintenance endpoint (MEP) and associate it with an SVLAN in a domain. When the MEP is enabled, it will generate CCM messages.
ethernet cfm mep level *level* **direction** {**up** | **down**} **mpid** *mep-id* **vlan** *vlan-id*	Enable a MEP at the specified level and direction.
ethernet cfm mep active	Set the administrative state of the MEP to active.
ethernet cfm mip level *level*	Create a MIP at the specified level on the interface.
exit	Exit to privileged Exec Mode
show ethernet cfm maintenance–points	Add the keywords **local domain**, **local interface**, **local level**, **remote domain**, or **remote level** to show information on local or remote endpoints.
show ethernet cfm statistics	Display statistics per MEP.

Dot1ag Ping and Traceroute

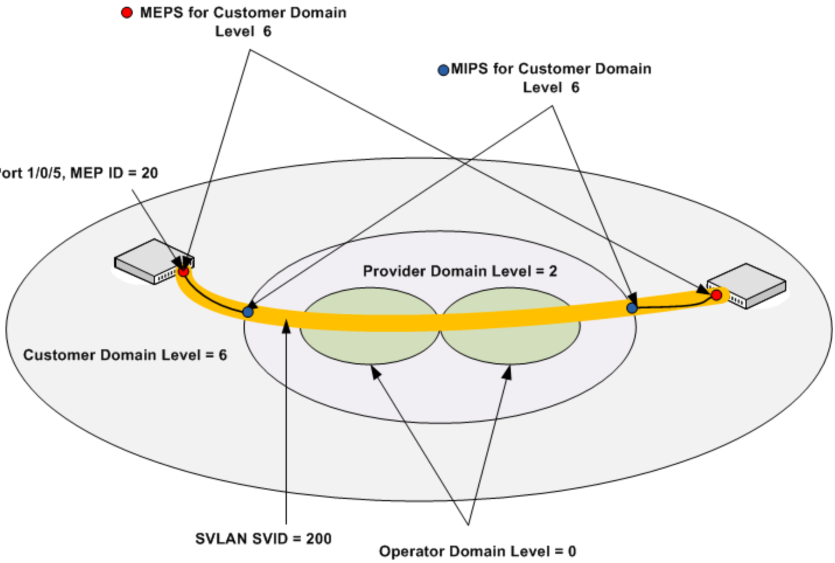
Beginning in Privileged Exec mode, use the following commands to help identify and troubleshoot Ethernet CFM settings.

CLI Command	Description
ping ethernet cfm mac *mac-addr*	Generate a loopback message from the MEP with the specified MAC address.
ping ethernet cfm remote–mpid *mep-id*	Generate a loopback message from the MEP with the specified MEP ID.
traceroute ethernet cfm mac *mac-addr*	Generate a Link Trace message from the MEP with the specified MAC address.
traceroute ethernet cfm remote–mpid *mep-id*	Generate a Link Trace message from the MEP with the specified MEP ID.
show ethernet cfm traceroute–cache	Show the link trace database.

Dot1ag Configuration Example

In the following example, the switch at the customer site is part of a Metro Ethernet network that is bridged to remote sites through a provider network. A service VLAN (SVID 200) identifies a particular set of customer traffic on the provider network.

Figure 26-14. Dot1ag Configuration for a Metro Ethernet Customer Network



To configure the switch:

1 Enable CFM globally on the switch, and then create a level-6 management domain named CustDom for end-to-end CFM on the Metro Ethernet network. VLAN 200 is associated with this domain.

```
console#config
console(config)#ethernet cfm enable
console(config)#ethernet cfm domain CustDom level 6
console(config-cfm-mdomain)#service vlan vlan 200
console(config-cfm-mdomain)#exit
```

2 Configure port 1/0/5 as an MEP for service VLAN 200 so that the port can exchange CFM PDUs with its counterpart MEPs on the customer network. The port is first configured as a MEP with MEP ID 20 on domain level 6 for VLAN 200. Then the port is enabled and activated as a MEP.

```
console(config)#interface gigabitethernet 1/0/5
console(config-if-Gi1/0/5)#ethernet cfm mep level 6 direction
down mpid 20 vlan 200
console(config-if-Gi1/0/5)#ethernet cfm mep enabled level 6
vlan 200 mpid 20
console(config-if-Gi1/0/5)#ethernet cfm mep active level 6
vlan 200 mpid 20
console(config-if-Gi1/0/5)#exit
```

3 On an intermediate switch, configure the MIP for the customer domain and enable CFM services on the CustDom domain to include local network devices.

27

Snooping and Inspecting Traffic

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes Dynamic Host Configuration Protocol (DHCP) Snooping, IP Source Guard (IPSG), and Dynamic ARP Inspection (DAI), which are layer-2 security features that examine traffic to help prevent accidental and malicious attacks on the switch or network.

The topics covered in this chapter include:

Traffic Snooping and Inspection Overview

DHCP Snooping is a security feature that monitors DHCP messages between a DHCP client and DHCP server to filter harmful DHCP messages and to build a bindings database. The IPSG and DAI features use the DHCP Snooping bindings database to help enforce switch and network security.

IP Source Guard allows the switch to drop incoming packets that do not match a binding in the bindings database. Dynamic ARP Inspection allows the switch to drop ARP packets whose sender MAC address and sender IP address do not match an entry in the DHCP snooping bindings database.

DHCP Snooping and IPSG are supported for both IPv4 and IPv6. DAI is supported for IPv4 only, as IPv6 does not use ARP.

What Is DHCP Snooping?

Dynamic Host Configuration Protocol (DHCP) Snooping is a security feature that monitors DHCP messages between a DHCP client and DHCP server to accomplish the following tasks:

- Ensure that only authorized DHCP clients are able to utilize the network.
- Designate which ports are connected to trusted DHCP servers and drop DHCP messages from servers connected to untrusted ports
- Build an authorized DHCP client bindings database with entries that consist of the following information:
  - MAC address
  - IP address
  - VLAN ID
  - Client port

Entries in the bindings database are considered to be authorized network clients. DHCP clients can exchange messages with DHCP servers connected via trusted ports. DHCP client messages are never forwarded to untrusted ports.

DHCP snooping can be enabled on VLANs, and the trust status (trusted or untrusted) is specified on individual physical ports or LAGS that are members of the VLAN. When a port or LAG is configured as untrusted, it could potentially be used to launch a network attack. DHCP snooping protects against attacks on untrusted ports. DHCP servers must be reached through trusted ports. DHCP clients are configured on untrusted ports.

DHCP snooping enforces the following security rules:

- DHCP packets from a DHCP server (DHCPOFFER, DHCPACK, DHCPNAK, DHCPRELEASEQUERY) are dropped if they are received on an untrusted port and a warning level message is logged if invalid DHCP packet logging is enabled. DHCP client originated messages are never forwarded over untrusted ports.
- DHCPRELEASE and DHCPDECLINE messages are dropped if the MAC addresses are found in the snooping database, but the binding's interface is other than the interface where the message was received.

- On untrusted DHCP client interfaces, the switch drops DHCP packets with a source MAC address that does not match the client hardware address. This is a configurable option.

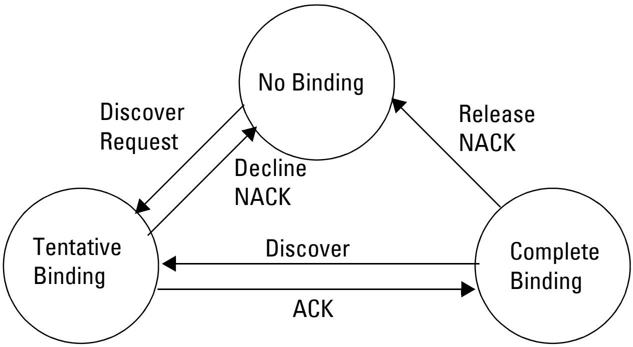
How Is the DHCP Snooping Bindings Database Populated?

The DHCP snooping application uses DHCP messages to build and maintain the binding's database. DHCP snooping creates a tentative binding from DHCP DISCOVER and REQUEST messages. Tentative bindings tie a client to a port (the port where the DHCP client message was received). Tentative bindings are completed when DHCP snooping learns the client's IP address from a DHCP ACK message on a trusted port. DHCP snooping removes bindings in response to DECLINE, RELEASE, and NACK messages. The DHCP snooping application ignores the ACK messages as a reply to the DHCP Inform messages received on trusted ports. Static bindings can also be entered into the binding database.

When a switch learns of new bindings or loses bindings, the switch immediately updates the entries in the database. The switch also updates the entries in the binding file. The frequency at which the file is updated is based on a configurable delay, and the updates are batched.

If the absolute lease time of the snooping database entry expires, that entry is removed. Make sure the system time is consistent across the reboots. Otherwise, the snooping entries will not expire properly. If a host sends a DHCP release while the switch is rebooting, when the switch receives the DHCP discovery or request, the client's binding goes to the tentative binding as shown in Figure 27-1.

Figure 27-1. DHCP Binding



The binding database includes data for clients only on untrusted ports.

DHCP Snooping and VLANs

DHCP snooping forwards valid DHCP client messages received on non-routing VLANs. The message is forwarded on all trusted interfaces in the VLAN.

DHCP snooping can be configured on switching VLANs and routing VLANs. When a DHCP packet is received on a routing VLAN, the DHCP snooping application applies its filtering rules and updates the bindings database. If a client message passes filtering rules, the message is placed into the software forwarding path where it may be processed by the DHCP relay agent, the local DHCP server, or forwarded as an IP packet.

DHCP Snooping Logging and Rate Limits

The DHCP snooping application processes incoming DHCP messages. For DHCPRELEASE and DHCPDECLINE messages, the application compares the receive interface and VLAN with the client interface and VLAN in the bindings database. If the interfaces do not match, the application logs the event and drops the message. For valid client messages, DHCP snooping compares the source MAC address to the DHCP client hardware address. When there is a mismatch, DHCP snooping drops the packet and generates a log message if logging of invalid packets is enabled.

If DHCP relay co-exists with DHCP snooping, DHCP client messages are sent to DHCP relay for further processing.

To prevent DHCP packets from being used as a DoS attack when DHCP snooping is enabled, the snooping application enforces a rate limit for DHCP packets received on interfaces. DHCP rate limiting can be configured on both trusted and untrusted interfaces. DHCP snooping monitors the receive rate on each interface separately. If the receive rate exceeds a configurable limit, DHCP snooping diagnostically disables the interface. Administrative intervention is necessary to enable the port, either by using the **no shutdown** command in Interface Config mode or on the **Switching → Ports → Port Configuration** page. Use the **ip dhcp snooping limit none** command to disable diagnostic disabling of the port due to DHCP snooping.

What Is IP Source Guard?

IPSG is a security feature that filters IP packets based on source ID. This feature helps protect the network from attacks that use IP address spoofing to compromise or overwhelm the network.

The source ID may be either the source IP address or a {source IP address, source MAC address} pair. The following can be configured:

- Whether enforcement includes the source MAC address
- Static authorized source IDs

The DHCP snooping bindings database and static IPSG entries identify authorized source IDs. IPSG can be enabled on physical and LAG ports.

If you enable IPSG on a port where DHCP snooping is disabled or where DHCP snooping is enabled but the port is trusted, all IP traffic received on that port is dropped depending on the admin-configured IPSG entries.

IPSG and Port Security

IPSG interacts with port security, also known as port MAC locking, (see "Port-based Security—Port MAC Locking " on page 623) to enforce the source MAC address. Port security controls source MAC address learning in the layer-2 forwarding database (MAC address table). When a frame is received with a previously unlearned source MAC address, port security queries the IPSG feature to determine whether the MAC address belongs to a valid binding.

If IPSG is disabled on the ingress port, IPSG replies that the MAC is valid. If IPSG is enabled on the ingress port, IPSG checks the bindings database. If the MAC address is in the bindings database and the binding matches the VLAN the frame was received on, IPSG replies that the MAC is valid. If the MAC is not in the bindings database, IPSG informs port security that the frame is a security violation.

In the case of an IPSG violation, port security takes whatever action it normally takes upon receipt of an unauthorized frame. Port security limits the number of MAC addresses to a configured maximum. If the limit n is less than the number of stations m in the bindings database, port security allows only n stations to use the port. If $n > m$, port security allows only the stations in the bindings database. For information about configuring the Port Security feature, see "Port and System Security " on page 623.

What is Dynamic ARP Inspection?

Dynamic ARP Inspection (DAI) is a security feature that rejects invalid and malicious ARP packets. DAI prevents a class of man-in-the-middle attacks where an unfriendly station intercepts traffic for other stations by poisoning the ARP caches of its unsuspecting neighbors. The malicious attacker sends ARP requests or responses mapping another station's IP address to its own MAC address.

When DAI is enabled, the switch drops ARP packets whose sender MAC address and sender IP address do not match an entry in the DHCP snooping bindings database. Additional ARP packet validation can optionally be configured.

When DAI is enabled on a VLAN, DAI is enabled on the interfaces (physical ports or LAGs) that are members of that VLAN. Individual interfaces are configured as trusted or untrusted. The trust configuration for DAI is independent of the trust configuration for DHCP snooping.

Optional DAI Features

If the network administrator has configured the option, DAI verifies that the sender MAC address equals the source MAC address in the Ethernet header. There is a configurable option to verify that the target MAC address equals the destination MAC address in the Ethernet header. This check applies only to ARP responses, since the target MAC address is unspecified in ARP requests. IP address checking can also be enabled. When this option is enabled, DAI drops ARP packets with an invalid IP address. The following IP addresses are considered invalid:

- 0.0.0.0
- 255.255.255.255
- all IP multicast addresses
- all class E addresses (240.0.0.0/4)
- loopback addresses (in the range 127.0.0.0/8)

DAI can also be configured to rate-limit ARP requests on untrusted interfaces. If the configured rate is exceeded, DAI diagnostically disables the port on which the rate limit was exceeded. Use the **no shutdown** command to re-enable the port. DAI rate limiting cannot be enabled on trusted interfaces. Use the **no ip arp inspection limit** command to disable diagnostic disabling of untrusted ports due to DAI.

Why Is Traffic Snooping and Inspection Necessary?

DHCP Snooping, IPSG, and DAI are security features that can help protect the switch and the network against various types of accidental or malicious attacks. It might be a good idea to enable these features on ports that provide network access to hosts that are in physically unsecured locations or if network users connect nonstandard hosts to the network.

For example, if an employee unknowingly connects a workstation to the network that has a DHCP server, and the DHCP server is enabled, hosts that attempt to acquire network information from the legitimate network DHCP server might obtain incorrect information from the rogue DHCP server. However, if the workstation with the rogue DHCP server is connected to a port that is configured as untrusted and is a member of a DHCP Snooping-enabled VLAN, the port discards the DHCP server messages.

Default Traffic Snooping and Inspection Values

DHCP snooping is disabled globally and on all VLANs by default. Ports are untrusted by default.

Table 27-1. Traffic Snooping Defaults

Parameter	Default Value
DHCP snooping mode	Disabled
DHCP snooping VLAN mode	Disabled on all VLANs
Interface trust state	Disabled (untrusted)
DHCP logging invalid packets	Disabled
DHCP snooping rate limit	15 packets per second
DHCP snooping burst interval	1 second
DHCP snooping binding database storage	Local
DHCP snooping binding database write delay	300 seconds

Table 27-1. Traffic Snooping Defaults (Continued)

Parameter	Default Value
Static DHCP bindings	None configured
IPSG mode	Disabled on all interfaces
IPSG port security	Disabled on all interfaces
Static IPSG bindings	None configured
DAI validate source MAC	Disabled
DAI validate destination MAC	Disabled
DAI validate IP	Disabled
DAI trust state	Disabled (untrusted)
DAI Rate limit	Disabled
DAI Burst interval	1 second
DAI mode	Disabled on all VLANs
DAI logging invalid packets	Disabled
DAI ARP ACL	None configured
DAI Static flag	Disabled (validation by ARP ACL and DHCP snooping binding database)

Configuring Traffic Snooping and Inspection (Web)

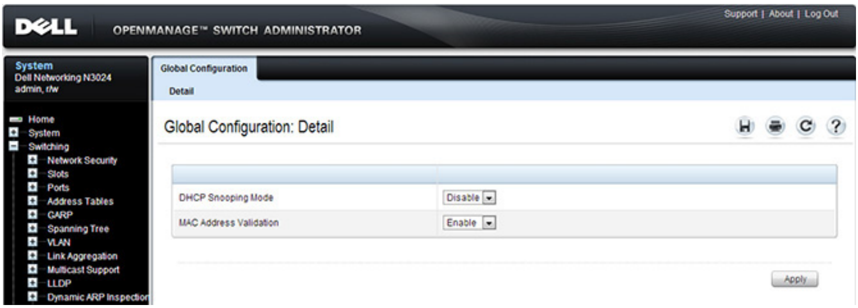
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring DHCP snooping, IPSG, and DAI features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

DHCP Snooping Configuration

Use the **DHCP Snooping Configuration** page to control the DHCP Snooping mode on the switch and to specify whether the sender MAC Address for DHCP Snooping must be validated.

To access the **DHCP Snooping Configuration** page, click **Switching** → **DHCP Snooping** → **Global Configuration** in the navigation panel.

Figure 27-2. DHCP Snooping Configuration

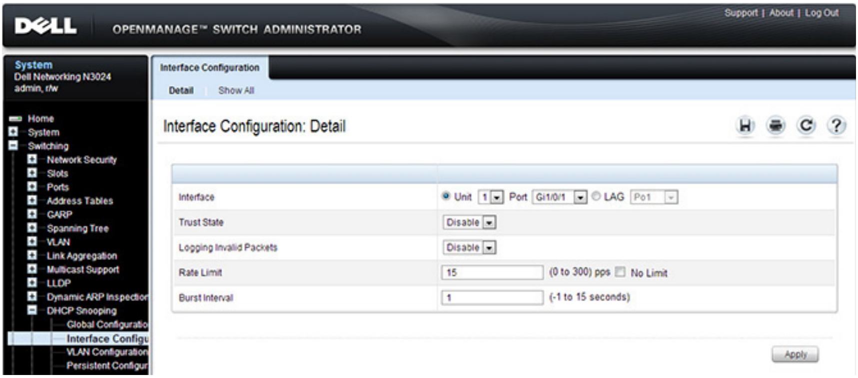


DHCP Snooping Interface Configuration

Use the **DHCP Snooping Interface Configuration** page to configure the DHCP Snooping settings on individual ports and LAGs.

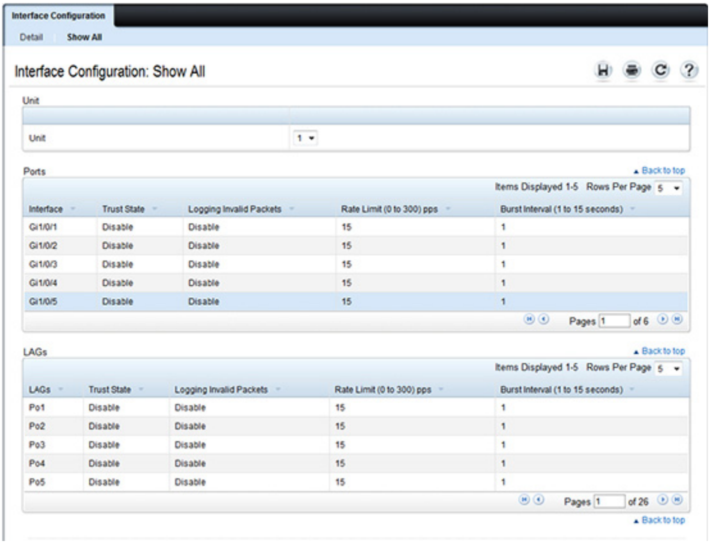
To access the **DHCP Snooping Interface Configuration** page, click **Switching** → **DHCP Snooping** → **Interface Configuration** in the navigation panel.

Figure 27-3. DHCP Snooping Interface Configuration



To view a summary of the DHCP snooping configuration for all interfaces, click **Show All**.

Figure 27-4. DHCP Snooping Interface Configuration Summary

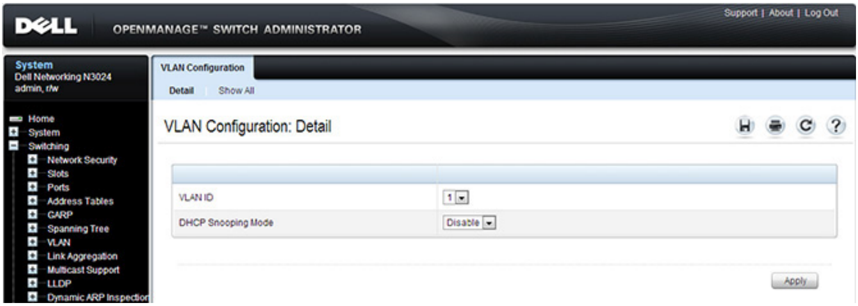


DHCP Snooping VLAN Configuration

Use the **DHCP Snooping VLAN Configuration** page to control the DHCP snooping mode on each VLAN.

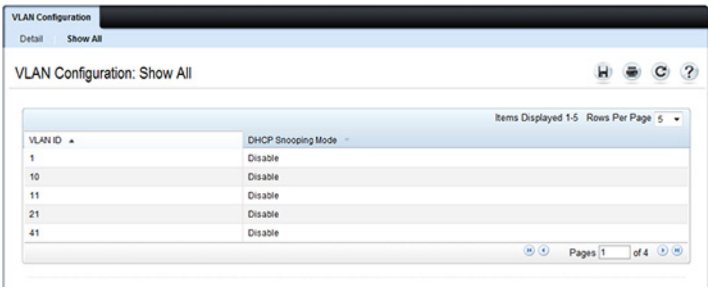
To access the **DHCP Snooping VLAN Configuration** page, click **Switching** → **DHCP Snooping** → **VLAN Configuration** in the navigation panel.

Figure 27-5. DHCP Snooping VLAN Configuration



To view a summary of the DHCP snooping status for all VLANs, click **Show All**.

Figure 27-6. DHCP Snooping VLAN Configuration Summary

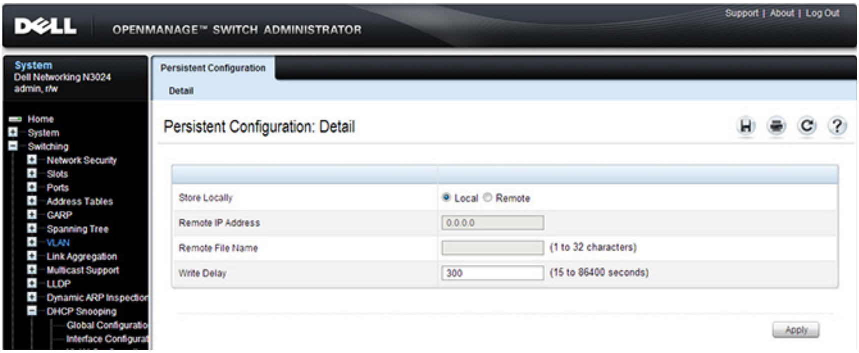


DHCP Snooping Persistent Configuration

Use the **DHCP Snooping Persistent Configuration** page to configure the persistent location of the DHCP snooping database. The bindings database can be stored locally on the switch or on a remote system somewhere else in the network. The switch must be able to reach the IP address of the remote system to send bindings to a remote database.

To access the **DHCP Snooping Persistent Configuration** page, click **Switching** → **DHCP Snooping** → **Persistent Configuration** in the navigation panel.

Figure 27-7. DHCP Snooping Persistent Configuration

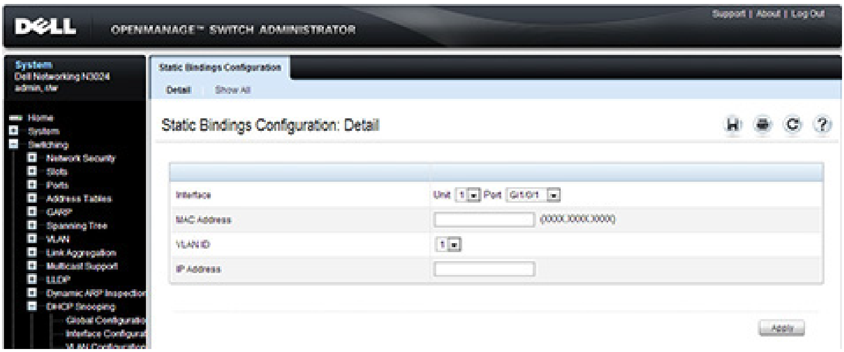


DHCP Snooping Static Bindings Configuration

Use the **DHCP Snooping Static Bindings Configuration** page to add static DHCP bindings to the binding database.

To access the **DHCP Snooping Static Bindings Configuration** page, click **Switching** → **DHCP Snooping** → **Static Bindings Configuration** in the navigation panel.

Figure 27-8. DHCP Snooping Static Bindings Configuration



To view a summary of the DHCP snooping status for all VLANs, click **Show All**.

Figure 27-9. DHCP Snooping Static Bindings Summary

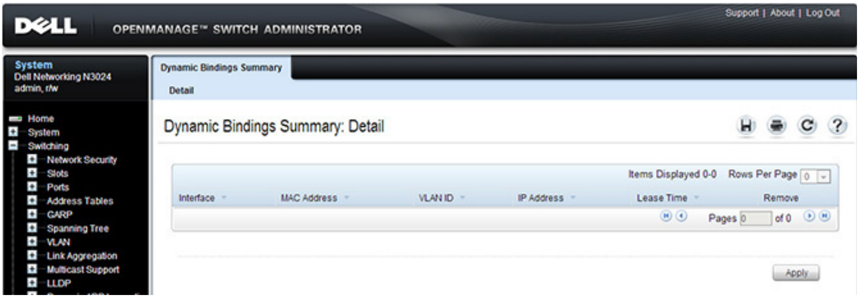
To remove a static binding, select the **Remove** checkbox associated with the binding and click **Apply**.

DHCP Snooping Dynamic Bindings Summary

The **DHCP Snooping Dynamic Bindings Summary** lists all the DHCP snooping dynamic binding entries learned on the switch ports.

To access the **DHCP Snooping Dynamic Bindings Summary** page, click **Switching → DHCP Snooping → Dynamic Bindings Summary** in the navigation panel.

Figure 27-10. DHCP Snooping Dynamic Bindings Summary

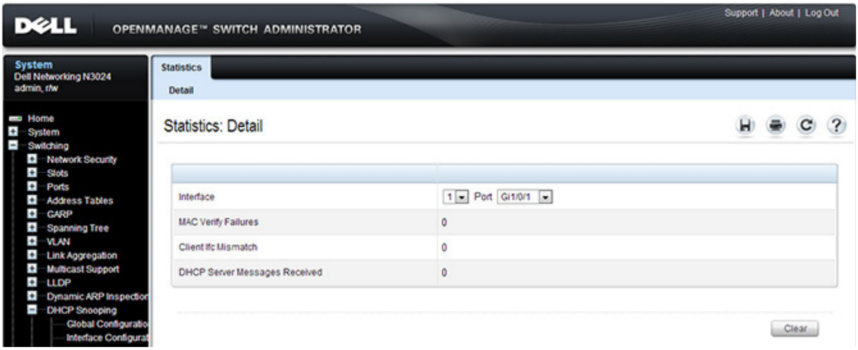


DHCP Snooping Statistics

The **DHCP Snooping Statistics** page displays DHCP snooping interface statistics.

To access the **DHCP Snooping Statistics** page, click **Switching** → **DHCP Snooping** → **Statistics** in the navigation panel.

Figure 27-11. DHCP Snooping Statistics

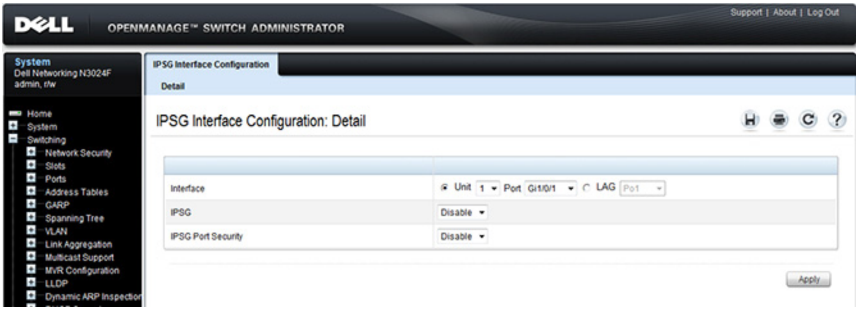


IPSG Interface Configuration

Use the **IPSG Interface Configuration** page to configure IPSG on an interface.

To access the **IPSG Interface Configuration** page, click **Switching** → **IP Source Guard** → **IPSG Interface Configuration** in the navigation panel.

Figure 27-12. IPSG Interface Configuration

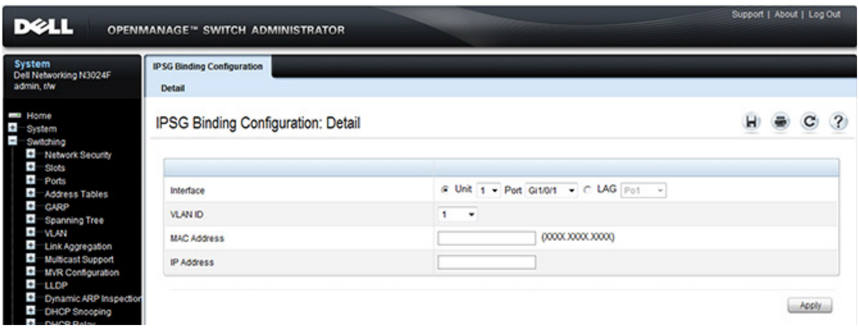


IPSG Binding Configuration

Use the **IPSG Binding Configuration** page displays DHCP snooping interface statistics.

To access the **IPSG Binding Configuration** page, click **Switching** → **IP Source Guard** → **IPSG Binding Configuration** in the navigation panel.

Figure 27-13. IPSG Binding Configuration

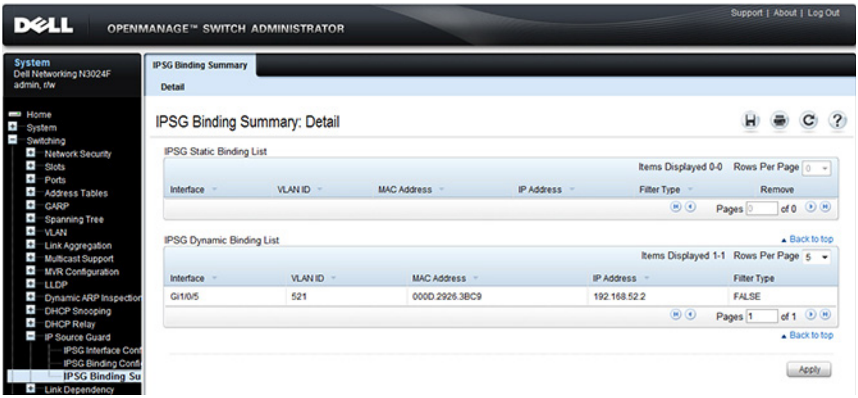


IPSG Binding Summary

The **IPSG Binding Summary** page displays the IPSG Static binding list and IPSG dynamic binding list (the static bindings configured in Binding configuration page).

To access the **IPSG Binding Summary** page, click **Switching** → **IP Source Guard** → **IPSG Binding Summary** in the navigation panel.

Figure 27-14. IPSG Binding Summary

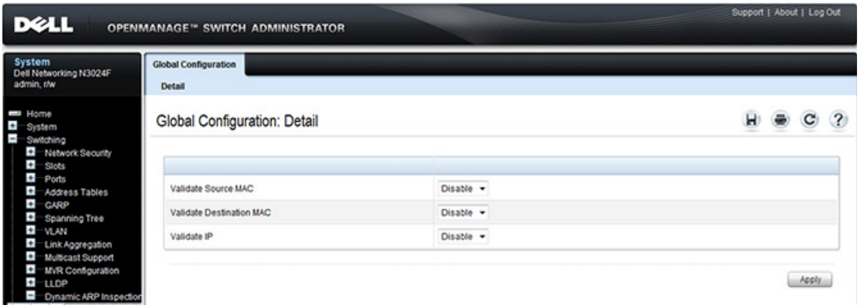


DAI Global Configuration

Use the **DAI Configuration** page to configure global DAI settings.

To display the **DAI Configuration** page, click **Switching** → **Dynamic ARP Inspection** → **Global Configuration** in the navigation panel.

Figure 27-15. Dynamic ARP Inspection Global Configuration

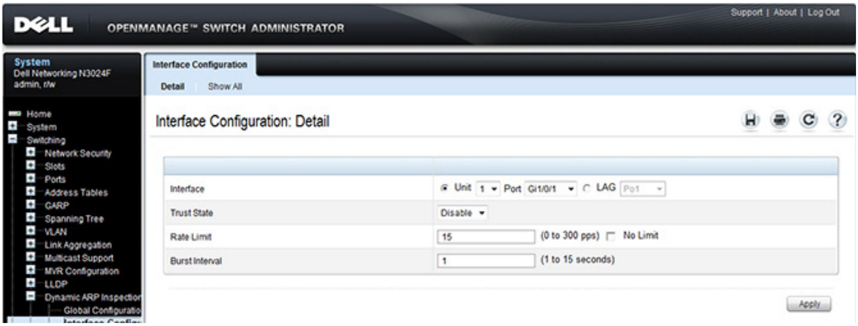


DAI Interface Configuration

Use the **DAI Interface Configuration** page to select the DAI Interface for which information is to be displayed or configured.

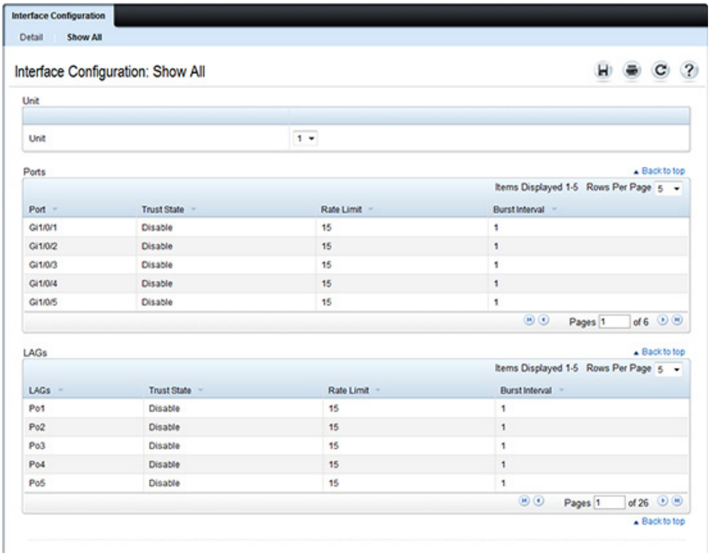
To display the **DAI Interface Configuration** page, click **Switching** → **Dynamic ARP Inspection** → **Interface Configuration** in the navigation panel.

Figure 27-16. Dynamic ARP Inspection Interface Configuration



To view a summary of the DAI status for all interfaces, click **Show All**.

Figure 27-17. DAI Interface Configuration Summary

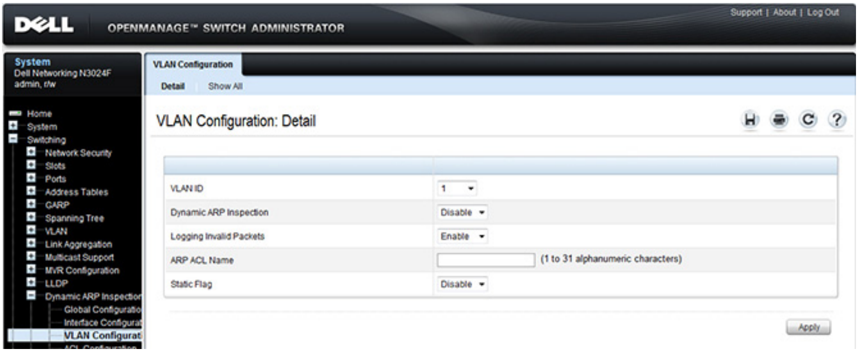


DAI VLAN Configuration

Use the DAI **VLAN Configuration** page to select the VLANs for which information is to be displayed or configured.

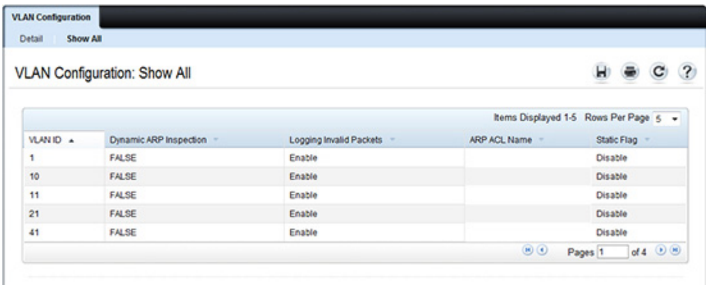
To display the DAI **VLAN Configuration** page, click **Switching** → **Dynamic ARP Inspection** → **VLAN Configuration** in the navigation panel.

Figure 27-18. Dynamic ARP Inspection VLAN Configuration



To view a summary of the DAI status for all VLANs, click **Show All**.

Figure 27-19. Dynamic ARP Inspection VLAN Configuration Summary

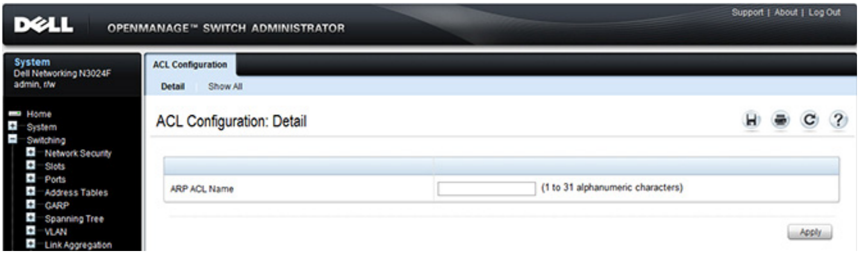


DAI ACL Configuration

Use the DAI **ACL Configuration** page to add or remove ARP ACLs.

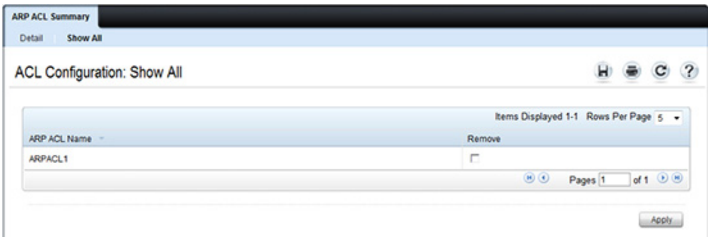
To display the DAI **ACL Configuration** page, click **Switching** → **Dynamic ARP Inspection** → **ACL Configuration** in the navigation panel.

Figure 27-20. Dynamic ARP Inspection ACL Configuration



To view a summary of the ARP ACLs that have been created, click **Show All**.

Figure 27-21. Dynamic ARP Inspection ACL Summary



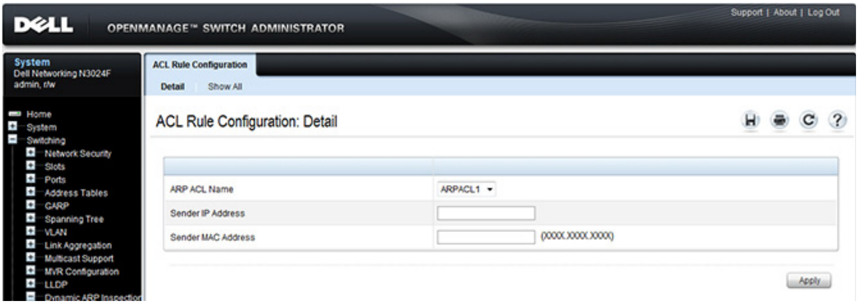
To remove an ARP ACL, select the **Remove** checkbox associated with the ACL and click **Apply**.

DAI ACL Rule Configuration

Use the DAI ARP **ACL Rule Configuration** page to add or remove DAI ARP ACL Rules.

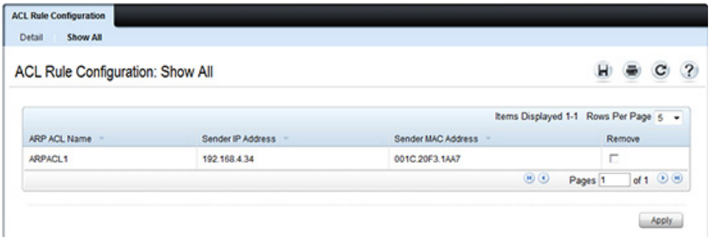
To display the DAI ARP **ACL Rule Configuration** page, click **Switching** → **Dynamic ARP Inspection** → **ACL Rule Configuration** in the navigation panel.

Figure 27-22. Dynamic ARP Inspection Rule Configuration



To view a summary of the ARP ACL rules that have been created, click **Show All**.

Figure 27-23. Dynamic ARP Inspection ACL Rule Summary



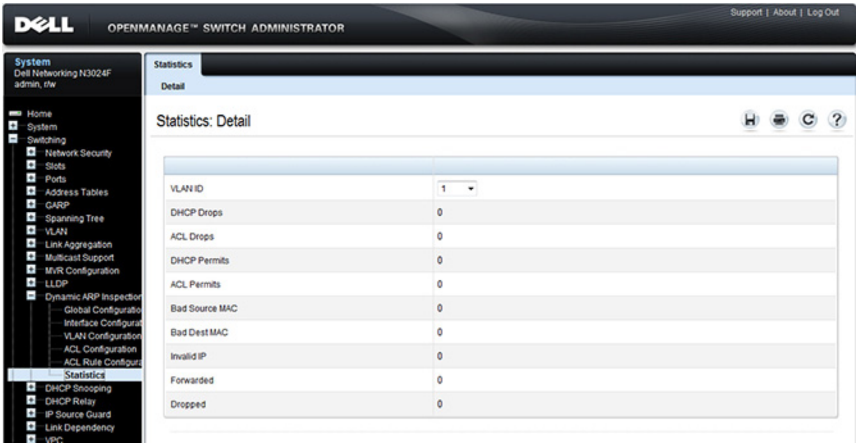
To remove an ARP ACL rule, select the **Remove** checkbox associated with the rule and click **Apply**.

DAI Statistics

Use the DAI **Statistics** page to display the statistics per VLAN.

To display the DAI **Statistics** page, click **Switching** → **Dynamic ARP Inspection** → **Statistics** in the navigation panel.

Figure 27-24. Dynamic ARP Inspection Statistics



Configuring Traffic Snooping and Inspection (CLI)

This section provides information about the commands used for configuring DHCP snooping, IPSG, and DAI settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring DHCP Snooping

Beginning in Privileged EXEC mode, use the following commands to configure and view DHCP snooping settings.

Command	Purpose
configure	Enter global configuration mode.
ip dhcp snooping	Enable DHCP snooping on the switch.
ipv6 dhcp snooping	Enable IPv6 DHCP snooping on the switch.
ip dhcp snooping verify mac-address	Enable the verification of the source MAC address with the client MAC address in the received DHCP message.
ip dhcp snooping log-invalid	Enable the logging of DHCP messages filtered by the DHCP Snooping application.
ip dhcp snooping binding mac-address vlan vlan-id ip-address interface interface	Configure a static binding in the DHCP snooping static bindings database. • *mac-address*—The client's MAC address. • *vlan-id*—The number of the VLAN the client is authorized to use. • *ip-address*—The IP address of the client. • *interface*—The interface on which the client is authorized. The form is unit/port.
ip dhcp snooping database {local | tftp://*hostIP/filename*}	Configure the persistent storage location of the DHCP snooping database. • *hostIP*—The IP address of the remote host. • *filename*—The name of the file for the database on the remote host.

Command	Purpose
ip dhcp snooping database write-delay *seconds*	Configure the interval, in seconds, at which the DHCP Snooping database will be stored in persistent storage. The number of seconds can range from 15–86400.
ip dhcp snooping limit {**none** | **rate** *rate* [**burst interval** *seconds*]}	Configure the maximum rate of DHCP messages allowed on the switch at any given time. • *rate*—The maximum number of packets per second allowed (Range: 0–300 pps). • *seconds*—The time allowed for a burst (Range: 1–15 seconds).
interface *interface*	Enter interface configuration mode for the specified port or LAG. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. For a LAG, the interface type is **port-channel**. A range of ports can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
ip dhcp snooping trust	Configure the interface (or range of interfaces) as a trusted port. DHCP server messages are not filtered on trusted ports.
exit	Exit to Global Configuration mode.
interface [**range**] **vlan** *vlan id*	Enter interface configuration mode for the specified VLAN or range of VLANs.
CTRL + Z	Exit to Privileged EXEC mode.
show ip dhcp snooping [**interfaces**]	View the DHCP snooping global and per port configuration.
show ip dhcp snooping binding [{**static** | **dynamic**}] [**interface** *port*] [**vlan** *vlan-id*]	View the entries in the DHCP snooping bindings database.
show ip dhcp snooping database	View information about the persistent database configuration.
show ip dhcp snooping statistics	View the DHCP snooping statistics.

Command	Purpose
clear ip dhcp snooping statistics	Reset the DHCP snooping statistics to zero.

Configuring IP Source Guard

Beginning in Privileged EXEC mode, use the following commands to configure IPSG settings on the switch.

Command	Purpose
configure	Enter global configuration mode.
interface *interface*	Enter interface configuration mode for the specified port or LAG. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. For a LAG, the interface type is **port-channel**. A range of ports can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
ip verify source [port-security] and **ipv6 verify source [port-security]**	Enable IPSG on the port or LAG to prevent packet forwarding if the source IP address in the packet is not in the DHCP snooping binding database. Use the optional **port-security keyword** to also prevent packet forwarding if the sender MAC address is not in forwarding database table or the DHCP snooping binding database. **NOTE:** To enforce filtering based on the source MAC address, port security must also be enabled on the interface by using the **port security** command in Interface Configuration mode.
exit	Exit to Global Config mode.
ip verify binding *mac_addr* **vlan** *vlan_id* *ipaddr* **interface** *interface* and **ipv6 verify binding** *mac_addr* **vlan** *vlan_id* *ipaddr* **interface** *interface*	Configure a static binding for IPSG.

Command	Purpose
exit	Exit to Privileged EXEC mode.
show ip verify interface *interface*	View IPSG parameters for a specific port or LAG. The *interface* parameter includes the interface type (**gigabitethernet**, **tengigabitethernet**, or **port-channel**) and number.
show ip verify source [**interface** *interface*]	View IPSG bindings configured on the switch or on a specific port or LAG.
show ip source binding	View IPSG bindings.

Configuring Dynamic ARP Inspection

Beginning in Privileged EXEC mode, use the following commands to configure DAI settings on the switch.

Command	Purpose
configure	Enter global configuration mode.
ip arp inspection vlan vlan-range [**logging**]	Enable Dynamic ARP Inspection on a single VLAN or a range of VLANs. Use the **logging** keyword to enable logging of invalid packets.
ip arp inspection validate {[**src-mac**] [**dst-mac**] [**ip**]}	Enable additional validation checks like source MAC address validation, destination MAC address validation, or IP address validation on the received ARP packets. Each command overrides the configuration of the previous command. For example, if a command enables source MAC address and destination validations and a second command enables IP address validation only, the source MAC address and destination MAC address validations are disabled as a result of the second command. • **src-mac**—For validating the source MAC address of an ARP packet. • **dst-mac**—For validating the destination MAC address of an ARP packet. • **ip**—For validating the IP address of an ARP packet.

Command	Purpose
arp access-list *acl-name*	Create an ARP ACL with the specified name (1–31 characters) and enter ARP Access-list Configuration mode for the ACL.
permit ip host *sender-ip* **mac host** *sender-mac*	Configure a rule for a valid IP address and MAC address combination used in ARP packet validation. • *sender-ip* — Valid IP address used by a host. • *sender-mac* —Valid MAC address in combination with the above sender-ip used by a host.
exit	Exit to Global Config mode.
ip arp inspection filter *acl-name* **vlan** *vlan-range* [**static**]	Configure the ARP ACL to be used for a single VLAN or a range of VLANs to filter invalid ARP packets. Use the **static** keyword to indicate that packets that do not match a permit statement are dropped without consulting the DHCP snooping bindings.
interface *interface*	Enter interface configuration mode for the specified port or LAG. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. For a LAG, the interface type is **port-channel**. A range of ports can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
ip arp inspection limit {**none** | **rate** *pps* [**burst interval** *seconds*]}	Configure the rate limit and burst interval values for an interface.Use the keyword **none** to specify that the interface is not rate limited for Dynamic ARP Inspection. • **none** — To set no rate limit. • *pps* — Packets per second (Range: 0–300). • *seconds* — The number of seconds (Range: 1–15).
ip arp inspection trust	Specify that the interface as trusted for Dynamic ARP Inspection.
CTRL + Z	Exit to Privileged EXEC mode.
show ip arp inspection interfaces [*interface*]	View the Dynamic ARP Inspection configuration on all the DAI-enabled interfaces or for the specified interface.

Command	Purpose
show ip arp inspection vlan [*vlan-range*]	View the Dynamic ARP Inspection configuration on the specified VLAN(s). This command also displays the global configuration values for source MAC validation, destination MAC validation and invalid IP validation.
show ip arp inspection statistics [**vlan** *vlan-range*]	View the statistics of the ARP packets processed by Dynamic ARP Inspection for the switch or for the specified VLAN(s).
show arp access-list [*acl-name*]	View all configured ARP ACL and their rules, or use the ACL name to view information about that ARP ACL only.

Traffic Snooping and Inspection Configuration Examples

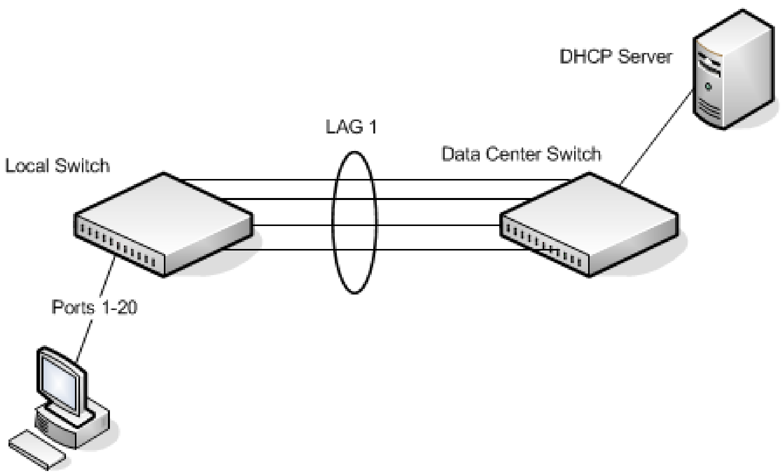
This section contains the following examples:

- Configuring DHCP Snooping
- Configuring IPSG

Configuring DHCP Snooping

In this example, DHCP snooping is enabled on VLAN 100. Ports 1-20 connect end users to the network and are members of VLAN 100. These ports are configured to limit the maximum number of DHCP packets with a rate limit of 100 packets per second. LAG 1, which is also a member of VLAN 100 and contains ports 21-24, is the trunk port that connects the switch to the data center, so it is configured as a trusted port.

Figure 27-25. DHCP Snooping Configuration Topology



The commands in this example also enforce rate limiting and remote storage of the bindings database. The switch has a limited amount of storage space in NVRAM and flash memory, so the administrator specifies that the DHCP snooping bindings database is stored on an external TFTP server.

To configure the switch:

1 Enable DHCP snooping on VLAN 100.

```
console#config
console(config)#ip dhcp snooping vlan 100
```

2 Configure LAG 1, which includes ports 21-24, as a trusted port. All other interfaces are untrusted by default.

```
console(config)#interface port-channel 1
console(config-if-Po1)#ip dhcp snooping trust
console(config-if-Po1)#exit
```

3 Enter interface configuration mode for all untrusted interfaces (ports 1-20) and limit the number of DHCP packets that an interface can receive to 100 packets per second. LAG 1 is a trusted port and keeps the default value for rate limiting (unlimited).

```
console(config)#interface range gi1/0/1-20
console(config-if)#ip dhcp snooping limit rate 100
console(config-if)#exit
```

4 Specify that the DHCP snooping database is to be stored remotely in a file called dsDb.txt on a TFTP server with and IP address of 10.131.11.1.

```
console(config)#ip dhcp snooping database
tftp://10.131.11.1/dsDb.txt
console(config)#exit
```

5 Enable DHCP snooping for the switch

```
console(config)#ip dhcp snooping
```

6 View DHCP snooping information.

```
console#show ip dhcp snooping

DHCP snooping is Enabled
DHCP snooping source MAC verification is enabled
DHCP snooping is enabled on the following VLANs:
100

 Interface    Trusted     Log Invalid Pkts
-----------  ----------   ----------------
```

Configuring IPSG

This example builds on the previous example and uses the same topology shown in Figure 27-25. In this configuration example, IP source guard is enabled on ports 1-20. DHCP snooping must also be enabled on these ports. Additionally, because the ports use IP source guard with source IP and MAC address filtering, port security must be enabled on the ports as well.

To configure the switch:

1 Enter interface configuration mode for the host ports and enable IPSG.

```
console(config)#interface range gi1/0/1-20
console(config-if)#ip verify source port-security
```

2 Enable port security on the ports.

```
console(config-if)#port security
```

3 View IPSG information.

```
console#show ip verify source

Interface   Filter    IP Address     MAC Address         Vlan
---------  -------    ----------
Gi1/0/1    ip-mac     192.168.3.45   00:1C:23:55:D4:8E   100
Gi1/0/2    ip-mac     192.168.3.40   00:1C:23:12:44:B6   100
Gi1/0/3    ip-mac     192.168.3.33   00:1C:23:AA:B8:01   100
Gi1/0/4    ip-mac     192.168.3.18   00:1C:23:67:D3:CC   100
Gi1/0/5    ip-mac     192.168.3.49   00:1C:23:55:1B:6E   100
--More-- or (q)uit
```

28

Link Aggregation

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to create and configure link aggregation groups (LAGs), which are also known as port-channels.

The topics covered in this chapter include:

- Link Aggregation
- Multi-Switch LAG (MLAG)

Link Aggregation

Overview

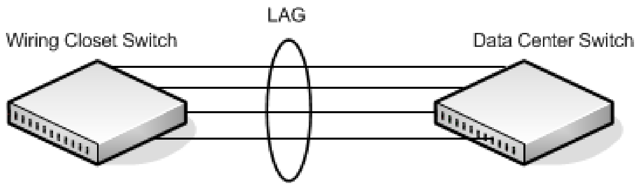
Link Aggregation allows one or more full-duplex Ethernet links of the same speed to be aggregated together to form a LAG. This allows the switch to treat the LAG as if it is a single link. The Dell Networking N-Series switches support industry-standard LAGs that adhere to the IEEE 802.3ad specification.

Assignment of interfaces to LAGs is based on a system limit of 144 interfaces assigned to LAGs, a maximum of 72 dynamic LAGs (or 128 static LAGs) and a maximum of 8 interfaces per LAG. For example, 72 dynamic LAGs may be assigned 2 interfaces each, or 18 dynamic or static LAGs may be assigned 8 interfaces each. Alternatively, 128 interfaces can be assigned to 128 static LAGs with each LAG containing a single port.

Each LAG can consist of up to eight 1 Gbps ports (for the Dell Networking N1500, N2000, and N3000 Series) or eight 10 Gbps (for the Dell Networking N4000 Series) ports (or even eight 40 Gbps interfaces). When eight Gigabit Ethernet ports are configured as a LAG, the maximum bandwidth for the single, logical interface is 8 Gbps, and when eight 10 Gbps ports are configured as a LAG, the maximum bandwidth for the single, logical interface is 80 Gbps.

Figure 28-1 shows an example of a switch in the wiring closet connected to a switch in the data center by a LAG that consists of four physical 10 Gbps links. The LAG provides full-duplex bandwidth of 40 Gbps between the two switches.

Figure 28-1. LAG Configuration



LAGs can be configured on stand-alone or stacked switches. In a stack of switches, the LAG can consist of ports on a single unit or across multiple stack members. When a LAG members span different units across a stack, and a unit fails, the remaining LAG members on the functional units continue to handle traffic for the LAG.

Why Are Link Aggregation Groups Necessary?

The primary purpose of LAGs is to increase the overall bandwidth between two switches. This is accomplished by effectively aggregating multiple ports together that act as a single, logical connection between the two switches.

LAGs also provide redundancy. If a link fails, traffic is automatically redistributed across the remaining links.

What Is the Difference Between Static and Dynamic Link Aggregation?

Link aggregation can be configured as either dynamic or static. Dynamic configuration is supported using the IEEE 802.3ad standard, which is known as Link Aggregation Control Protocol (LACP). Static configuration is used when connecting a Dell Networking N1500, N2000, N3000, and N4000 Series switches to an external Gigabit Ethernet switch that does not support LACP.

One advantage of LACP is that the protocol enables the switch to confirm that the external switch is also configured for link aggregation. When using static configuration, a cabling or configuration mistake involving the Dell Networking N2000/N3000/N4000 switch or the external switch could go undetected and thus cause undesirable network behavior. Both static and dynamic LAGs (via LACP) can detect physical link failures within the LAG and continue forwarding traffic through the other connected links within that same LAG. LACP can also detect switch or port failures that do not result in loss of link. This provides a more resilient LAG. Best practices suggest using dynamic link aggregation instead of static link aggregation. When a port is added to a LAG as a static member, it neither transmits nor receives LACP PDUs.

What is LAG Hashing?

Dell Networking N-Series switches support configuration of hashing algorithms for each LAG interface. The hashing algorithm is used to distribute traffic load among the physical ports of the LAG while preserving the per-flow packet order. Enhanced hashing mode is the recommended and default hashing mode for Dell Networking N-Series switches.

The hashing algorithm uses various packet attributes to determine the outgoing physical port.

The switch supports the following set of packet attributes to be used for hash computation:

- Source MAC, VLAN, EtherType, and incoming port.
- Destination MAC, VLAN, EtherType, and incoming port.
- Source IP and Source TCP/UDP port numbers.
- Destination IP and Destination TCP/UDP port numbers.
- Source/Destination MAC, VLAN, EtherType, and incoming port.
- Source/Destination IP and Source/Destination TCP/UDP port numbers.
- Enhanced hashing mode

Enhanced hashing mode has following advantages:

- MODULO-N operation based on the number of ports in the LAG.
- Packet attributes selection based on the packet type. For layer-2 packets, Source and Destination MAC address plus physical source port are used for hash computation. For IP packets, Source IP, Destination IP address, TCP/UDP ports, and physical source port are used.
- Non-Unicast traffic and Unicast traffic is hashed using a common hash algorithm.

- Excellent load balancing performance.

How Do LAGs Interact with Other Features?

From a system perspective, a LAG is treated just as a physical port, with the same configuration parameters for administrative enable/disable, spanning tree port priority, path cost as may be for any other physical port.

VLAN

When members are added to a LAG, they are removed from all existing VLAN membership. When members are removed from a LAG they are added back to the VLANs that they were previously members of as per the configuration file. Note that a port's VLAN membership can still be configured when it's a member of a LAG. However this configuration is only actually applied when the port leaves the LAG.

The LAG interface can be a member of a VLAN complying with IEEE 802.1Q.

STP

Spanning tree does not maintain state for members of a LAG, but does maintain state for the LAG interface. As far as STP is concerned, members of a LAG do not have individual link state. (Internally, the STP state of the LAG interface is replicated for the member links.)

When members are deleted from a LAG they become normal links, and spanning tree maintains their individual link state information.

Statistics

Statistics are collected for LAGs in the same manner as they are collected for the physical ports, in addition to the statistics collected for individual members as per the 802.3ad MIB statistics.

LAG Configuration Guidelines

Ports to be aggregated must be configured so that they are compatible with the link aggregation feature and with the partner switch to which they connect.

Ports to be added to a LAG must meet the following requirements:

- Interface must be a physical Ethernet link.

- Each member of the LAG must be running at the same speed and must be in full duplex mode.
- The port cannot be a mirrored port

The following are the interface restrictions

- The configured speed of a LAG member cannot be changed.
- An interface can be a member of only one LAG.

Default Link Aggregation Values

The LAGs on the switch are created by default, but no ports are members. Table 28-1 summarizes the default values for the MAC address table.

Table 28-1. MAC Address Table Defaults

Parameter	Default Value
LACP system priority	1
LACP port priority	1
LACP timeout	Long
LAG hash algorithm type	Enhanced (7)

Configuring Link Aggregation (Web)

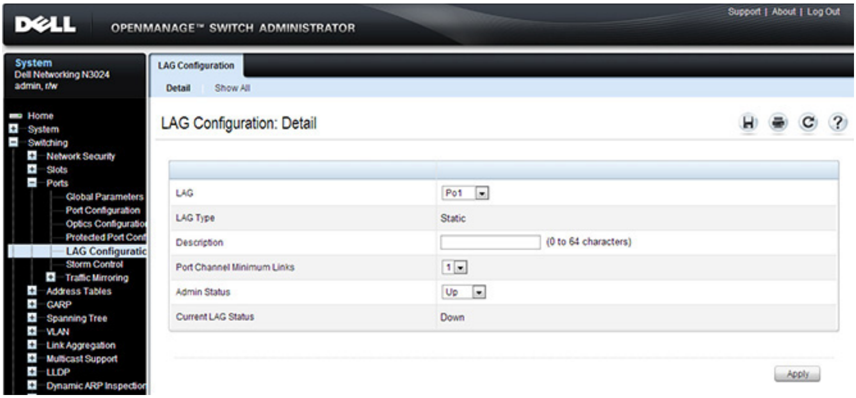
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring LAGs on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

LAG Configuration

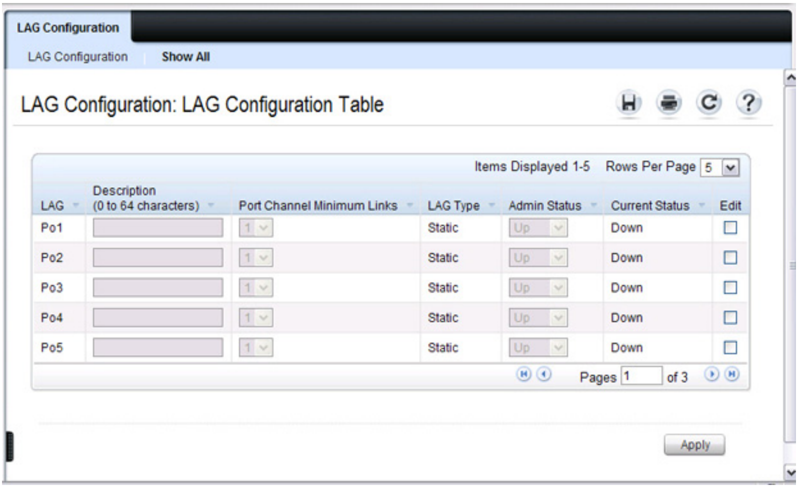
Use the **LAG Configuration** page to set the name and administrative status (up/down) of a LAG.

To display the **LAG Configuration** page, click **Switching → Ports → LAG Configuration** in the navigation panel.

Figure 28-2. LAG Configuration



To view or edit settings for multiple LAGs, click **Show All**.

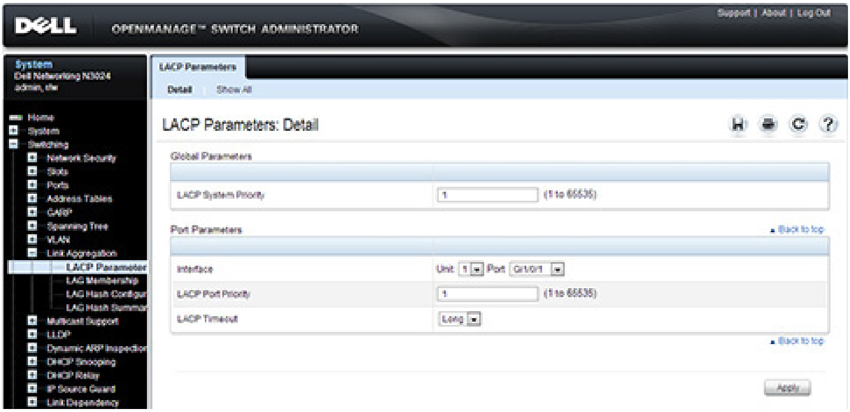


LACP Parameters

Dynamic link aggregation is initiated and maintained by the periodic exchanges of LACP PDUs. Use the **LACP Parameters** page to configure LACP LAGs.

To display the **LACP Parameters** page, click **Switching** → **Link Aggregation** → **LACP Parameters** in the navigation panel.

Figure 28-3. LACP Parameters



Configuring LACP Parameters for Multiple Ports

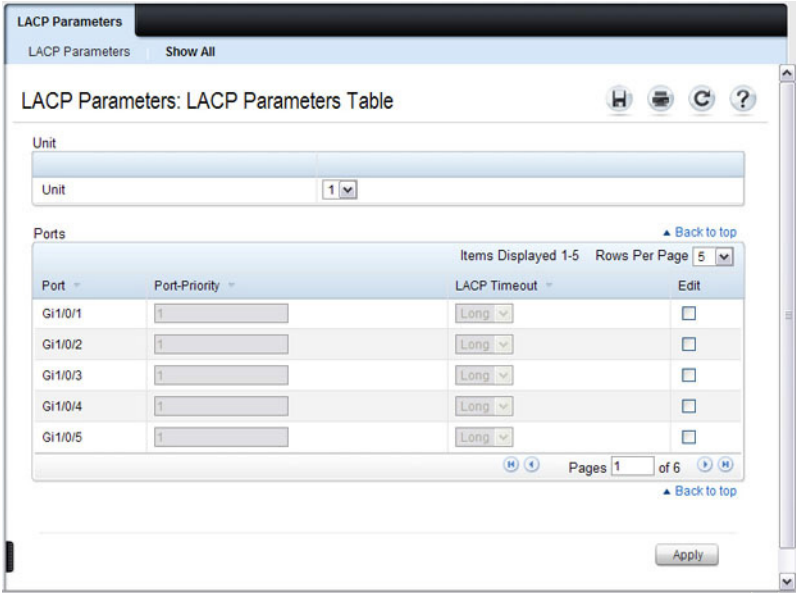
To configure LACP settings:

1 Open the **LACP Parameters** page.

2 Click **Show All**.

The **LACP Parameters Table** page displays.

Figure 28-4. LACP Parameters Table



3 Select the **Edit** check box associated with each port to configure.

4 Specify the LACP port priority and LACP timeout for each port.

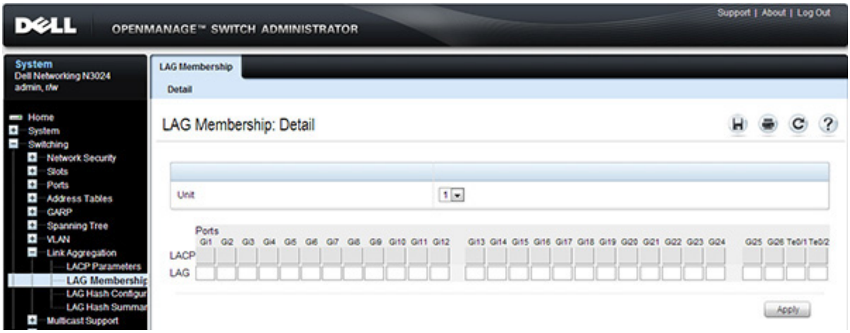
5 Click **Apply**.

LAG Membership

Your switch supports 48 LAGs per system, and eight ports per LAG. Use the **LAG Membership** page to assign ports to static and dynamic LAGs.

To display the **LAG Membership** page, click **Switching** → **Link Aggregation** → **LAG Membership** in the navigation panel.

Figure 28-5. LAG Membership



Adding a Port to a Static LAG

To add a static LAG member:

1 Open the **LAG Membership** page.

2 Click in the **LAG** row to toggle the port to the desired LAG.

  The LAG number displays for that port. The LAG number increases each time you click until the number reaches the maximum LAG number and then returns to blank (no LAG assigned).

3 Click **Apply**.

  The port is assigned to the selected LAG, and the device is updated.

Adding a LAG Port to a Dynamic LAG by Using LACP

To add a dynamic LAG member:

1 Open the **LAG Membership** page.

2 Click in the **LACP** row to toggle the desired LAG port to **L**.

  NOTE: The port must be assigned to a LAG before it can be aggregated to an LACP.

3 Click **Apply**.

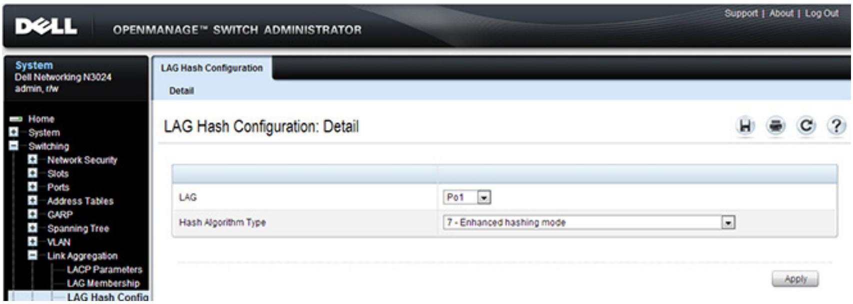
  The LAG port is added as a dynamic LAG member to the selected LAG.

LAG Hash Configuration

Use the LAG hash algorithm to set the traffic distribution mode on the LAG. The hash type can be set for each LAG.

To display the **LAG Hash Configuration** page, click **Switching** → **Link Aggregation** → **LAG Hash Configuration** in the navigation panel.

Figure 28-6. LAG Hash Configuration

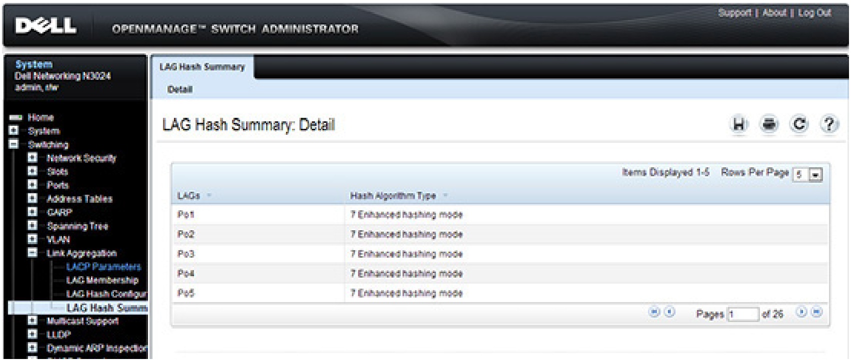


LAG Hash Summary

The **LAG Hash Summary** page lists the channels on the system and their assigned hash algorithm type.

To display the **LAG Hash Summary** page, click **Switching** → **Link Aggregation** → **LAG Hash Summary** in the navigation panel.

Figure 28-7. LAG Hash Summary



Configuring Link Aggregation (CLI)

This section provides information about the commands used for configuring link aggregation settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring LAG Characteristics

Beginning in Privileged EXEC mode, use the following commands to configure a few of the available LAG characteristics. Many of the commands described in "Configuring Port Characteristics (CLI) " on page 611 are also applicable to LAGs.

Command	Purpose
configure	Enter global configuration mode.
interface *interface*	Enter interface configuration mode for the specified LAG. The *interface* variable includes the interface type, which is **port-channel**, and the LAG number, for example **port-channel 3**. A range of LAGs can be specified using the **interface range port-channel** command. For example, **interface range port-channel 3-6** configures LAGs 3, 4, 5, and 6.
description *description*	Configure a description for the LAG or range of LAGs
port-channel min-links *minimum*	Set the minimum number of links that must be up in order for the port-channel interface to be declared up.
exit	Exit to Global Config mode.
CTRL + Z	Exit to Privileged EXEC mode.
show interfaces description port-channel *port-channel number*	View the configured description for the specified LAG.
show interfaces port-channel [*port-channel number*]	View LAG information for the specified LAG or for all LAGs.

Configuring Link Aggregation Groups

Beginning in Privileged EXEC mode, use the following commands to add ports as LAG members and to configure the LAG hashing mode.

Command	Purpose
configure	Enter global configuration mode.
interface *interface*	Enter interface configuration mode for the specified port. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of ports can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
channel-group *port-channel-number* **mode** {**on** | **auto**}	Add the port(s) to the LAG specified with the *port-channel-number* value. Use the auto keyword to add the port(s) as dynamic members, or use **on** to specify that the LAG membership is static. • *port-channel-number* — Number of a valid port-channel for the current port to join. • **on** — Forces the port to join a channel without LACP (static LAG). • **active** — Forces the port to join a channel with LACP (dynamic LAG).
exit	Exit to Global Config mode.
interface port-channel *number*	Enter interface configuration mode for the specified LAG. A range of LAGs to configure can be specified using the **interface range port-channel** command. For example, **interface range port-channel 1-3,10** configures LAGs 1, 2, 3, and 10.

Command	Purpose
hashing-mode *mode*	Set the hashing algorithm on the LAG. The *mode* value is a number from 1 to 7. The numbers correspond to the following algorithms: • 1 — Source MAC, VLAN, EtherType, source module, and port ID • 2 — Destination MAC, VLAN, EtherType, source module, and port ID • 3 — Source IP and source TCP/UDP port • 4 — Destination IP and destination TCP/UDP port • 5 — Source/destination MAC, VLAN, EtherType, and source MODID/port • 6 — Source/destination IP and source/destination TCP/UDP port • 7 — Enhanced hashing mode
CTRL + Z	Exit to Privileged EXEC mode.
show interfaces port-channel [*port-channel number*]	View LAG information for the specified LAG or for all LAGs.
show statistics port-channel *port-channel-number*	View interface statistics for the specified LAG.

Configuring LACP Parameters

Beginning in Privileged EXEC mode, use the following commands to configure system and per-port LACP parameters.

Command	Purpose
configure	Enter global configuration mode.
lacp system-priority *value*	Set the Link Aggregation Control Protocol priority for the switch. the priority value range is 1–65535.

Command	Purpose
interface port-channel *number*	Enter interface configuration mode for the specified LAG. A range of LAGs to configure can be specified using the **interface range port-channel** command. For example, **interface range port-channel 1-3,10** configures LAGs 1, 2, 3, and 10.
lacp port-priority *value*	Set the Link Aggregation Control Protocol priority for the port or range of ports. The priority value range is 1–65535.
lacp timeout {long | short}	Specify whether to wait a long or short time between LACP PDU transmissions.
exit	Exit to Privileged EXEC mode.
show lacp *interface*	View LACP parameters for an Ethernet interface or a LAG. The *interface* parameter includes the interface type (**gigabitethernet**, **tengigabitethernet**, or **port-channel**) and number.

Link Aggregation Configuration Examples

This section contains the following examples:

- Configuring Dynamic LAGs
- Configuring Static LAGs

NOTE: The examples in this section show the configuration of only one switch. Because LAGs involve physical links between two switches, the LAG settings and member ports must be configured on both switches.

Configuring Dynamic LAGs

The commands in this example show how to configure a static LAG on a switch. The LAG number is 1, and the member ports are 1, 2, 3, 6, and 7.

To configure the switch:

1 Enter interface configuration mode for the ports that are to be configured as LAG members.

```
console(config)#interface range te1/0/1-3,te1/0/6-7
```

2 Add the ports to LAG 2 with LACP.

```
console(config-if)#channel-group 1 mode active
```

3 View information about LAG 1.

```
console#show interfaces po1

Channel  Ports         Ch-Type  Hash Type  Min-links Local Prf
-------  ------------- -------  ---------  --------- ---------
Po1      Active:       Dynamic  7          1         Disabled
         Te1/0/1
         Inactive:
         Te1/0/2,
         Te1/0/3,
         Te1/0/6,
         Te1/0/7
Hash Algorithm Type
1 - Source MAC, VLAN, EtherType, source module and port Id
2 - Destination MAC, VLAN, EtherType, source module and port Id
3 - Source IP and source TCP/UDP port
4 - Destination IP and destination TCP/UDP port
5 - Source/Destination MAC, VLAN, EtherType, source MODID/port
6 - Source/Destination IP and source/destination TCP/UDP port
7 - Enhanced hashing mode
```

Configuring Static LAGs

The commands in this example show how to configure a static LAG on a switch. The LAG number is 2, and the member ports are 10, 11, 14, and 17.

To configure the switch:

1 Enter interface configuration mode for the ports that are to be configured as LAG members.

```
console(config)#interface range te1/0/10-12, te1/0/14,te1/0/17
```

2 Add the ports to LAG 2 without LACP.

```
console(config-if)#channel-group 2 mode on
```

3 View information about LAG 2.

```
console#show interfaces po2

Channel  Ports         Ch-Type  Hash Type Min-links Local Prf
-------  ------------- -------  --------- --------- ---------
Po1      Active:       Static   7         1         Disabled
         Te1/0/1
         Inactive:
         Te1/0/2,
         Te1/0/3,
         Te1/0/6,
         Te1/0/7
Hash Algorithm Type
1 - Source MAC, VLAN, EtherType, source module and port Id
2 - Destination MAC, VLAN, EtherType, source module and port Id
3 - Source IP and source TCP/UDP port
4 - Destination IP and destination TCP/UDP port
5 - Source/Destination MAC, VLAN, EtherType, source MODID/port
6 - Source/Destination IP and source/destination TCP/UDP port
7 - Enhanced hashing mode
```

Multi-Switch LAG (MLAG)

NOTE: This feature is not available on the Dell Networking N1500 Series switches.

Overview

In a typical layer-2 network, the Spanning Tree Protocol (STP) is deployed to avoid packet storms due to loops in the network. To perform this function, STP sets ports into either a forwarding state or a blocking state. Ports in the blocking state do not carry traffic. In the case of a topology change, STP re-converges to a new loop-free network and updates the port states. STP is relatively successful mitigating packet storms in the network, but redundant links in the network are blocked from carrying traffic by the spanning tree protocol.

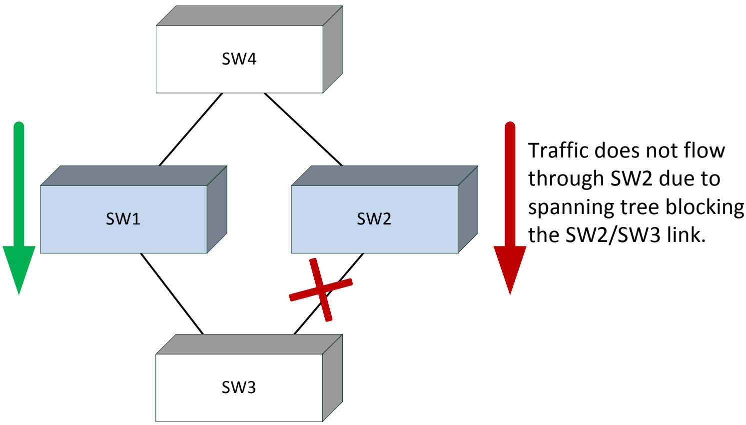
In some network deployments, redundant links between two switches are bundled together in a Link Aggregation Group (LAG) and appear as a single link in the spanning tree topology. The advantage is that all LAG member links can be in the forwarding state and a link failure can be recovered in milliseconds. This allows the bandwidth on the redundant links to be utilized. However, LAGs are limited to connecting multiple links between two partner switches, which leaves the switch as a single point of failure in the topology.

Dell Networking MLAG extends the LAG bandwidth advantage across multiple Dell Networking N-Series switches connected to a LAG partner device. The LAG partner device is unaware that it is connected over a LAG to two peer Dell Networking N-Series switches; instead, the two switches appear as a single switch with a single MAC address to the partner. All links can carry data traffic across a physically diverse topology and in the case of a link or switch failure, traffic can continue to flow with minimal disruption.

Deployment Scenarios

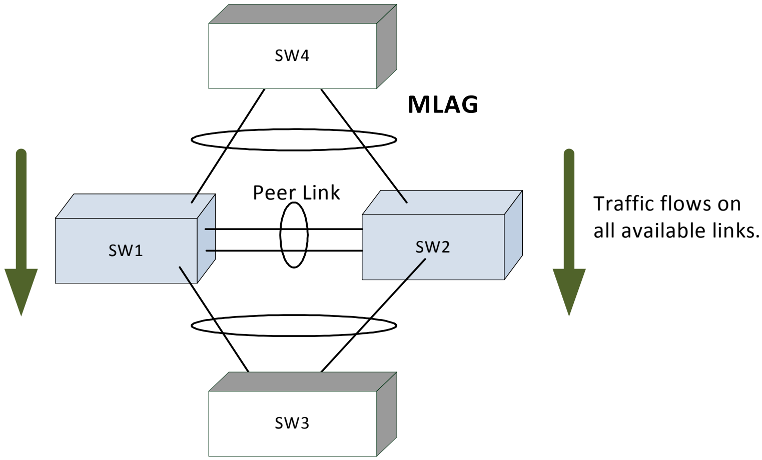
MLAG is intended to support higher bandwidth utilization in scenarios where a redundant layer-2 network is desired. In such scenarios the effects of STP on link utilization are profound. Large percentages of links do not carry data because they are blocked and only a single path through the network carries traffic.

Figure 28-8. STP Blocking



MLAG reduces some of the bandwidth shortcomings of STP in a layer-2 network. It provides a reduced convergence period when a port-channel link goes down and provides more bandwidth because all links can forward traffic. In the figure below, if SW1 and SW2 form an MLAG with SW3 and SW4, none of the links are blocked, which means traffic can flow over both links from SW4 through to SW1 and SW2 over both links from SW1 and SW2 to SW3.

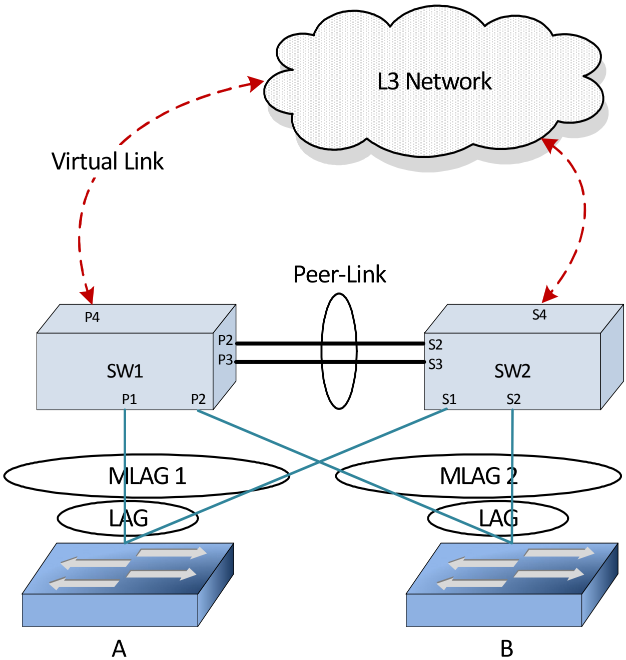
Figure 28-9. MLAG in a Layer-2 Network



Definitions

Refer to Figure 28-10 for the definitions that follow.

Figure 28-10. MLAG Components



MLAG switches: MLAG aware switches running Dell Networking Series switch firmware. No more than two MLAG aware switches can pair to form one end of the LAG. Stacked switches do not support MLAGs. In the above figure, SW1 and SW2 are MLAG peer switches. These two switches form a single logical end point for the MLAG from the perspective of switch A.

MLAG interfaces: MLAG functionality is a property of port-channels. Port-channels configured as MLAGs are called MLAG interfaces. Administrators can configure multiple instances of MLAG interfaces on the peer MLAG switches. Port-channel limitations and capabilities like min-links and maximum number of ports supported per LAG also apply to MLAG interfaces.

MLAG member ports: Ports on the peer MLAG switches that are part of the MLAG interface (P1 on SW1 and S1 on SW2).

Non-redundant ports: Ports on either of the peer switches that are not part of the MLAG (ports P4 and S4). MLAG interfaces and non-redundant ports cannot be members of the same VLAN, i.e. a VLAN may contain MLAG interfaces or a VLAN may contain non-redundant ports, but not both. To attach a host or switch to a non-redundant port, configure the port to be a member of the non-MLAG VLANs. This port is not part of the MLAG and is not considered an MLAG partner. Packets on non-MLAG VLANs are never passed over the peer link.

MLAG peer-link: A link between the two MLAG peer switches (ports P2,P3,S2,S3). Only one peer-link can be configured per device. The peer-link is crucial for the operation of the MLAG component. A port-channel must be configured as the peer-link. All VLANs configured on MLAG interfaces must be configured on the peer-link as well.

MLAG Dual Control Plane Detection link: A virtual link that is used to advertise the Dual Control Plane Detection Protocol (DCPDP) packets between the two MLAG switches (ports P4, S4). DCPDP is optional but should be used with caution. The protocol is used as a secondary means of detecting the presence of the peer switch in the network. The DCPDP protocol must not be configured on MLAG interfaces.

Configuration Consistency

The administrator must ensure that the neighboring devices connected to MLAG switches perceive the two switches as a single spanning tree and Link Aggregation Control Protocol (LACP) entity. To achieve this end, the following configuration settings must be identical for MLAG links on the MLAG peer switches:

1. Link aggregation
   - Hashing mode
   - Minimum links
   - Static/dynamic LAG
   - LACP parameters
     - Actor parameters
     - Admin key

- Collector max-delay
- Partner parameters

2 STP

The default STP mode for Dell Networking N-Series switches is RSTP. VLANs cannot be configured to contain both MLAG ports and non-MLAG (non-redundant) ports. RSTP, MSTP, and STP-PV/RSTP-PV are supported with MLAG. The following STP configuration parameters must be the identical on both MLAG peers.

– Spanning-tree version (RSTP, MSTP, or RSTP-PV)

– Bpdufilter

– Bpduflood

– Auto-edge

– TCN-guard

– Cost

– Edgeport

– STP Version

– MSTP or RSTP-PV VLAN configuration

– MST instance configuration (MST instance ID/port priority/port cost/mode) if MSTP is configured

– Root guard

– Loop guard

3 Port-channel interface

The following port-channel attributes must be identical for MLAG port-channels:

– Port-channel mode

– Link speed

– Duplex mode

– MTU

– Bandwidth

– VLAN configuration

The administrator should also ensure that the following are identical before enabling MLAG:

– FDB entry aging timers

– Static MAC entries.

– ACL configuration

4 Interface Configuration

– PFC configuration

– CoS queue assignments

5 VLAN configuration in an L2 topology

– MLAG VLANs must span the MLAG topology and be configured on both MLAG peers. This means that every MLAG VLAN must connect to two partner LAGs.

– VLAN termination of an MLAG VLAN on an MLAG peer is not supported.

6 Switch firmware versions

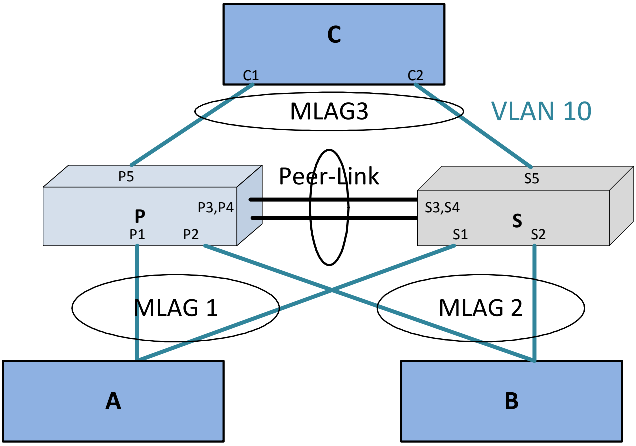
Except during firmware upgrade, the peer switch firmware versions must be identical, as subtle differences between versions may cause instability.

The administrator must ensure that the above configuration items are configured identically on the MLAG interfaces on both of the MLAG peers before enabling the MLAG feature. If the configuration settings are not in sync, the MLAG behavior is undefined. Once the above configuration is in place and consistent, the two switches will form an MLAG that operates in the desired manner. The MLAG may form even if the configuration is not consistent, however, it may not operate consistently in all situations.

Operation in the Network

Below is a sample MLAG topology and discussion:

Figure 28-11. Example MLAG Topology



In Figure 28-11:

1. VLAN 10 spans the MLAG network.
2. P and S are MLAG-aware peer devices. P stands for primary and S stands for secondary. The roles are elected after the DUTs exchange keep-alive messages. The two devices are connected with a peer-link {P3/P4–S3/S4}. Ports P1, S1 are members of MLAG1 and ports P2, S2 are members of MLAG2.
3. A port-channel must be configured as the peer-link. In Figure 28-11, P3, P4 and S3, S4 are the port-channel ports that form the peer-link.
4. MLAG devices select the roles based on keep-alive messages that run over the peer-link.
5. A, B, and C are MLAG-unaware devices.
6. A, B, and C are partner devices that form an MLAG with P and S. On A, B, and C, the aggregation is a regular LAG.
7. MLAG links are shown in blue.

Supported topologies and the way traffic is handled in these topologies is explained in the following sections.

The MLAG component uses the keep-alive protocol to select a primary and a secondary device. The primary switch owns the MLAG member ports on the secondary device. It handles the control plane functionality of supported protocols for the MLAG member ports on the secondary.

Peer-Link

The peer-link is a crucial for MLAG operation. The peer-link must be configured on a port-channel interface. Only one peer-link aggregation group is allowed per peer switch and this peer-link is shared by all instances of MLAG running on the two peer switches.

The peer-link is used for the following purposes:

- To transport keep-alive messages to the peer.
- To sync FDB entries learned on MLAG interfaces between the two MLAG peer switches.
- To forward STP BPDUs and LACPDUs received on secondary MLAG member ports to the primary MLAG switch.
- To send interface events related to MLAG interface and member ports that occur on the secondary switch to the primary switch.
- To transfer MLAG control information between the primary and secondary MLAG switches.
- To support a redundant forwarding plane in the case that all member ports of an MLAG interface are down on an MLAG peer. In this case, traffic received on the peer switch destined to the MLAG peer with the downed ports is sent over the peer-link to the peer MLAG switch for forwarding to the partner switch.

The peer-link is not utilized for partner traffic unless all LAG links connected to an MLAG partner on a single MLAG peer are disrupted. It is strongly recommended that the MLAG peer LAG consist of multiple physical links with sufficient bandwidth to carry all traffic expected to be carried by either of the MLAG peers.

The MLAG component internally configures filters so that traffic ingressing a peer-link is blocked from egress on the peer MLAG switch. The filters are modified when there is a failure of all the MLAG member interfaces on an

MLAG switch and traffic must egress through selected ports on the MLAG peer. These filters block incoming traffic on all VLANs configured on the peer link, not just those configured as part of an MLAG. Therefore, there is no connectivity between non-redundant ports across the peer-link.

Control Plane Election in MLAG Switches

The MLAG component uses the keep-alive protocol running on the peer link to select a primary and a secondary switch. The keep-alive protocol is mandatory. The selection of the primary switch is non-preemptive and is not configurable.

Once elected, the primary switch owns the MLAG member ports on the secondary device. It handles the control plane functionality of supported protocols for the MLAG member ports on the secondary switch. Protocol status is not sent from the primary to the secondary switch. Always use the management interface on the primary switch to examine MLAG status.

Peer-Link Keep-alive Protocol

MLAG peers exchange keep-alive packets over the peer-link. The keep-alive protocol is layer-2-based. Keep-alive messages are used for electing roles and to inform the MLAG peer that the MLAG switch is alive and functioning properly. The keep-alive protocol sends messages with an Ether-type of 0x88E8 addressed to destination MAC 01:00:B5:00:00:00.

Dual Control Plane Detection Protocol

The MLAG component may optionally run the Dual Control Plane Detection Protocol (DCPDP) to detect the presence of the peer switch independently of the keep-alive protocol running on the peer link.

The Dual Control Plane Detection Protocol is a UDP-based layer-3 protocol. DCPDP may be configured on a routed VLAN that does not contain any MLAG port-channel interfaces. When enabled, the DCPDP sends an layer-3 control plane detection message to the peer once every second. The message is unidirectional and contains the senders MAC address in the payload. The state of the primary and secondary MLAG switches is maintained on both MLAG peers.

DCPDP runs over an IP interface when enabled.

DCPDP and Peer Link Failures

DCPDP is intended to provide a secondary layer of protection against peer link failures. If the peer-link goes down while the DCPDP protocol is enabled and remains up, the MLAG links on the MLAG secondary peer are disabled. The primary switch continues to forward traffic and, if LACP is enabled, send LACPDUs using the system MAC of the MLAG. Spanning tree reconvergence on the partner devices is avoided.

In the case where there are no keep-alive messages detected from the peer and DCPDP is disabled, but both peer units remain up, two primary switches result, each with the MLAG system MAC address, and each operating over its part of the former MLAG. In this situation, the selection of dynamic or static LAGs determines the MLAG behavior.

On a peer-link failure with DCPDP disabled and the MLAG configured with dynamic LAGs to the partners, traffic forwarding continues through the primary MLAG switch. The secondary switch brings down the MLAG interfaces and brings them up with a new (different from the primary) MLAG system ID. LACP running on the partner device detects that the links in the port-channel connected to the secondary MLAG switch are sending LACPDUs with a different system ID and does not bring up the links connected to secondary MLAG peer. This behavior reduces or eliminates spanning tree reconvergence due to the MLAG switches sending BPDUs with different bridge IDs to the partner switch.

On a peer-link failure with DCPDP disabled and the MLAG configured with static LAGs to the partners, traffic forwarding continues through both the primary and secondary MLAG switches. Spanning tree sends BPDUs with different bridge IDs to the partner switch, resulting in serial spanning tree reconvergence. For this reason, dynamic MLAGs are strongly recommended.

MLAG and Redundant Peer Links

A redundant peer link is a link other than the peer link between the MLAG peer switches. Typically, these links cause a loop in the network and may cause the peer link to become blocked by spanning-tree. Dell Networking can support traffic flow over redundant links with some additional configuration. Multiple spanning tree must be configured, the redundant link must be assigned a VLAN that is NOT in the MLAG domain (but can be configured on the peer link), and the VLAN assigned to the redundant link must be configured in a unique MST instance not shared with the MLAG domain. If the VLAN assigned to the redundant link is also configured on the peer link, traffic on that VLAN is blocked by MLAG.

To configure the redundant link to be the forwarding for the redundant MST instance, the link cost needs to be reduced in order to be the root port. However, with MLAG, even changing the cost of the redundant port will not make the desired port as root port, as the MLAG peer link cost is internally set to 0 for all instances. Also it is not permitted by spanning-tree to set a link cost to 0. The lowest possible value allowed is 1. The spanning tree cost of the peer link is set to 0 with a specific intent—to ensure that the peer link is always elected as the designated /root forwarding port in the case of any redundant links between the primary and secondary switches. It is not possible to configure a non-peer-link/redundant port cost lower than the peer-link cost. The alternative is to configure the priority of the bridge (which is currently a non-root bridge) for the desired MST instance and increase the internal cost of the peer link interface for that same instance on the bridge (which is originally the root bridge) so that the port becomes designated forwarding instead of root forwarding. A example configuration is given later in this chapter.

Layer-2 Configuration Steps

This section describes how to configure two MLAG peers in a basic layer-2 switching configuration with the default spanning-tree configuration.

1 Enable MLAG globally and create the MLAG VLANs:

```
console#config
console(config)#feature vpc
console(config)#vlan 10-17
console(config-vlan10-17)#exit
```

2 Configure the keep-alive protocol:

a Configure the device priority using the **role** command, if desired. This should be configured differently for each of the MLAG peers.

b Configure the VPC domain ID and also configure the local priority, if desired. The VPC domain ID is not important except that it MUST be the same as the MLAG peer and SHOULD be different than any other MLAG partner. The MLAG system MAC address may optionally be configured in the step.

```
console(config)#vpc domain 1
```

```
console(config-vpc 1)#role 10
console(config-vpc 1)#exit
```

Modifications to priority and timeout interval are effective only before the keep-alive protocol is enabled. Once enabled, MLAG switches contest in an election to select the primary and secondary switch. The election is non-preemptive.

If configured, the system virtual MAC address MUST be the same on both of the MLAG peers.

3 Configure the peer-link. On each MLAG peer:

- Configure a port-channel as the peer-link for the MLAG devices. It is recommended that the administrator use dynamic LAGs as port-channels.
- It is strongly recommended that the MLAG peer LAG consist of multiple physical links with sufficient bandwidth to carry all MLAG traffic expected to be carried by either MLAG peer.
- Enable trunking on the peer-link. Remove any non-MLAG VLANs from the peer-link trunk port. VLANs cannot be configured to contain both MLAG ports and non-MLAG (non-redundant) ports.
- Ensure that the peer-link has a native VLAN configured.
- Optionally, configure UDLD on the peer-link to detect and shut down unidirectional links. UDLD should be used on any fiber ports, as fiber ports can operate in a unidirectional mode.
- Associate the port-channel with physical links.

```
console(config)#interface port-channel 1
console(config-if-Po1)#description "MLAG-Peer-Link"
console(config-if-Po1)#switchport mode trunk
console(config-if-Po1)#vpc peer-link
console(config-if-Po1)#exit
console(config)#interface range te1/0/1-2
console(config-if)#channel-group 1 mode active
console(config-if)#description "MLAG-Peer-Link"
console(config-if)#udld enable
console(config-if)#udld port aggressive
console(config-if)#exit
```

When the peer-link is configured, the MLAG component disables MAC learning on the port-channel configured as the peer-link.

4 Configure DCPDP (optional):

   a Configure a VLAN routing interface and assign a local IP address (different from the peer address).

   b Configure the peer-switch IP address (the destination IP address)

   c If needed, configure the UDP port number to send and receive the protocol messages.

   d Configure the source IP address

   e Enable the protocol. The protocol starts running if MLAG is globally enabled.

```
console(config)#vlan 100
console(config-vlan100)#exit
console(config)#interface vlan 100
console(config-if-vlan100)#ip address 192.168.0.2
255.255.255.0
console(config-if-vlan100)#exit
console(config)#vpc domain 1
console(config-vpc 1)#peer-keepalive destination
192.168.0.1 source 192.168.0.2
console(config-vpc 1)#peer detection enable
console(config-vpc 1)#exit
```

5 Configure the MLAG partner interfaces:

Configure a port-channel as an MLAG interface and assign to a VPC. Each MLAG must have a unique VPC ID and the VPC configuration must be identical on both switches. The port-channels on the MLAG peer must be assigned to the same VPC ID. However, the member ports for the port-channel may be different. The administrator must ensure that the port-channel configuration on both the switches is in sync before enabling MLAG. After the MLAG interfaces are enabled, the MLAG interfaces are operationally disabled for a brief period while the MLAG component exchanges information regarding the port members that constitute the port-channel on each device. Once this information is populated on both devices, the MLAG interfaces are operationally enabled and traffic forwarding on MLAG interfaces is allowed.

The secondary switch forwards all BPDUs/LACPDUs received on the port members of the MLAG interface to the primary switch over the peer-link. Events related to MLAG interface and their port members are forwarded to the primary switch for handling. FDB entries learned on MLAG interfaces are synced between the two devices.

```
console(config)#interface range gi1/0/1-4
console(config-if)#channel-group 2 mode active
console(config-if)#exit
console(config)#interface range gi1/0/5-8
console(config-if)#channel-group 3 mode active
console(config-if)#exit
console(config)#interface port-channel 2
console(config-if-Po2)#switchport mode trunk
console(config-if-Po2)#vpc 1
console(config-if-Po2)#exit
console(config)#interface port-channel 3
console(config-if-Po3)#switchport mode trunk
console(config-if-Po3)#vpc 2
console(config-if-Po3)#exit
```

Switch Firmware Upgrade Procedure

MLAG supports minimally intrusive firmware upgrade of the MLAG peer switches. In most cases, protocols with retransmission capability, e.g., TCP, will experience a limited interruption of service. Network operators must ensure that the aggregate bandwidth in use on the MLAG can be supported on a single MLAG peer.

Use the **show vpc brief** command to determine which switch is the primary switch. This procedure upgrades the standby switch first, followed by the primary switch. Following this order reduces the reconvergence time to the minimum.

Upgrade Steps

Copy the new firmware to both the primary and secondary switches and activate it. Disable DCPDP if enabled:

1 On the secondary switch, disable DCPDP if enabled and save the configuration.

2 On the primary switch, disable DCPDP if enabled and save the configuration.

Upgrade the MLAG secondary switch:

1 On the MLAG secondary switch, shut down the MLAG-enabled physical links (not the port-channel). Do not save the running-config.

2 On the MLAG secondary switch, shut down the MLAG peer-link.
3 Reload the secondary switch.
4 Re-enable the peer-link, if disabled, and ensure that it is up. Re-enable the MLAG-associated physical ports.
5 Wait until traffic is re-established on the standby switch.

Repeat the upgrade procedure on the MLAG primary peer:

1 On MLAG primary switch, shut down the MLAG enabled physical links.
2 On MLAG primary switch, shut down the MLAG peer-link.
3 Reload the primary switch.
4 Re-enable the peer link, if disabled, and ensure that it is up. Re-enable the MLAG-associated physical ports.
5 Verify that traffic is re-established on the primary switch after the reconvergence.

At this point, the switch firmware is upgraded and the MLAG is fully functional.

Static Routing on MLAG Interfaces

MLAG interfaces can be enabled as layer-3 VLANs; that is, they can be assigned IP addresses. There is no support for routing protocols such as OSPF, RIP, etc. on MLAG interfaces. VRRP can be configured on these routing interfaces to provide Virtual IP/Virtual MAC redundancy. Routing is supported only on the edge of the MLAG towards the partner network, in support of implementing a subnet per VLAN towards which the partner network can route. The interior MLAG VLANs, and especially the MLAG peer links, must be configured for switching and must span the MLAG topology.

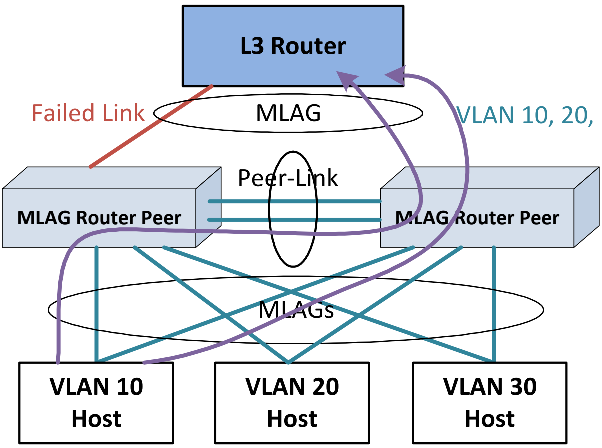
MLAGs and Routing

MLAG is supported as a replacement for spanning tree in layer-2 switched network topologies. When connecting to routed networks, the links/VLANs on the router must be part of the MLAG domain, and the links/VLANs leading to the rest of the layer-2 network or to layer-3 hosts must be part of the MLAG domain for the MLAG feature to automatically utilize the peer-link to forward packets around failures. MLAG VLANs may have IP addresses assigned, but MLAG VLANs cannot be used to route across MLAG or non-redundant VLANs, as the MLAG feature does not correlate failures in one VLAN with another VLAN to unblock packets crossing the MLAG peer-link.

Recommended Layer-3 Connectivity

The topology shown in Figure 28-12 uses the MLAG switches as layer-2 switches. All VLANs traverse the MLAG topology from the top switches/routers to the bottom switches/routers. The LAGs for each VLAN host are in a separate VPC. The router sees the port-channel as a single logical interface with multiple VLANs. This topology is highly recommended as it utilizes MLAG in the scenario for which it was intended (redundant full-bandwidth replacement for spanning tree) and allows the MLAG peers to detect failures and unblock the appropriate VLANs on the peer link so that traffic flow can continue unimpeded.

Figure 28-12. Recommended Layer-3 Connectivity

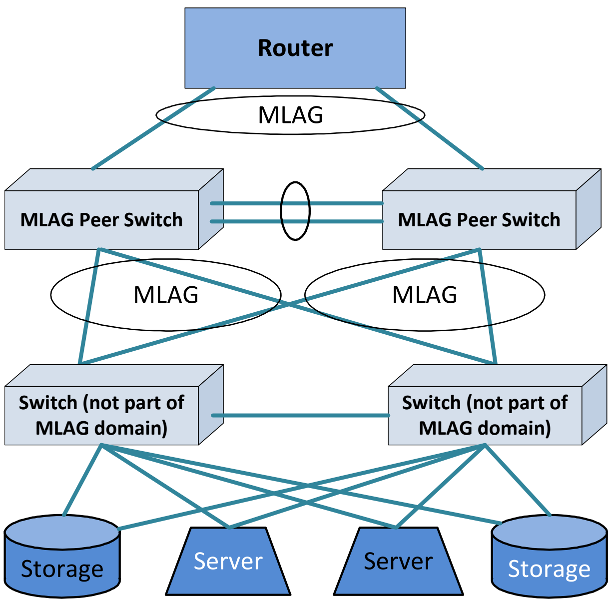


Alternative Recommended Layer-3 Connectivity

The loop-free topology shown in Figure 28-13 uses the MLAG switches as layer-2 switches in an EOR role. The single VLAN traverses the MLAG topology from the top router to the bottom storage and servers. Multiple VLANs in different VPCs may be used to isolate clusters of storage/servers from each other. This topology is highly recommended, as it utilizes MLAG in the scenario for which it was intended (redundant full-bandwidth replacement for spanning tree in a fully layer-2 topology) and allows the MLAG peers to detect failures and unblock the appropriate VLANs on the peer link so that traffic flow can continue unimpeded.

The lower pair of switches connects clusters of storage and servers in a TOR role in support of devices that do not support link aggregation. Switching between the storage and the servers within the rack proceeds in the normal manner and remains isolated on the lower switch pair that is not part of the MLAG domain. Traffic entering or exiting the rack proceeds over the EOR MLAG.

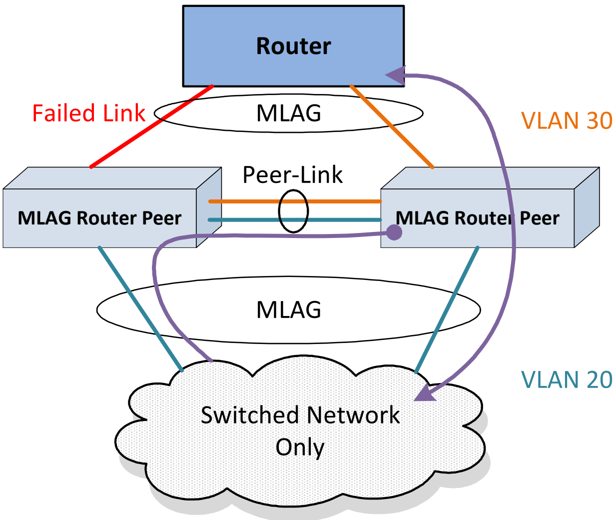
Figure 28-13. Alternative Recommended Layer-3 Connectivity



Layer-3 VLAN Termination on MLAG Not Supported

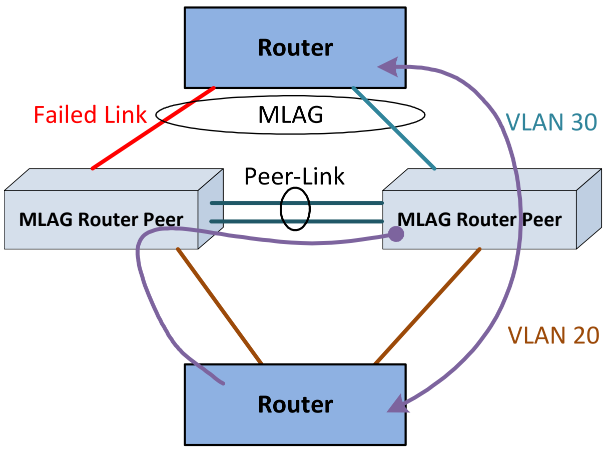
In the “two-armed” fully routed scenario shown in Figure 28-14, both the routed network and the switched network are in the MLAG. Switched traffic to and from the upstream network is automatically unblocked over the peer-link when an MLAG link fails. But, because the failed link is part of a layer-3 port-channel (with a unique VLAN), the peer-link does not automatically unblock the downstream routed VLANs (which are not correlated with the upstream MLAG VLANs) across the peer link. Specifically, MLAG does not correlate the failure in VLAN 30 with VLAN 20. This leads to a black hole. Adding a backup routed link solves the black hole issue, but it also makes the MLAG solution unnecessary. Layer-3-routed VLAN termination on the MLAG peers is not supported—VLANs must extend across the MLAG peers to two MLAG partners.

Figure 28-14. Layer-3 VLAN Termination on MLAG, Example 1



In the scenario shown in Figure 28-15(similar to the previous scenario), the downstream router is not configured with port-channel and uses ECMP or some other load sharing scheme to send packets to routed MLAG peers. MLAG cannot react appropriately to a link failure on the upstream router because the VLANs are routed across the MLAG peers. MLAG cannot logically connect the failure on VLAN 30 with non-redundant VLAN 20. Consequently, MLAG does not unblock VLAN 20 from traversing the peer link. The downstream router continues to send packets on VLAN 20 to the MLAG peer with the failed link. But because routed VLAN 20 is not part of the MLAG, packets remain blocked when transiting the MLAG peer link. Layer-3 routed VLANs termination on the MLAG peers is not supported—VLANs must extend across the MLAG peers.

Figure 28-15. Layer-3 VLAN Termination on MLAG, Example 2

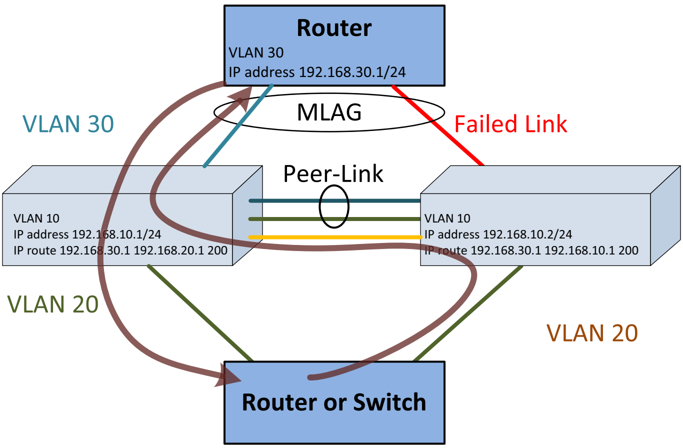


Degenerate Routing Topology

In a “one-armed” topology, the MLAG is partnered with a single router or switch. The router is configured with a LAG toward the MLAG peer switches and has an IP address configured on the router LAG. The peer switches may be configured with VRRP and have IP addresses assigned to both the routed VLANs. If a multi-tier MLAG topology is used below the MLAG peers, these switches must not have layer-3 port-channels configured as part of the MLAG. An additional backup routed link between the MLAG peers is required to handle the case where a link from the router to one of the MLAG peers fails. Static routes must be added to the primary and secondary MLAG peers to route traffic addressed to the connected router across the backup routed link in the case of a failure of an MLAG link to the router.

This is not a recommended topology, as the same scenario can be achieved without the use of MLAG by simply configuring the middle switches as routers and using ECMP to load-balance across the links to the redundant router pair. In this type of solution, MLAG adds no value, as the redundancy is provided by layer-3 routing, not by the MLAG.

Figure 28-16. Degenerate Routing Topology



In the one-armed scenario in Figure 28-16, the MLAG cannot associate the failure of the VLAN 30 link with VLAN 20. Traffic from the routed or switched network towards the upstream router is routed over the backup router link when the MLAG link fails solely based on the routing configuration. Traffic from the upstream router on VLAN 30 to the switched/routed network is handled by the MLAG failover scenario and is switched across the peer-link on VLAN 30, but it could just as easily be handled by layer-3 routing.

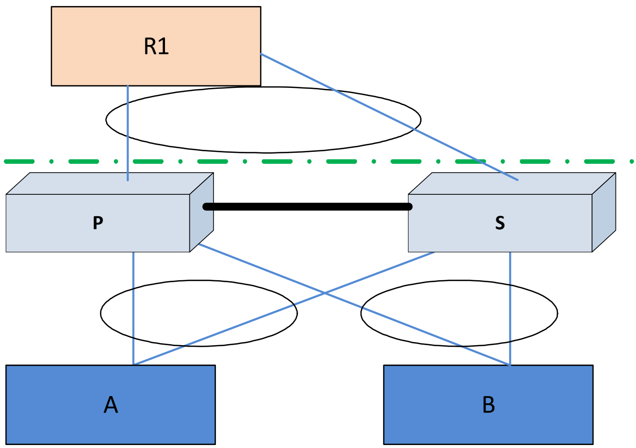
Virtual Router Redundancy Protocol

If VRRP is enabled on a VLAN that has an MLAG port as its member, both VRRP routers become VRRP masters operationally in the VLAN. This is to allow load balancing of the northbound layer-3 traffic on the MLAG.

Since the peer-link is a member of the same routing VLANs as all MLAGs, both the primary and secondary MLAG routers see VRRP advertisements sent by the other router. The internal switch packet filters are modified to drop VRRP advertisement packets if they are received on a VLAN that has an MLAG as its member. This ensures that both MLAG devices become VRRP Masters.

Consider the scenario in Figure 28-17.

Figure 28-17. MLAG with VRRP



When layer-3 data is received by the MLAG primary switch destined to A, P would trigger an ARP request to learn A's MAC address. In this case, the ARP request originated by P would have its interface MAC address as the source MAC address (MAC-P, for example) and interface IP address as the source IP address. If the ARP reply from A, with destination MAC address as MAC-P, is hashed to S, the packet will be forwarded by S to P over the peer-link so that P can learn A's MAC address. **This requires that the virtual IP address of the VRRP "routers" be different from the physical IP of either of the MLAG peers.** This is to ensure that packets generated by either of the MLAG peers is transmitted with the source MAC address as the physical MAC address and not the virtual MAC address. In the example in Figure 28-17, if the virtual MAC address is used as the source MAC address in the ARP from P to A, then S will consume the packet, as it is operationally a VRRP master too. The packet is forwarded to P if the physical MAC address is used.

Note that the VLANs connecting A and B to the MLAG peers are extended to R1. P and S do not actually route packets. Within the MLAG domain, packets are switched only.

Caveats and Limitations

Traffic to and from non-redundant ports is filtered and never crosses the MLAG peer-link, and such ports/VLANs need to obtain connectivity via an alternative other than the MLAG-connected ports/VLANs.

To achieve east-west traffic, the administrator can connect an additional interface between the primary and secondary MLAG switches with spanning tree disabled on that link. The administrator must configure the interface to be a member of the required non-MLAG VLANs. The administrator should ensure that there are no loops with this connection.

The MLAG peer link does not become operational until both MLAG peers establish communication, and are configured with the same domain identifier and the same virtual MLAG MAC address. MLAG partner links remain disabled until the MLAG peers establish connectivity. Shutting down an MLAG peer and rebooting the other peer will result in the MLAG partner links remaining disabled until the MLAG peer that was shut down is re-enabled.

Port-channel numbers within a VPC must be identical on the MLAG peers. The associated physical interfaces in the port-channels need not be the same.

It is a requirement of a port-channel that the link members operate at the same duplex and speed settings. In addition, copper ports have larger latencies than fiber ports. If fiber and copper ports are aggregated together, packets sent over the fiber ports would arrive significantly sooner at the destination than packets sent over the copper ports. This can cause significant issues in the receiving device, as it would be required to buffer a potentially large number of frames. Devices unable to buffer the requisite number of frames will show excessive frame discard.

Routing is not supported across multiple MLAGs (i.e., in two-tier topology). This is a fundamental limitation of MLAG, which is intended as a replacement for other, less efficient layer-2 topologies. Should a multi-tier layer-3 topology be desired, other well established and well understood techniques, such as ECMP and redundant router pairs, will allow a layer-3 routed network to utilize bandwidth efficiently. Layer-3 routing is capable of routing packets around failed links and failed routers.

Spanning tree (and LACP) PDUs are proxied from the secondary MLAG peer to the MLAG primary switch. This implies that at least two spanning tree roots will exist in the MLAG network: the root bridge for the MLAG member ports/VLANs on the primary switch and the root bridge for the non-redundant ports/VLANs that are not part of the MLAG.

The peer link requires a native VLAN to be configured. This is a limitation of the peer-link keep alive protocol.

On primary switch failover, the secondary switch flushes the FDB MAC addresses and uses the system virtual MAC address in spanning tree BPDUs and in the LACP actor ID. This avoids rebuilding the link aggregation group followed by spanning tree reconvergence.

MLAG-supported protocols are active only on the MLAG primary switch. The protocols are proxied from the secondary peer switch to the primary switch. The primary switch receives state information from the secondary peer switch and programs the secondary peer hardware. It does not send protocol state information to the secondary peer. This leads to a number of seemingly inconsistent behaviors if these facts are ignored:

- MLAG port-channel state is maintained on the primary peer only. The MLAG secondary peer has accurate state for the member links, but not for an MLAG port-channel. The operator can shut down a MLAG port-channel only from the primary MLAG peer.
- Shutting down a MLAG port-channel on the primary peer shuts down the port-channel on both the primary and secondary MLAG peers.
- Shutting down a MLAG port-channel on the secondary MLAG peer has no effect. The operator can shut down the individual links instead.
- The spanning tree status is only shown correctly on the primary MLAG peer for the redundant links and associated VLANs. The spanning-tree state on the secondary switch is accurate only for the non-redundant links and associated VLANs.

- On a failover from the primary MLAG peer to the secondary MLAG peer, the ports are made members of the secondary MLAG peer switch's spanning tree and spanning tree reconvergence may occur.. The forwarding database and ARP cache are flushed and relearned.
- MLAG (VPC) status only shows correctly on the primary MLAG peer and does not show correctly on the secondary MLAG peer. Status is not forwarded from the primary MLAG peer to the secondary MLAG peer.
- If it is desired to run a redundant link between the MLAG peers, Multiple Spanning Tree must be used, a separate VLAN must be configured for the redundant link which cannot be configured on the peer link, and the VLAN must be assigned to a unique MST instance not used by any MLAG VLAN.

The Dell Networking MLAG solution is not peer-compatible with other vendor's multichassis LAG solutions. Dell Networking N-Series switches configured for MLAG cannot peer with another vendor switch.

IGMP/MLD snooping is not supported on MLAG-enabled switches. Disable IGMP/MLD snooping before enabling MLAG.

MLAG interfaces and non-redundant ports cannot be members of the same VLAN; i.e., a VLAN may contain MLAG interfaces or a VLAN may contain non-redundant ports, but not both.

The Dell Networking MLAG solution supports MSTP, RSTP, and RSTP-PV spanning tree modes. Spanning tree may also be disabled on the MLAG peers, although this may lead to a network loop. If MSTP is configured, the MSTP domain MUST be named or the peer-link will remain blocked.

Only two switches are supported as MLAG peers. These switches may not be stacked with other switches.

The MLAG peer switches synchronize the state of spanning tree, the layer-2 forwarding cache, and other protocols to support reduced convergence times during MLAG link and MLAG switch failures. The synchronized state is only available on the MLAG primary switch. Table 28-1 indicates which switch features synchronize their state across the MLAG peers.

- An N/A entry indicates that state synchronization is not required (usually for a link local protocol) and the feature can be configured on an MLAG VLAN or MLAG-associated links. In some cases, it may be necessary to configure an N/A feature identically on the MLAG peer switches for it to work properly; e.g., port mirroring for an MLAG link must be configured on both MLAG peer switches to capture the conversation from the MLAG partner switch.

- A Yes entry indicates that the feature may be configured on an MLAG VLAN and will synchronize state across the MLAG peers. The configuration for features marked Yes must be identical on both switches. MLAG does not synchronize configuration with the MLAG peer.
- A No entry indicates that the switch feature does not synchronize state across the MLAG peers and the feature may not be configured on an MLAG VLAN.

Table 28-2. MLAG State Synchronization Per Feature

Components	MLAG State Synchronization Support
DOT1Q	Yes
Protocol Based VLANs/802.1v	No
GARP	No
GVRP	No
GMRP	No
DOT1P	No
Unauthenticated VLAN	No
Voice VLAN	No
Guest VLAN	No
MAC Authentication Bypass	No
Broadcast Storm Recovery	No
DOT3AD	Yes
LAG Hashing	Yes
Port Mirroring	N/A
MAC Filter	N/A
MFDB	No
IGMP/MLD Snooping	No
DOT1Qbb	No

Table 28-2. MLAG State Synchronization Per Feature (Continued)

Components	MLAG State Synchronization Support
DOT1S	Yes
Loop Guard	No
FDB	Yes
MACLOCK	No
DVLAN	No
DOT1AB	No
IP Subnet-based VLANs	N/A
MACVLAN	N/A
Protected Port	No
DHCP Snooping	No
IP Source Guard	No
Dynamic ARP Inspection	No
Auto-Negotiation	N/A
L2-Relay	No
MRP	No
MMRP	No
DOT1AS	No
802.1qav	No
DOT1AG	No
ACL	N/A
DiffServ	N/A
CoS	N/A
ACL Logging	N/A
Flow-based port mirroring	N/A
VOIP	No
iSCSI	No
DOT1AD	No

Table 28-2. MLAG State Synchronization Per Feature (Continued)

Components	MLAG State Synchronization Support
DOT3AH	No
DCBX	N/A
ETS	N/A
FIP Snooping	No
MVRP	No
Management ACL	No
UDLD	N/A
Private VLAN	No
LLPF	No
Port Aggregator	No
EAV	No
MSRP	No
MVR	No
Class-Based VLAN	No
DHCP Filtering	No
EASY_ACL	No
Media VLAN	No
PBVLAN	No
VLAN-Rate Limit	No
Flow Control	N/A
LLDP	N/A
Jumbo Frames	N/A

Basic Configuration Example

This example shows the configuration of the two MLAG peers and a single MLAG partner in the simplest possible configuration. No MLAG peer priorities are configured, nor is UDLD enabled on the peer-link. DCPDP is not enabled. The default spanning tree configuration is used and spanning-tree is disabled on the peer link. A system MAC address is assigned to both MLAG peers. The system virtual MAC address is used in the spanning-tree BPDUs and LACPDUs.

MLAG Peer A

Current Configuration:

- System Description "Dell Networking N3024F, 6.0.0.0, Linux 3.6.5-858bcf6e"
- System Software Version 6.0.0.0

```
console#configure
console(config)#vlan 10
console(config-vlan10)#exit
console(config)#hostname "MLAG-Peer-A"
MLAG-Peer-A(config)#slot 1/0 2    ! Dell Networking N3024F
MLAG-Peer-A(config)#stack
MLAG-Peer-A(config-stack)#member 1 2    ! N3024F
MLAG-Peer-A(config-stack)#exit

MLAG-Peer-A(config)#interface Gi1/0/23
MLAG-Peer-A(config-if-Gi1/0/23)#channel-group 2 mode active
MLAG-Peer-A(config-if-Gi1/0/23)#description "MLAG-Partner-Link"
MLAG-Peer-A(config-if-Gi1/0/23)#exit

MLAG-Peer-A(config)#interface Te1/0/1
MLAG-Peer-A(config-if-Te1/0/1)#channel-group 1 mode active
MLAG-Peer-A(config-if-Te1/0/1)#description "MLAG-Peer-Link"
MLAG-Peer-A(config-if-Te1/0/1)#exit

MLAG-Peer-A(config)#interface port-channel 1
MLAG-Peer-A(config-if-Po1)#description "MLAG-Peer-Link"
MLAG-Peer-A(config-if-Po1)#switchport mode trunk
MLAG-Peer-A(config-if-Po1)#vpc peer-link
MLAG-Peer-A(config-if-Po1)#exit

MLAG-Peer-A(config)#interface port-channel 2
MLAG-Peer-A(config-if-Po2)#switchport mode trunk
MLAG-Peer-A(config-if-Po2)#switchport trunk native vlan 10
```

```
MLAG-Peer-A(config-if-Po2)#vpc 1
MLAG-Peer-A(config-if-Po2)#exit

MLAG-Peer-A(config)#snmp-server engineid local
800002a203001ec9dec52b
MLAG-Peer-A(config)#snmp-server agent boot count 2
MLAG-Peer-A(config)#feature vpc
MLAG-Peer-A(config)#vpc domain 3
MLAG-Peer-A(config-vpc 3)#system-mac 0011.2233.4455
MLAG-Peer-A(config-vpc 3)#peer-keepalive enable
MLAG-Peer-A(config-vpc 3)#exit
MLAG-Peer-A(config)#exit
```

MLAG Peer B

Current Configuration:

- System Description “Dell Networking N3024F, 6.0.0.0, Linux 3.6.5-858bcf6e”
- System Software Version 6.0.0.0

```
console#configure
console(config)#vlan 10
console(config-vlan10)#exit
console(config)#hostname "MLAG-Peer-B"
MLAG-Peer-B(config)#slot 1/0 2    ! Dell Networking N3024F
MLAG-Peer-B(config-stack)#stack
MLAG-Peer-B(config-stack)#member 1 2    ! N3024F
MLAG-Peer-B(config-stack)#exit

MLAG-Peer-B(config)#interface Gi1/0/23
MLAG-Peer-B(config-if-Gi1/0/23)#channel-group 2 mode active
MLAG-Peer-B(config-if-Gi1/0/23)#description "MLAG-Partner-Link"
MLAG-Peer-B(config-if-Gi1/0/23)#exit

MLAG-Peer-B(config)#interface Te1/0/1
MLAG-Peer-B(config-if-Te1/0/1)#channel-group 1 mode active
MLAG-Peer-B(config-if-Te1/0/1)#description "MLAG-Peer-Link"
MLAG-Peer-B(config-if-Te1/0/1)#exit

MLAG-Peer-B(config)#interface port-channel 1
MLAG-Peer-B(config-if-Po1)#description "MLAG-Peer-Link"
MLAG-Peer-B(config-if-Po1)#switchport mode trunk
MLAG-Peer-B(config-if-Po1)#vpc peer-link
MLAG-Peer-B(config-if-Po1)#exit

MLAG-Peer-B(config)#interface port-channel 2
MLAG-Peer-B(config-if-Po2)#switchport mode trunk
MLAG-Peer-B(config-if-Po2)#switchport trunk native vlan 10
MLAG-Peer-B(config-if-Po2)#vpc 1
MLAG-Peer-B(config-if-Po2)#exit

MLAG-Peer-B(config)#snmp-server engineid local
800002a203001ec9dec513
MLAG-Peer-B(config)#snmp-server agent boot count 3
MLAG-Peer-B(config)#feature vpc
MLAG-Peer-B(config)#vpc domain 3
MLAG-Peer-B(config-vpc 3)#system-mac 0011.2233.4455
MLAG-Peer-B(config-vpc 3)#peer-keepalive enable
MLAG-Peer-B(config-vpc 3)#exit
```

```
MLAG-Peer-B(config)#exit
```

MLAG Partner

Current Configuration:

- System Description "Dell Networking N2048, 6.0.0.0, Linux 3.6.5-858bcf6e"
- System Software Version 6.0.0.0

```
console#configure
console(config)#hostname "LAG-SW"
LAG-SW(config)#slot 1/0 5    ! Dell Networking N2048
LAG-SW(config)#stack
LAG-SW(config-stack)#member 1 8    ! N2048
LAG-SW(config-stack)#exit

LAG-SW(config)#interface vlan 1
LAG-SW(config-if-vlan1)#ip address dhcp
LAG-SW(config-if-vlan1)#exit

LAG-SW(config)#interface Gi1/0/1
LAG-SW(config-if-Gi1/0/1)#channel-group 1 mode active
LAG-SW(config-if-Gi1/0/1)#exit

LAG-SW(config)#interface Gi1/0/2
LAG-SW(config-if-Gi1/0/2)#channel-group 1 mode active
LAG-SW(config-if-Gi1/0/2)#exit

LAG-SW(config)#interface port-channel 1
LAG-SW(config-if-Po1)#switchport mode trunk
LAG-SW(config-if-Po1)#exit

LAG-SW(config)#snmp-server engineid local 800002a203001ec9deb777
LAG-SW(config)#snmp-server agent boot count 3
LAG-SW(config)#exit
```

Status Reporting

The status outputs of the various VPC commands are self-explanatory. Both the configured and operational status is shown in the outputs. Additional commands are shown below that may be useful in troubleshooting MLAG configuration or operational issues. All of the commands below are run on the MLAG primary switch except as noted otherwise.

```
MLAG-Peer-A(config)#show vpc brief

VPC admin status............................... Enabled
Keep-alive admin status........................ Enabled
VPC operational status......................... Enabled
Self role...................................... Primary
Peer role...................................... Secondary
Peer detection admin status.................... Disabled

Peer-Link details
-----------------

Interface...................................... Po1
Peer-link admin status......................... Enabled
Peer-link STP admin status..................... Disabled
Configured VLANs............................... 1,10
Egress tagged VLANs............................ 10

VPC Details
-----------

Number of VPCs configured...................... 1
Number of VPCs operational..................... 1


VPC id# 1
-----------
Interface...................................... Po2
Configured VLANs............................... 1,10
VPC interface state............................ Active

Local Members        Status
-----------------    ------
Gi1/0/23             Up

Peer Members         Status
----------------     ------
Gi1/0/23             Up
```

```
LAG-SW(config)#show vpc role

Self
----

Keep-alive admin status........................ Disabled
Keep-alive operational status.................. Disabled
Priority....................................... 100
System MAC address............................. 001E.C9DE.B777
Time-out....................................... 5
VPC admin status............................... Disabled
VPC role....................................... None

Peer
----

Priority....................................... 0
VPC role....................................... None
System MAC address............................. 0000.0000.0000


LAG-SW(config)#show vpc peer-keepalive

Peer IP address................................ 0.0.0.0
Source IP address.............................. 0.0.0.0
UDP port....................................... 50000
Peer detection................................. Disabled
Peer detection operational status.............. Down
Peer is detected............................... False

MLAG-Peer-A(config)#show interfaces status po1

Port    Description
Channel
------- ------------------------------
Po1     MLAG-Peer-Link

Operational State.............................. Up
Admin Mode..................................... Enabled
Port Channel Flap Count........................ 1

Member    Device/       Port   Port   Flap
Ports     Timeout       Speed  Active Count
--------- ------------- ------ ------ ------
Te1/0/1   actor/long    10000  True   1
          partner/long
```

```
MLAG-Peer-A(config)#show interfaces status po2

Port    Description
Channel
------- ------------------------------
Po2

Operational State.............................. Up
Admin Mode..................................... Enabled
Port Channel Flap Count........................ 0

Member    Device/       Port   Port   Flap
Ports     Timeout       Speed  Active Count
--------- ------------- ------ ------ ------
Gi1/0/23  actor/long    1000   True   0
          partner/long


MLAG-Peer-A(config)#show interfaces utilization po1

Port    Load   Rx bits/s    Rx packets/s  Tx bits/s    Tx packets/s
Channel Interval
------- ------ ------------- -----------  ---------    -----------
Po1     300    792           1            1192         2

MLAG-Peer-A(config)#show vpc role

Self
----

Keep-alive admin status........................ Enabled
Keep-alive operational status.................. Enabled
Priority....................................... 100
System MAC address............................. 001E.C9DE.C52B
Timeout........................................ 5
VPC state...................................... Primary
VPC role....................................... Primary

Peer
----

Priority....................................... 100
VPC role....................................... Secondary
System MAC address............................. 001E.C9dE.C513
```

```
MLAG-Peer-B#show vpc statistics peer-link

Peer link control messages transmitted......... 95
Peer link control messages Tx errors........... 0
Peer link control messages Tx timeout.......... 0
Peer link control messages ACK transmitted..... 37
Peer link control messages ACK Tx errors....... 0
Peer link control messages received............ 37
Peer link data messages transmitted............ 777
Peer link data messages Tx errors.............. 0
Peer link data messages Tx timeout............. 0
Peer link data messages received............... 878
Peer link BPDU's transmitted to peer........... 2
Peer link BPDU's Tx errors..................... 0
Peer link BPDU's received from peer............ 11
Peer link BPDU's Rx errors..................... 0
Peer link LACPDU's transmitted to peer......... 775
Peer link LACPDU's Tx errors................... 0
Peer link LACPDU's received from peer.......... 867
Peer link LACPDU's Rx errors................... 0

MLAG-Peer-B#show vpc statistics peer-keepalive

Total transmitted.............................. 15545
Tx successful.................................. 15545
Tx errors...................................... 0
Total received................................. 15542
Rx successful.................................. 15542
Rx Errors...................................... 0
Timeout counter................................ 0
```

A Complete MLAG Example

The following example configures eight VLANs (10–17) across two VPCs. VPC 1 is connected to a Dell Networking N2048 over two links (gi1/0/23-24) over port-channel 2 on each MLAG peer. Interfaces Te1/0/1-2 on each MLAG peer connect to each other on port-channel 1 utilizing LACP. UDLD is enabled on the two MLAG peer-links and the timers are configured to the minimum values. DCPDP is enabled on VLAN 100 (interface gi1/0/8 on each MLAG peer). VLAN 100 is excluded from any MLAG interface, including the peer-link.

VPC 2 is connected to a legacy Cisco 3750 over port-channel 3 on each MLAG peer, and is also running LACP. The Cisco configuration is shown for completeness.

Spanning tree instance 0 is configured for VLAN 1. Spanning tree instance 1 is configured for VLANs 10–17. Spanning-tree instance 2 carries VLAN 100 traffic on a redundant link between the two MLAG peers. The Cisco 3750 acts as the root bridge for the topology.

To support the redundant link using VLAN 200 running in MST instance 2 across gi1/0/8, configure the peer link with a high path cost for instance 2 on the primary switch to discourage forwarding across the peer link. Likewise, on the MLAG secondary switch, set the bridge priority to 0 for instance 2 to encourage the secondary switch to select the root path. Be sure to name the MST domain.

MLAG Peer A Configuration

Current Configuration:

- System Description “Dell Networking N3024F, 6.0.0.0, Linux 3.6.5-858bcf6e”
- System Software Version 6.0.0.0

```
console#configure
console(config)#vlan 10-17,100
console(config-vlan10-17)#exit

console(config)#hostname “MLAG-Peer-A”
MLAG-Peer-A(config)#slot 1/0 2    ! Dell Networking N3024F
MLAG-Peer-A(config-stack)#stack
MLAG-Peer-A(config-stack)#member 1 2    ! N3024F
MLAG-Peer-A(config-stack)#exit
MLAG-Peer-A(config)#interface vlan 100
```

```
MLAG-Peer-A(config-if-vlan100)#ip address 192.168.0.1 255.255.255.0
MLAG-Peer-A(config-if-vlan100)#exit

MLAG-Peer-A(config)#spanning-tree mode mst
MLAG-Peer-A(config)#spanning-tree mst configuration
MLAG-Peer-A(config-mst)#instance 1 add vlan 10-17
MLAG-Peer-A(config-mst)#name "MLAG-A"
MLAG-Peer-A(config-mst)#revision 0
MLAG-Peer-A(config-mst)#instance 2 add vlan 100
MLAG-Peer-A(config-mst)#exit

MLAG-Peer-A(config)#udld enable
MLAG-Peer-A(config)#udld message time 7
MLAG-Peer-A(config)#udld timeout interval 9

MLAG-Peer-A(config)#interface Gi1/0/1
MLAG-Peer-A(config-if-Gi1/0/1)#channel-group 3 mode active
MLAG-Peer-A(config-if-Gi1/0/1)#description "Old-Iron-Partner-Link"
MLAG-Peer-A(config-if-Gi1/0/1)#exit

MLAG-Peer-A(config)#interface Gi1/0/8
MLAG-Peer-A(config-if-Gi1/0/8)#switchport access vlan 100
MLAG-Peer-A(config-if-Gi1/0/8)#exit

MLAG-Peer-A(config)#interface Gi1/0/23
MLAG-Peer-A(config-if-Gi1/0/23)#channel-group 2 mode active
MLAG-Peer-A(config-if-Gi1/0/23)#description "MLAG-Partner-Link"
MLAG-Peer-A(config-if-Gi1/0/23)#exit

MLAG-Peer-A(config)#interface Gi1/0/24
MLAG-Peer-A(config-if-Gi1/0/24)#channel-group 2 mode active
MLAG-Peer-A(config-if-Gi1/0/24)#description "MLAG-Partner-Link"
MLAG-Peer-A(config-if-Gi1/0/24)#exit

MLAG-Peer-A(config)#interface Te1/0/1
MLAG-Peer-A(config-if-Te1/0/1)#channel-group 1 mode active
MLAG-Peer-A(config-if-Te1/0/1)#description "MLAG-Peer-Link"
MLAG-Peer-A(config-if-Te1/0/1)#udld enable
MLAG-Peer-A(config-if-Te1/0/1)#udld port aggressive
MLAG-Peer-A(config-if-Te1/0/1)#exit

MLAG-Peer-A(config)#interface Te1/0/2
MLAG-Peer-A(config-if-Te1/0/2)#channel-group 1 mode active
MLAG-Peer-A(config-if-Te1/0/2)#description "MLAG-Peer-Link"
MLAG-Peer-A(config-if-Te1/0/2)#udld enable
MLAG-Peer-A(config-if-Te1/0/2)#udld port aggressive
```

```
MLAG-Peer-A(config-if-Te1/0/2)#exit

MLAG-Peer-A(config)#interface port-channel 1
MLAG-Peer-A(config-if-Po1)#description "MLAG-Peer-Link"
MLAG-Peer-A(config-if-Po1)#switchport mode trunk
MLAG-Peer-A(config-if-Po1)#switchport trunk allowed vlan 1-99,101-
4093
MLAG-Peer-A(config-if-Po1)#vpc peer-link
MLAG-Peer-A(config-if-Po1)#spanning-tree mst 2 cost 50000
MLAG-Peer-A(config-if-Po1)#exit

MLAG-Peer-A(config)#interface port-channel 2
MLAG-Peer-A(config-if-Po2)#switchport mode trunk
MLAG-Peer-A(config-if-Po2)#switchport trunk allowed vlan 1-99,101-
4093
MLAG-Peer-A(config-if-Po2)#vpc 1
MLAG-Peer-A(config-if-Po2)#exit

MLAG-Peer-A(config)#interface port-channel 3
MLAG-Peer-A(config-if-Po3)#description "Old-Iron-Partner-Link"
MLAG-Peer-A(config-if-Po3)#switchport mode trunk
MLAG-Peer-A(config-if-Po3)#switchport trunk allowed vlan 1-99,101-
4093
MLAG-Peer-A(config-if-Po3)#vpc 2
MLAG-Peer-A(config-if-Po3)#exit

MLAG-Peer-A(config)#snmp-server engineid local
800002a203001ec9dec52b
MLAG-Peer-A(config)#snmp-server agent boot count 2
MLAG-Peer-A(config)#feature vpc
MLAG-Peer-A(config)#vpc domain 1
MLAG-Peer-A(config-vpc 1)#peer-keepalive enable
MLAG-Peer-A(config-vpc 1)#peer-keepalive destination 192.168.0.2
source 192.168.0.1
MLAG-Peer-A(config-vpc 1)#peer detection enable
MLAG-Peer-A(config-vpc 1)#exit
MLAG-Peer-A(config)#exit
```

MLAG Peer B Configuration

Current Configuration:

- System Description "Dell Networking N3024F, 6.0.0.0, Linux 3.6.5-858bcf6e"
- System Software Version 6.0.0.0

```
console#configure
console(config)#vlan 10-17,100
console(config-vlan10-17)#exit

console(config)#hostname "MLAG-Peer-B"
MLAG-Peer-B(config)#slot 1/0 2    ! Dell Networking N3024F
MLAG-Peer-B(config-stack)#stack
MLAG-Peer-B(config-stack)#member 1 2    ! N3024F
MLAG-Peer-B(config-stack)#exit

MLAG-Peer-B(config)#interface vlan 100
MLAG-Peer-B(config-if-vlan100)#ip address 192.168.0.2 255.255.255.0
MLAG-Peer-B(config-if-vlan100)#exit

MLAG-Peer-B(config)#spanning-tree mode mst
MLAG-Peer-B(config)#spanning-tree mst configuration
MLAG-Peer-B(config-mst)#instance 1 add vlan 10-17
MLAG-Peer-B(config-mst)#spanning-tree mst 2 priority 0
MLAG-Peer-B(config-mst)#name "MLAG-A"
MLAG-Peer-B(config-mst)#revision 0
MLAG-Peer-B(config-mst)#instance 2 add vlan 100
MLAG-Peer-B(config-mst)#exit

MLAG-Peer-B(config)#udld enable
MLAG-Peer-B(config)#udld message time 7
MLAG-Peer-B(config)#udld timeout interval 9

MLAG-Peer-B(config)#interface Gi1/0/1
MLAG-Peer-B(config-if-Gi1/0/1)#channel-group 3 mode active
MLAG-Peer-B(config-if-Gi1/0/1)#description "Old-Iron-Partner-Link"
MLAG-Peer-B(config-if-Gi1/0/1)#exit

MLAG-Peer-B(config)#interface Gi1/0/8
MLAG-Peer-B(config-if-Gi1/0/8)#switchport access vlan 100
MLAG-Peer-B(config-if-Gi1/0/8)#exit

MLAG-Peer-B(config)#interface Gi1/0/23
MLAG-Peer-B(config-if-Gi1/0/23)#channel-group 2 mode active
```

```
MLAG-Peer-B(config-if-Gi1/0/23)#description "MLAG-Partner-Link"
MLAG-Peer-B(config-if-Gi1/0/23)#exit

MLAG-Peer-B(config)#interface Gi1/0/24
MLAG-Peer-B(config-if-Gi1/0/24)#channel-group 2 mode active
MLAG-Peer-B(config-if-Gi1/0/24)#description "MLAG-Partner-Link"
MLAG-Peer-B(config-if-Gi1/0/24)#exit

MLAG-Peer-B(config)#interface Te1/0/1
MLAG-Peer-B(config-if-Te1/0/1)#channel-group 1 mode active
MLAG-Peer-B(config-if-Te1/0/1)#description "MLAG-Peer-Link"
MLAG-Peer-B(config-if-Te1/0/1)#udld enable
MLAG-Peer-B(config-if-Te1/0/1)#udld port aggressive
MLAG-Peer-B(config-if-Te1/0/1)#exit

MLAG-Peer-B(config)#interface Te1/0/2
MLAG-Peer-B(config-if-Te1/0/2)#channel-group 1 mode active
MLAG-Peer-B(config-if-Te1/0/2)#description "MLAG-Peer-Link"
MLAG-Peer-B(config-if-Te1/0/2)#udld enable
MLAG-Peer-B(config-if-Te1/0/2)#udld port aggressive
MLAG-Peer-B(config-if-Te1/0/2)#exit

MLAG-Peer-B(config)#interface port-channel 1
MLAG-Peer-B(config-if-Po1)#description "MLAG-Peer-Link"
MLAG-Peer-B(config-if-Po1)#switchport mode trunk
MLAG-Peer-B(config-if-Po1)#switchport trunk allowed vlan 1-99,101-
4093
MLAG-Peer-B(config-if-Po1)#vpc peer-link
MLAG-Peer-B(config-if-Po1)#exit

MLAG-Peer-B(config)#interface port-channel 2
MLAG-Peer-B(config-if-Po2)#switchport mode trunk
MLAG-Peer-B(config-if-Po2)#switchport trunk allowed vlan 1-99,101-
4093
MLAG-Peer-B(config-if-Po2)#vpc 1
MLAG-Peer-B(config-if-Po2)#exit

MLAG-Peer-B(config)#interface port-channel 3
MLAG-Peer-B(config-if-Po3)#description "Old-Iron-Partner-Link"
MLAG-Peer-B(config-if-Po3)#switchport mode trunk
MLAG-Peer-B(config-if-Po3)#switchport trunk allowed vlan 1-99,101-
4093
MLAG-Peer-B(config-if-Po3)#vpc 2
MLAG-Peer-B(config-if-Po3)#exit
```

```
MLAG-Peer-B(config)#snmp-server engineid local
800002a203001ec9dec513
MLAG-Peer-B(config)#snmp-server agent boot count 3
MLAG-Peer-B(config)#feature vpc
MLAG-Peer-B(config)#vpc domain 1
MLAG-Peer-B(config-vpc 1)#peer-keepalive enable
MLAG-Peer-B(config-vpc 1)#peer-keepalive destination 192.168.0.1
source 192.168.0.2
MLAG-Peer-B(config-vpc 1)#peer detection enable
MLAG-Peer-B(config-vpc 1)#exit
MLAG-Peer-B(config)#exit
```

MLAG Partner Configuration

Current Configuration:

- System Description "Dell Networking N2048, 6.0.0.0, Linux 3.6.5-858bcf6e"
- System Software Version 6.0.0.0

```
console#configure
console(config)#hostname "LAG-SW"
LAG-SW(config)#slot 1/0 5    ! Dell Networking N2048
LAG-SW(config-stack)#stack
LAG-SW(config-stack)#member 1 8    ! N2048
LAG-SW(config-stack)#exit

LAG-SW(config)#interface vlan 1
LAG-SW(config-if-vlan1)#ip address dhcp
LAG-SW(config-if-vlan1)#exit

LAG-SW(config)#spanning-tree mode mst
LAG-SW(config)#spanning-tree mst configuration
LAG-SW(config-mst)#instance 1 add vlan 10-17
LAG-SW(config-mst)#name MLAG-A
LAG-SW(config-mst)#revision 0
LAG-SW(config-mst)#exit

LAG-SW(config)#interface Gi1/0/1
LAG-SW(config-if-Gi1/0/1)#channel-group 1 mode active
LAG-SW(config-if-Gi1/0/1)#exit

LAG-SW(config)#interface Gi1/0/2
LAG-SW(config-if-Gi1/0/2)#channel-group 1 mode active
LAG-SW(config-if-Gi1/0/2)#exit

LAG-SW(config)#interface Gi1/0/3
```

```
LAG-SW(config-if-Gi1/0/3)#channel-group 1 mode active
LAG-SW(config-if-Gi1/0/3)#exit

LAG-SW(config)#interface Gi1/0/4
LAG-SW(config-if-Gi1/0/4)#channel-group 1 mode active
LAG-SW(config-if-Gi1/0/4)#exit

LAG-SW(config)#interface port-channel 1
LAG-SW(config-if-Po1)#switchport mode trunk
LAG-SW(config-if-Po1)#exit

LAG-SW(config)#snmp-server engineid local 800002a203001ec9deb777
LAG-SW(config)#snmp-server agent boot count 3
LAG-SW(config)#exit
```

Cisco 3750 MLAG Partner Configuration

Current configuration: 1913 bytes

- version 12.2
- no service pad
- service timestamps debug datetime msec
- service timestamps log datetime msec
- no service password-encryption
- service unsupported-transceiver

```
config
hostname Switch

boot-start-marker
boot-end-marker

no aaa new-model
switch 1 provision ws-c3750g-24ts
system mtu routing 1500
ip subnet-zero

spanning-tree mode mst
spanning-tree extend system-id

spanning-tree mst configuration
 instance 1 vlan 10-17
 name MLAG-A
 revision 0
```

```
vlan internal allocation policy ascending
interface Port-channel1
 switchport trunk encapsulation dot1q
 switchport mode trunk
interface GigabitEthernet1/0/1
interface GigabitEthernet1/0/2
interface GigabitEthernet1/0/3
interface GigabitEthernet1/0/4
interface GigabitEthernet1/0/5
interface GigabitEthernet1/0/6
interface GigabitEthernet1/0/7
interface GigabitEthernet1/0/8
interface GigabitEthernet1/0/9
interface GigabitEthernet1/0/10
interface GigabitEthernet1/0/11
interface GigabitEthernet1/0/12
interface GigabitEthernet1/0/13
interface GigabitEthernet1/0/14
interface GigabitEthernet1/0/15
interface GigabitEthernet1/0/16
interface GigabitEthernet1/0/17
interface GigabitEthernet1/0/18
interface GigabitEthernet1/0/19
interface GigabitEthernet1/0/20
interface GigabitEthernet1/0/21
interface GigabitEthernet1/0/22
interface GigabitEthernet1/0/23
interface GigabitEthernet1/0/24
interface GigabitEthernet1/0/25

 description "MLAG-Peer-Link"
 switchport trunk encapsulation dot1q
 switchport mode trunk
 channel-group 1 mode active

interface GigabitEthernet1/0/26
 description "MLAG-Peer-Link"
 switchport trunk encapsulation dot1q
 switchport mode trunk
 channel-group 1 mode active

interface GigabitEthernet1/0/27
interface GigabitEthernet1/0/28
interface Vlan1
 no ip address
```

```
ip classless
ip http server
ip http secure-server

control-plane

line con 0
line vty 5 15

end
```

Status Reporting

The following shows the status of various components of the switches in the above configuration. The switch prompts identify the switch on which the status is shown. To obtain accurate status, the commands below are run on the primary MLAG switch unless noted otherwise.

Spanning Tree Status

```
Old-Iron-3750#show spanning-tree

MST0
  Spanning tree enabled protocol mstp
  Root ID    Priority    32768
             Address     0013.c4bd.f080
             This bridge is the root
            Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID  Priority    32768  (priority 32768 sys-id-ext 0)
             Address     0013.c4bd.f080
            Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- --------------------
Po1                 Desg FWD 10000     128.488  P2p Bound(STP)



MST1
  Spanning tree enabled protocol mstp
  Root ID    Priority    32769
             Address     0013.c4bd.f080
             This bridge is the root
            Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID  Priority    32769  (priority 32768 sys-id-ext 1)
             Address     0013.c4bd.f080
            Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- --------------------
Po1                 Desg FWD 10000     128.488  P2p Bound(STP)
```

```
LAG-SW#show spanning-tree

Spanning tree Enabled  BPDU flooding Disabled Portfast BPDU
filtering  Disabled mode mst
CST Regional Root:        80:00:00:1E:C9:DE:B7:77
Regional Root Path Cost:  0

###### MST 0 Vlan Mapped:   1
ROOT ID
              Priority        32768
              Address         0013.C4BD.F080
              Path Cost       5000
              Root Port       Po1
              Hello Time 2 Sec Max Age 20 sec Forward Delay 15 sec
              Bridge Max Hops 20
Bridge ID
              Priority        32768
              Address         001E.C9DE.B777
              Hello Time 2 Sec Max Age 20 sec Forward Delay 15 sec
TxHoldCount 6 sec

Name      State    Prio.Nbr  Cost      Sts  Role  RestrictedPort
--------- -------- --------- --------- ---- ----- --------------
Gi1/0/1   Enabled  128.1     0         DIS  Disb  No
Gi1/0/2   Enabled  128.2     0         DIS  Disb  No
Gi1/0/3   Enabled  128.3     0         DIS  Disb  No
Gi1/0/4   Enabled  128.4     0         DIS  Disb  No
Gi1/0/5   Enabled  128.5     0         DIS  Disb  No
Gi1/0/6   Enabled  128.6     0         DIS  Disb  No
Gi1/0/7   Enabled  128.7     0         DIS  Disb  No
Gi1/0/8   Enabled  128.8     0         DIS  Disb  No
Gi1/0/9   Enabled  128.9     0         DIS  Disb  No
Gi1/0/10  Enabled  128.10    0         DIS  Disb  No
Gi1/0/11  Enabled  128.11    0         DIS  Disb  No
Gi1/0/12  Enabled  128.12    0         DIS  Disb  No
Gi1/0/13  Enabled  128.13    0         DIS  Disb  No
Gi1/0/14  Enabled  128.14    0         DIS  Disb  No
Gi1/0/15  Enabled  128.15    0         DIS  Disb  No
Gi1/0/16  Enabled  128.16    0         DIS  Disb  No
Gi1/0/17  Enabled  128.17    0         DIS  Disb  No
Gi1/0/18  Enabled  128.18    0         DIS  Disb  No
Gi1/0/19  Enabled  128.19    0         DIS  Disb  No
Gi1/0/20  Enabled  128.20    0         DIS  Disb  No
Gi1/0/21  Enabled  128.21    0         DIS  Disb  No
Gi1/0/22  Enabled  128.22    0         DIS  Disb  No
Gi1/0/23  Enabled  128.23    0         DIS  Disb  No
```

```
Gi1/0/24  Enabled  128.24    0         DIS  Disb  No
Gi1/0/25  Enabled  128.25    0         DIS  Disb  No
Gi1/0/26  Enabled  128.26    0         DIS  Disb  No
Gi1/0/27  Enabled  128.27    0         DIS  Disb  No
Gi1/0/28  Enabled  128.28    0         DIS  Disb  No
Gi1/0/29  Enabled  128.29    0         DIS  Disb  No
Gi1/0/30  Enabled  128.30    0         DIS  Disb  No
Gi1/0/31  Enabled  128.31    0         DIS  Disb  No
Gi1/0/32  Enabled  128.32    0         DIS  Disb  No
Gi1/0/33  Enabled  128.33    0         DIS  Disb  No
Gi1/0/34  Enabled  128.34    0         DIS  Disb  No
Gi1/0/35  Enabled  128.35    0         DIS  Disb  No
Gi1/0/36  Enabled  128.36    0         DIS  Disb  No
Gi1/0/37  Enabled  128.37    0         DIS  Disb  No
Gi1/0/38  Enabled  128.38    0         DIS  Disb  No
Gi1/0/39  Enabled  128.39    0         DIS  Disb  No
Gi1/0/40  Enabled  128.40    0         DIS  Disb  No
Gi1/0/41  Enabled  128.41    0         DIS  Disb  No
Gi1/0/42  Enabled  128.42    0         DIS  Disb  No
Gi1/0/43  Enabled  128.43    0         DIS  Disb  No
Gi1/0/44  Enabled  128.44    0         DIS  Disb  No
Gi1/0/45  Enabled  128.45    0         DIS  Disb  No
Gi1/0/46  Enabled  128.46    0         DIS  Disb  No
Gi1/0/47  Enabled  128.47    0         DIS  Disb  No
Gi1/0/48  Enabled  128.48    0         DIS  Disb  No
Te1/0/1   Enabled  128.49    0         DIS  Disb  No
Te1/0/2   Enabled  128.50    0         DIS  Disb  No
Tw1/0/1   Enabled  128.51    0         DIS  Disb  No
Tw1/0/2   Enabled  128.52    0         DIS  Disb  No
Po1       Enabled  96.650    5000      FWD  Root  No
Po2       Enabled  96.651    0         DIS  Disb  No
Po3       Enabled  96.652    0         DIS  Disb  No
Po4       Enabled  96.653    0         DIS  Disb  No
Po5       Enabled  96.654    0         DIS  Disb  No
Po6       Enabled  96.655    0         DIS  Disb  No
Po7       Enabled  96.656    0         DIS  Disb  No
Po8       Enabled  96.657    0         DIS  Disb  No
Po9       Enabled  96.658    0         DIS  Disb  No
Po10      Enabled  96.659    0         DIS  Disb  No
Po11      Enabled  96.660    0         DIS  Disb  No
Po12      Enabled  96.661    0         DIS  Disb  No
Po13      Enabled  96.662    0         DIS  Disb  No
Po14      Enabled  96.663    0         DIS  Disb  No
Po15      Enabled  96.664    0         DIS  Disb  No
Po16      Enabled  96.665    0         DIS  Disb  No
Po17      Enabled  96.666    0         DIS  Disb  No
```

```
Po18      Enabled  96.667    0          DIS  Disb  No
Po19      Enabled  96.668    0          DIS  Disb  No
Po20      Enabled  96.669    0          DIS  Disb  No
Po21      Enabled  96.670    0          DIS  Disb  No
Po22      Enabled  96.671    0          DIS  Disb  No
Po23      Enabled  96.672    0          DIS  Disb  No
Po24      Enabled  96.673    0          DIS  Disb  No
Po25      Enabled  96.674    0          DIS  Disb  No
Po26      Enabled  96.675    0          DIS  Disb  No
Po27      Enabled  96.676    0          DIS  Disb  No
Po28      Enabled  96.677    0          DIS  Disb  No
Po29      Enabled  96.678    0          DIS  Disb  No
Po30      Enabled  96.679    0          DIS  Disb  No
Po31      Enabled  96.680    0          DIS  Disb  No
Po32      Enabled  96.681    0          DIS  Disb  No
Po33      Enabled  96.682    0          DIS  Disb  No
Po34      Enabled  96.683    0          DIS  Disb  No
Po35      Enabled  96.684    0          DIS  Disb  No
Po36      Enabled  96.685    0          DIS  Disb  No
Po37      Enabled  96.686    0          DIS  Disb  No
Po38      Enabled  96.687    0          DIS  Disb  No
Po39      Enabled  96.688    0          DIS  Disb  No
Po40      Enabled  96.689    0          DIS  Disb  No
Po41      Enabled  96.690    0          DIS  Disb  No
Po42      Enabled  96.691    0          DIS  Disb  No
Po43      Enabled  96.692    0          DIS  Disb  No

MLAG-Peer-A#show spanning-tree

Spanning tree Enabled  BPDU flooding Disabled Portfast BPDU
filtering  Disabled mode mst
CST Regional Root:        80:00:00:13:C4:BD:F0:80
Regional Root Path Cost:  200

###### MST 0 Vlan Mapped:   1, 100
ROOT ID
              Priority        32768
              Address         0013.C4BD.F080
              Path Cost       0
              Root Port       Po3
              Hello Time 2 Sec Max Age 20 sec Forward Delay 15 sec
              Bridge Max Hops 20
Bridge ID
              Priority        32768
              Address         001E.C9DE.C52B
```

```
             Hello Time 2 Sec Max Age 20 sec Forward Delay 15 sec
TxHoldCount 6 sec

Name      State    Prio.Nbr  Cost      Sts  Role  RestrictedPort
--------- -------- --------- --------- ---- ----- --------------
Gi1/0/1   Enabled  128.1     0         DIS  Disb  No
Gi1/0/2   Enabled  128.2     0         DIS  Disb  No
Gi1/0/3   Enabled  128.3     0         DIS  Disb  No
Gi1/0/4   Enabled  128.4     0         DIS  Disb  No
Gi1/0/5   Enabled  128.5     0         DIS  Disb  No
Gi1/0/6   Enabled  128.6     0         DIS  Disb  No
Gi1/0/7   Enabled  128.7     0         DIS  Disb  No
Gi1/0/8   Enabled  128.8     20000     FWD  Desg  No
Gi1/0/9   Enabled  128.9     0         DIS  Disb  No
Gi1/0/10  Enabled  128.10    0         DIS  Disb  No
Gi1/0/11  Enabled  128.11    0         DIS  Disb  No
Gi1/0/12  Enabled  128.12    0         DIS  Disb  No
Gi1/0/13  Enabled  128.13    0         DIS  Disb  No
Gi1/0/14  Enabled  128.14    0         DIS  Disb  No
Gi1/0/15  Enabled  128.15    0         DIS  Disb  No
Gi1/0/16  Enabled  128.16    0         DIS  Disb  No
Gi1/0/17  Enabled  128.17    0         DIS  Disb  No
Gi1/0/18  Enabled  128.18    0         DIS  Disb  No
Gi1/0/19  Enabled  128.19    0         DIS  Disb  No
Gi1/0/20  Enabled  128.20    0         DIS  Disb  No
Gi1/0/21  Enabled  128.21    0         DIS  Disb  No
Gi1/0/22  Enabled  128.22    0         DIS  Disb  No
Gi1/0/23  Enabled  128.23    0         DIS  Disb  No
Gi1/0/24  Enabled  128.24    0         DIS  Disb  No
Te1/0/1   Enabled  128.25    0         DIS  Disb  No
Te1/0/2   Enabled  128.26    0         DIS  Disb  No
Tw1/0/1   Enabled  128.27    0         DIS  Disb  No
Tw1/0/2   Enabled  128.28    0         DIS  Disb  No
Po1       Disabled 96.650    0         FWD  Disb  No
Po2       Enabled  96.651    5000      FWD  Desg  No
Po3       Enabled  96.652    200       FWD  Root  No
Po4       Enabled  96.653    0         DIS  Disb  No
```

MLAG Status

```
MLAG-Peer-A#show vpc brief

VPC config Mode................................ Enabled
Keepalive config mode.......................... Enabled
VPC operational Mode........................... Enabled
Self Role...................................... Primary
Peer Role...................................... Secondary
Peer detection................................. Peer detected, VPC
Operational

Peer-Link details
-----------------

Interface...................................... Po1
Peer link status............................... UP
Peer-link STP Mode............................. Enabled
Configured Vlans...............................
1,10,11,12,13,14,15,16,17
Egress tagging.................................
10,11,12,13,14,15,16,17

VPC Details
-----------

Number of VPCs configured...................... 2
Number of VPCs operational..................... 2


VPC id# 1
-----------
Interface...................................... Po2
Configured Vlans...............................
1,10,11,12,13,14,15,16,17
VPC Interface State............................ Active

Local MemberPorts    Status
-----------------    ------
Gi1/0/23             UP
Gi1/0/24             UP

Peer MemberPorts     Status
----------------     ------
Gi1/0/23             UP
Gi1/0/24             UP
```

```
VPC id# 2
-----------
Interface...................................... Po3
Configured Vlans...............................
1,10,11,12,13,14,15,16,17
VPC Interface State............................ Active

MLAG-Peer-A#show vpc 1

VPC id# 1
-----------------
Config mode.................................... Enabled
Operational mode............................... Enabled
Port channel................................... Po2

Local MemberPorts    Status
-----------------    ------
Gi1/0/23             UP
Gi1/0/24             UP

Peer MemberPorts     Status
----------------     ------
Gi1/0/23             UP
Gi1/0/24             UP

MLAG-Peer-A#show vpc 2

VPC id# 2
-----------------
Config mode.................................... Enabled
Operational mode............................... Enabled
Port channel................................... Po3

Local MemberPorts    Status
-----------------    ------
Gi1/0/1              UP

Peer MemberPorts     Status
----------------     ------
Gi1/0/1              UP
```

```
MLAG-Peer-A#show vpc peer-keepalive

Peer IP address................................ 192.168.0.2
Source IP address.............................. 192.168.0.1
UDP port....................................... 50000
Peer detection................................. Enabled
Peer detection operational status.............. Up
Peer is detected............................... TRUE


MLAG-Peer-A#show vpc statistics peer-keepalive

Total transmitted.............................. 20908
Tx successful.................................. 20908
Tx errors...................................... 0
Total received................................. 20835
Rx successful.................................. 20835
Rx Errors...................................... 0
Timeout counter................................ 1

MLAG-Peer-A#show vpc statistics peer-link

Peer link control messages transmitted......... 75
Peer link control messages Tx errors........... 0
Peer link control messages Tx timeout.......... 0
Peer link control messages ACK transmitted..... 119
Peer link control messages ACK Tx errors....... 0
Peer link control messages received............ 119
Peer link data messages transmitted............ 1294
Peer link data messages Tx errors.............. 0
Peer link data messages Tx timeout............. 0
Peer link data messages received............... 1886
Peer link BPDU's transmitted to peer........... 11
Peer link BPDU's Tx errors..................... 0
Peer link BPDU's received from peer............ 751
Peer link BPDU's Rx errors..................... 0
Peer link LACPDU's tranmsitted to peer......... 1283
Peer link LACPDU's Tx errors................... 0
Peer link LACPDU's received from peer.......... 1135
Peer link LACPDU's Rx errors................... 0
```

29

Data Center Bridging Features

Dell Networking N4000 Series Switches

This chapter describes how to manage the features developed for use in data center environments but often used in a variety of 10G applications.

NOTE: The data center bridging features described in this chapter are available on the Dell Networking N4000 Series switches only.

The topics covered in this chapter include:

- Data Center Bridging Technology Overview
- Priority Flow Control (PFC)
- DCB Capability Exchange (DCBX)
- Enhanced Transmission Selection (ETS)

Data Center Bridging Technology Overview

The Dell Networking N4000 Series switches support Data Center Bridging (DCB) features to increase the reliability of Ethernet-based networks in the data center.

The Dell Networking N4000 Series switches support PFC, ETS, and DCBX capability exchange, with the ability to autoconfigure from a peer switch.

The Ethernet enhancements that DCB provides are well suited for iSCSI applications.

Table 29-1 provides a summary of the features this chapter describes.

Table 29-1. Data Center Features

Feature	Description
PFC	Provides a way to distinguish which traffic on a physical link is paused when congestion occurs based on the priority of the traffic.
DCBx	Allows DCB devices to exchange configuration information, using type-length-value (TLV) information elements over LLDP, with directly connected peers.

Table 29-1. Data Center Features (Continued)

Feature	Description
ETS	Supports the ETS configuration and Application Priority TLVs, which are accepted from auto-upstream devices and propagated to auto-downstream devices. The Dell Networking N4000 Series switches support the automatic configuration of the switch with received ETS parameters.

Default DCB Values

Table 29-2 lists the default values for the DCB features that this chapter describes.

Table 29-2. Default Port-Based Traffic Control Values

Feature	Default
PFC	Disabled, no priority classifications are configured.
DCBx version	Auto detect

Priority Flow Control

Ordinarily, when flow control is enabled on a physical link, it applies to all traffic on the link. When congestion occurs, the hardware sends pause frames that temporarily suspend traffic flow to help prevent buffer overflow and dropped frames.

PFC provides a means of pausing individual priorities within a single physical link. By pausing the congested priority or priorities independently, protocols that are highly loss-sensitive can share the same link with traffic that has different loss tolerances.

This feature is used in networks where the traffic has differing loss tolerances. For example, Fibre Channel traffic is highly sensitive to traffic loss. If a link contains both loss-sensitive data and other less loss-sensitive data, the loss-sensitive data should use a no-drop priority that is enabled for flow control.

Priorities are differentiated by the priority field of the IEEE 802.1Q VLAN header, which identifies an IEEE 802.1p priority value. These priority values must be mapped to internal class-of-service (CoS) values.

The PFC feature allows you to specify the CoS values that should be paused (due to greater loss sensitivity) instead of dropped when congestion occurs on a link. Unless configured as no-drop, all CoS priorities are considered non-pausable ("drop") when priority-based flow control is enabled until no-drop is specifically turned on.

PFC Operation and Behavior

PFC uses a control packet newly defined in IEEE 802.1Qbb and, therefore, is not compatible with IEEE 802.3 Annex 31B flow control. An interface that is configured for PFC is automatically disabled for flow control. When PFC is disabled on an interface, the flow control configuration for the interface becomes active. Any IEEE 802.3 Annex 31B link-layer flow-control frames received on a PFC configured interface are ignored.

Each priority is configured as either *drop* or *no-drop*. If a priority that is designated as no-drop is congested, the priority is paused. Drop priorities do not participate in pause. You must configure the same no-drop priorities and enable VLAN tagging for the no-drop priorities across the network to ensure end-to-end lossless behavior.

Operator configuration of PFC is used only when the port is configured in a manual role. When interoperating with other equipment in a manual role, the peer equipment must be configured with identical PFC priorities and VLAN assignments. Interfaces not enabled for PFC ignore received PFC frames. Ports configured in auto-upstream or auto-downstream roles receive their PFC configuration from the configuration source and ignore any manually configured information.

When using links in a port-channel, PFC must be configured on all of the individual links in the port-channel, as it operates on a link-local basis.

NOTE: This feature is configurable on physical full duplex interfaces only. To enable PFC on a LAG interface, the member interfaces must have the same configuration.

When PFC is disabled, the interface defaults to the IEEE 802.3 Annex 31B flow control setting for the interface. PFC is disabled by default.

If you enable priority-based flow control for a particular priority value on an interface, ensure that VLAN tagging is enabled on the interface so that the 802.1p priority values are carried through the network (see "VLAN Tagging " on page 704). Additionally, make sure that 802.1p priority values are mapped to CoS values (see "Class-of-Service " on page 1473). If DCBX is enabled, the manually configured PFC parameters (no-drop priorities) must match the peers PFC parameters. If they do not match, PFC will not be operationally enabled and the port will continue to operate in the configured legacy flow control mode (IEEE 802.3 Annex 31B mode or none).

PFC can be configured using the web interface and the command line interface.

Configuring PFC Using the Web Interface

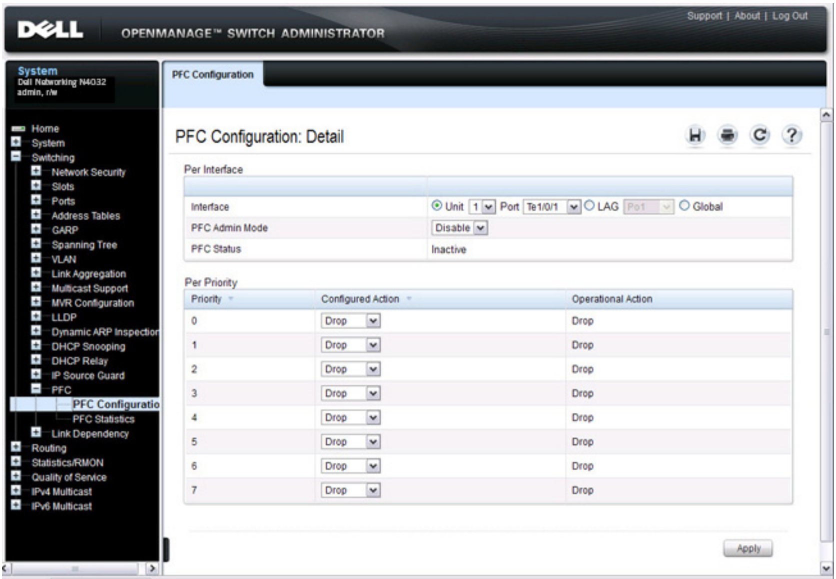
This section provides information about the OpenManage Switch Administrator pages to use to view and configure PFC on Dell Networking N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

PFC Configuration Page

Use the PFC Configuration page to enable priority flow control on one or more interfaces and to configure which priorities are subject to being paused to prevent data loss.

To display the **PFC Configuration** page, click **Switching** → **PFC** → **PFC Configuration** in the navigation menu.

Figure 29-1. PFC Configuration

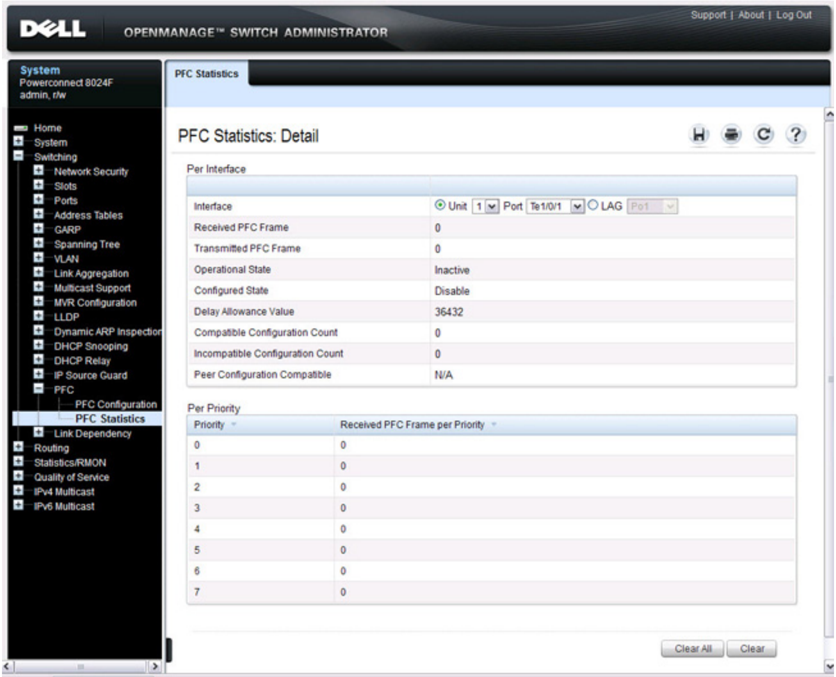


PFC Statistics Page

Use the **PFC Statistics** page to view the PFC statistics for interfaces on the switch.

To display the **PFC Statistics** page, click **Switching** → **PFC** → **PFC Statistics** in the navigation menu.

Figure 29-2. PFC Statistics



Configuring PFC Using the CLI

Beginning in Privileged EXEC mode, use the following commands to configure PFC.

NOTE: If DCBx is enabled and the switch is set to autoconfigure from a DCBX peer, configuring PFC is not necessary because the DCBx protocol automatically configures the PFC parameters.

Command	Purpose
configure	Enter global configuration mode.

Command	Purpose
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
datacenter-bridging	Enter the Data Center Bridging mode. PFC commands are issued from within this mode.
priority-flow-control mode on	Enable PFC on the interface(s) **NOTE:** It is unnecessary to set the priority flow control to enable if the **lldp dcbx port-role auto-down** or **lldp dcbx port-role auto-up** command has already been applied.
priority-flow-control priority *priority-id* **{drop | no-drop}**	Use the **no-drop** option to enable the priority group for lossless behavior. To enable lossy behavior, use the **drop** form of the command. *priority-id* — Specify the IEEE 802.1p priority value (range: 0–7) **NOTE:** Only two queues can be set to no-drop at one time.
CTRL + Z	Exit to Privileged EXEC mode.
show interfaces datacenter-bridging [*interface* | **port-channel** *port-channel-id*]	Display the datacenter-bridging configuration, status and counters for a given interface.
clear priority-flow-control statistics [*interface* | **port-channel** *port-channel-id*]	Clear all PFC statistics or the PFC statistics for the specified interface.

PFC Configuration Example

The network in this example handles both data and voice traffic. Because the voice traffic is time sensitive, it requires a higher priority than standard data traffic. The voice traffic uses VLAN 100 and has an 802.1p priority of 5, which is mapped to hardware queue 4. IP phones are connected to ports 3, 5, and 10, so PFC is enabled on these ports with 802.1p priority 5 traffic as no-drop. The configuration also enables VLAN tagging so that the 802.1p priority is identified. This example assumes the voice VLAN (VLAN 100) has already been configured. The administrator may need to perform other configuration steps to ensure suitable network operation, including configuring the bandwidth allocation for the drop and no-drop priorities or to enable or disable the DCBX protocol.

CAUTION: All ports may be briefly shutdown when modifying either flow control (FC) or PFC settings. PFC uses a control packet defined in 802.1Qbb and is not compatible with IEEE 802.3 Annex 31B flow control.

1 Map 802.1p priority 5 to traffic class 4. The following command changes the priority to traffic class mapping to be one-to-one, based upon the default switch settings. For lossless service, a priority must be mapped one-to-one to a traffic class. For more information about traffic classes, see "Class-of-Service " on page 1473. For a complete example of manually configuring a Dell Networking N4000 Series switch for iSCSI with PFC, refer to "iSCSI Optimization Configuration Examples " on page 587.

```
console#configure
console(config)#classofservice dot1p-mapping 5 4
```

2 Enter Interface Configuration mode for ports 3, 5, and 10, and then enter Data Center Bridging mode for these ports.

```
console(config)#interface range te1/0/3, te1/0/5,    te1/0/10
console(config-if)#datacenter-bridging
```

3 Enable PFC and configure traffic marked with 802.1p priority 5 to be paused rather than dropped when congestion occurs.

```
console(config-dcb)#priority-flow-control mode on
console(config-dcb)#priority-flow-control    priority 5 no-
drop
console(config-dcb)#exit
```

4 Enable VLAN tagging on the ports so the 802.1p priority is identified. Trunk mode can also be enabled on port-channels.

```
console(config-if)#switchport mode trunk
console(config-if)#exit
```

DCB Capability Exchange

The Data Center Bridging Exchange Protocol (DCBx) is used by DCB devices to exchange configuration information with directly connected peers. DCBx uses type-length-value (TLV) information elements over LLDP to exchange information, so LLDP must be enabled on the port to enable the information exchange. By default, LLDP is enabled on all ports. For more information, see "Discovering Network Devices " on page 825.

The main objective of DCBx is to perform the following operations:

- **Discovery of DCB capability in a peer**: DCBx is used to learn about the capabilities of the peer device. It is a means to determine if the peer device supports a particular feature such as PFC.
- **DCB feature misconfiguration detection**: DCBx can be used to detect misconfiguration of a feature between the peers on a link. Misconfiguration detection is feature-specific because some features may allow asymmetric configuration.
- **Peer configuration of DCB features**: DCBx can be used by a device to perform configuration of DCB features in its peer device if the peer device is willing to accept configuration.

For discussion and examples of configuring iSCSI with DCBX, refer to "iSCSI Optimization " on page 573.

The DCBx protocol supports the propagation of configuration information for the following features:

- Enhanced Transmission Selection (ETS)
- Priority-based Flow Control (PFC)
- Application Priorities

These features use DCBx to send and receive device configuration and capability information to the peer DCBx device.

The Application Priorities information is simply captured from the peer and potentially propagated to other peers by the DCBx component, as well as being configured when iSCSI is enabled on an operationally active PFC port.

DCBX information is carried over LLDP, which is a link-local protocol. When configuring links in a port-channel to use DCBX, the DCBX settings should be the same for all links in the port-channel.

Interoperability with IEEE DCBx

To be interoperable with legacy industry implementations of the DCBx protocol, The Dell Networking N4000 Series switches use a hybrid model to support both the IEEE version of DCBx (IEEE 802.1Qaz) and legacy DCBx versions.

The Dell Networking N4000 Series switch automatically detects whether a peer is operating with either of the two CEE DCBx versions or the IEEE standard DCBx version (the default mode). DCBx can also be configured to manually select one of the legacy versions or IEEE standard mode. In auto-detect mode, the switch starts operating in IEEE DCBx mode on a port, and if it detects a legacy DCBx device based on the OUI of the organization TLV, then the switch changes its DCBx mode on that port to support the version detected. There is no timeout mechanism to move back to IEEE mode. If the DCBx peer times out, multiple peers are detected, the link is reset (link down/up) or if commanded by the operator, DCBx resets its operational mode to IEEE.

The interaction between the DCBx component and other components remains the same irrespective of the operational mode it is executing. For instance, the DCBx component interacts with PFC to get needed information to pack the TLVs to be sent out on the interface. Based on the operational control mode of the port, DCBx packs it in the proper frame format.

DCBx and Port Roles

The behavior of each port is dependent on its operational mode and that of other ports in the stack. The port mode is a DCBx configuration item that is passed to the DCBx clients to control the processing of their configuration information. There are four port roles:

1. Manual
2. Auto-Upstream
3. Auto-Downstream
4. Configuration Source

Ports operating in the manual role do not have their configuration affected by peer devices or by internal propagation of configuration. These ports have their operational mode, traffic classes, and bandwidth information specified explicitly by the operator. These ports advertise their configuration to their peer if DCBx is enabled on that port. Incompatible peer configurations are logged and counted with an error counter.

The default operating mode for each port is manual. A port that is set to manual mode sets the willing bit for DCBx client TLVs to false. Manually-configured ports never internally propagate or accept internal or external configuration from other ports; in other words, a manual configuration discards any automatic configuration. Manually-configured ports may notify the operator of incompatible configurations if client configuration exchange over DCBx is enabled. Manually configured ports are always operationally enabled for DCBx clients, regardless of whether DCBx is enabled. Operationally enabled means that the port reports that it is able to operate using the current configuration.

A port operating in the auto-upstream role advertises a configuration, but is also willing to accept a configuration from the link-partner and propagate it internally to the auto-downstream ports, as well as receive configuration propagated internally by other auto-upstream ports. Specifically, the willing parameter is enabled on the port and the recommendation TLV is sent to the peer and processed if received locally. The first auto-upstream port to successfully accept a compatible configuration becomes the configuration source. The configuration source propagates its configuration to other auto-upstream and auto-downstream ports. Only the configuration source may propagate configuration to other ports internally. Auto-upstream ports that receive internally propagated information ignore their local configuration and utilize the internally propagated information.

Peer configurations received on auto-upstream ports other than the configuration source result in one of two possibilities. If the configuration is compatible with the configuration source, then the DCBx client becomes operationally active on the upstream port. If the configuration is not compatible with the configuration source, then a message is logged indicating an incompatible configuration, an error counter is incremented, and the DCBx client is operationally disabled on the port. The expectation is that the network administrator configures the upstream devices appropriately so that all such devices advertise a compatible configuration.

A port operating in the auto-downstream role advertises a configuration but is not willing to accept one from the link partner. However, the port will accept a configuration propagated internally by the configuration source. Specifically, the willing parameter is disabled on auto-downstream. By default, auto-downstream ports have the recommendation TLV parameter enabled. Auto-downstream ports that receive internally propagated information ignore their local configuration and utilize the internally propagated information. Auto-downstream ports propagate PFC, ETS, and application priority information received from the configuration source.

In the Configuration Source role, the port has been manually selected to be the configuration source. Configuration received over this port is propagated to the other auto configuration ports, however, no automatic election of a new configuration source port is allowed. Events that cause selection of a new configuration source are ignored. The configuration received over the configuration source port is maintained until cleared by the operator (set the port to the manual role).

Since it is not possible to configure the port role for a port-channel, it is recommended that the individual links have an identical port role configured on all links in the port-channel (auto-up or auto-down). Since only one port in the system can be configured as the configuration source, configuring interfaces as auto-up is a preferable alternative to a config-source setting.

Configuration Source Port Selection Process

When an auto-upstream or auto-downstream port receives a configuration from a peer, the DCBx client first checks if there is an active configuration source. If there is a configuration source already selected, the received configuration is checked against the local port operational values as received from the configuration source, and if compatible, the client marks the port as operationally enabled. If the configuration received from the peer is determined to not be compatible, a message is logged, an error counter is incremented and the DCBx clients become operationally disabled on the port. Operationally disabled means that PFC will not operate over the port. The port continues to keep link up and exchanges DCBx packets. If a compatible configuration is later received, the DCBx clients will become operationally enabled.

If there is no configuration source, a port may elect itself as the configuration source on a first-come, first-serve basis from the set of eligible ports. A port is eligible to become the configuration source if the following conditions are true:

- No other port is the configuration source.

- The port role is auto-upstream.
- The port is enabled with link up and DCBx enabled.
- The port has negotiated a DCBx relationship with the partner.
- The switch is capable of supporting the received configuration values, either directly or by translating the values into an equivalent configuration.

Whether or not the peer configuration is compatible with the configured values is NOT considered.

The newly elected configuration source propagates DCBx client information to the other ports and is internally marked as being the port over which configuration has been received. Configuration changes received from the peer over the configuration source port are propagated to the other auto configuration ports. Ports receiving auto configuration information from the configuration source ignore their current settings and utilize the configuration source information.

When a configuration source is selected, all auto-upstream ports other than the configuration source are marked as willing disabled.

To reduce flapping of configuration information, if the configuration source port is disabled, disconnected, or loses LLDP connectivity, the system clears the selection of configuration source port (if not manually selected) and enables the willing bit on all auto-upstream ports. The configuration on the auto configuration ports is not cleared (configuration holdover). If the user wishes to clear the configuration on the system in this scenario, the user can put the configuration source port into manual mode.

When a new port is selected as the configuration source, it is marked as the configuration source, the DCBx configuration is refreshed on all auto configuration ports, and each port may begin configuration negotiation with their peer again (if any information has changed).

Disabling DCBX

If it is desired to disable DCBX, the network operator can use the following commands to eliminate the transmission of DCBX TLVs in the LLDP frames on an interface:

```
no lldp tlv-select dcbxp application-priority
no lldp tlv-select dcbxp congestion-notification
no lldp tlv-select dcbxp ets-config
no lldp tlv-select dcbxp ets-recommend
```

```
no lldp tlv-select dcbxp pfc
```

These commands eliminate only the DCBX TLVs from use by LLDP. They do not otherwise affect any manually configured DCBX capabilities or the normal operation of LLDP.

Configuring DCBx

The CLI can be used to configure DCBX on Dell Networking N4000 Series switches.

Beginning in Privileged EXEC mode, use the following commands to configure DCBx.

Command	Purpose
configure	Enter global configuration mode.
lldp dcbx version {auto | cin | cee | ieee}	Optionally configure the administrative version for the DCBx protocol: • auto—Automatically select the version based on the peer response (default) • cin—Force the mode to Cisco-Intel-Nuova. (DCBx 1.0) • cee—Force the mode to CEE (DCBx 1.06) • ieee—Force the mode to IEEE 802.1Qaz
lldp tlv-select dcbxp [pfc | application-priority]	Enable LLDP to send specific DCBx TLVs if LLDP is enabled to transmit on the given interface. Entering the command with no parameters enables transmission of all TLVs. • pfc—Transmit the PFC configuration TLV • application-priority—Transmit the application priority TLV
interface *interface*	Enter interface configuration mode for the specified interface. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. A range of interfaces can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.

Command	Purpose
lldp tlv-select dcbxp [pfc | application-priority]	Override the global configuration for the LLDP DCBx TLVs on this interface. Entering the command with no parameters enables transmission of all TLVs. • pfc—Transmit the PFC configuration TLV. • application-priority—Transmit the application priority TLV.
lldp dcbx port-role {auto-up |auto-down | manual | configuration-source}	Configure the DCBx port role on the interface: • **auto-up**—Advertises a configuration, but is also willing to accept a configuration from the link-partner and propagate it internally to the auto-downstream ports as well as receive configuration propagated internally by other auto-upstream ports. These ports have the willing bit enabled. These ports should be connected to FCFs. • **auto-down**—Advertises a configuration but is not willing to accept one from the link partner. However, the port will accept a configuration propagated internally by the configuration source. These ports have the willing bit set to disabled. Selection of a port based upon compatibility of the received configuration is suppressed. These ports should be connected to a trusted FCF. • **manual**—Ports operating in the Manual role do not have their configuration affected by peer devices or by internal propagation of configuration. These ports will advertise their configuration to their peer if DCBx is enabled on that port. The willing bit is set to disabled on manual role ports. • **configuration-source**—In this role, the port has been manually selected to be the configuration source. Configuration received over this port is propagated to the other auto configuration ports. Selection of a port based upon compatibility of the received configuration is suppressed. These ports should be connected to a trusted FCF. These ports have the willing bit enabled.
CTRL + Z	Exit to Privileged EXEC mode.

Command	Purpose
show lldp tlv-select interface {**all** | *interface*}	Display the interface TLV configuration for all interfaces or for the specified interface.
show lldp dcbx interface {**all status** | *interface* [**detail**]}	Display the interface TLV configuration for all interfaces or for the specified interface.

Enhanced Transmission Selection

Networks classify and prioritize traffic to provide different service characteristics to end user traffic flows. Administrators may wish to guarantee or limit bandwidth for certain traffic, ensure lossless behavior for other traffic, and control the queue discipline/drop characteristics for best-effort traffic. Additionally, it is desirable for a switch to support sharing bandwidth among bursty sources while still enabling prioritization of time-sensitive or management traffic that requires minimum latency.

Enhanced transmission selection (ETS) provides uniform management for sharing bandwidth between congestion managed and traditional classes on a single bridged network. Using priority-based processing and weight allocations, Traffic Class Groups (TCGs) carrying different types of traffic such as LAN, SAN, and management traffic can be configured to provide minimum bandwidth guarantees, unused bandwidth sharing, and lossless or best-effort transmit characteristics. Dell Networking N4000 Series switches support strict priority and Weighted Deficit Round Robin (WDRR) scheduling with up to two lossless traffic classes. WDRR schedules traffic based on average bandwidth consumed vs. frame counts.

ETS Operation

The normal (default) operation of Dell Networking N-Series switches, when uncongested, is that packets are scheduled for output in the order in which they are received, that is, using FIFO scheduling. The class of service (CoS) mechanism enables the administrator to schedule packets for output ahead of other packets when the switch is congested, choose which type of packets to drop when the switch is congested, and assign a minimum bandwidth guarantee to ensure scheduling fairness. These mechanisms operate at the CoS queue level; that is, the minimum bandwidth guarantees are made across all configured CoS queues.

NOTE: Minimum bandwidth guarantees and scheduling mechanisms apply only when the switch is congested. When the switch is not congested, packets egress the switch as soon as they are received.

ETS provides a second level of scheduling for packets selected for transmission by the CoS scheduler. ETS operates at the traffic class group (TCG) level and supports sharing of bandwidth across TCGs, bandwidth assignment for each TCG, and queue discipline (drop behavior) for each TCG. Dell Networking N4000 Series switches support three TCGs internally, up to two of which may be configured as lossless.

When a packet arrives on an ingress port, it is forwarded to the appropriate CoS queue based upon the 802.1p mapping. The 802.1p mapping maps the user priority contained in the received VLAN Priority Tag to the traffic class or uses the default port assignment. A traffic class identifies a particular CoS queue. If the ingress port is configured to use 802.1p CoS mapping and the port is configured to trust the user priority value in the received frame, then the frame is forwarded to the appropriate CoS queue per its 802.1p user priority value.

At the first level of egress scheduling, each of the configured attributes of a CoS queue, namely scheduler algorithm, min-bandwidth and drop mechanism, are honored, and the packet is either dropped or forwarded to next level. Only frames selected by the first level scheduler are forwarded to the second level.

Strict priority traffic classes are serviced first in order of traffic class number. A strict priority traffic class is one that is configured as strict priority or has a traffic class group weight of 0 (unlimited bandwidth). Within a strict priority traffic class, frames are serviced (as determined by the traffic class number) in CoS queue priority order from the highest (6) to the lowest (0).

Each traffic class has an internal weight equal to the traffic class number plus one. After the strict priority traffic classes have been serviced, the remaining traffic classes are serviced according to their internal weight. For example, if CoS queues 0, 1, and 2 have an equal offered load toward a congested output port, CoS queue 2 will receive 3/6 of the bandwidth, CoS queue 1 will receive 2/6 of the bandwidth, and CoS queue 0 will receive 1/6 of the bandwidth.

The minimum bandwidth setting can be used to override the strict priority and weighted settings. The highest numbered strict priority queue will receive no more bandwidth than 100 percent minus the sum of the minimum bandwidth percentages assigned to the other queues. If used, it is recommended that minimum bandwidth percentages only be set high enough to ensure a minimum level of service for any queue; i.e., the sum of the minimum bandwidth percentages is a fraction of 100%. This ensures that the system can respond to bursts in traffic. Setting the minimum bandwidths such that they sum to 100% effectively sets the scheduler such that no queue can exceed its minimum bandwidth percentage when congestion occurs.

NOTE: CoS queue 7 is reserved for internal traffic. Non-strict priority CoS queues are serviced with WDRR scheduling using the bandwidth available after strict priority traffic is serviced.

Each CoS queue in the first level scheduler is mapped to one of the three traffic class groups in the second level scheduler. There, frames are serviced using the TCG configuration. The minimum bandwidth guarantee is first calculated across the TCGs. Strict priority TCGs are scheduled first but have their bandwidth reduced by the minimum bandwidth guarantees configured on other TCGs. Strict priority TCGs are scheduled from highest numbered TCG to lowest. When all TCGs have met their minimum bandwidth limits (or the queues are empty), TCGs that have not met their maximum bandwidth limit are scheduled. Once the limits for a TCG are satisfied (maximum bandwidth, no frames available for transmission, etc.), the scheduler moves to the next TCG.

If no minimum or maximum bandwidth limits are configured, TCGs are serviced by the second-level scheduler using the configured TCG weights to define the relative bandwidth allocation among the TCGs. When an egress port is congested, packets are selected for discard using the configured tail-drop or WRED discipline. Minimum TCG bandwidth, maximum TCG bandwidth, and TCG weights are metered to within approximately 3% of the link bandwidth.

In the case that all TCGs are configured as strict priority, inter-TCG scheduling reverts to Weighted Deficit Round Robin scheduling based on the configured TCG weights.

Commands

This section provides information about the commands you use to manually configure and monitor ETS. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support. On Dell Networking N4000 Series switches, the following steps are not required if using the DCBX protocol to obtain ETS configuration from an auto-configuration source.

For information on configuring iSCSI and DCBX, refer to "iSCSI Optimization " on page 573.

Command	Purpose
classofservice traffic-class-group	Maps the internal Traffic Class to an internal Traffic Class Group (TCG). The Traffic Class can range from 0-6, although the actual number of available traffic classes depends on the platform.
traffic-class-group max-bandwidth	Specifies the maximum transmission bandwidth limit for each TCG as a percentage of the interface rate. Also known as rate shaping, this has the effect of smoothing temporary traffic bursts over time so that the transmitted traffic rate is bound.
traffic-class-group min-bandwidth	Specifies the minimum transmission bandwidth guaranteed for each TCG before processing frames from lower numbered TCGs on an interface.
traffic-class-group strict	Activates the strict priority scheduler mode for each specified TCG.
traffic-class-group weight	Specifies the scheduling weight for each TCG. The scheduler attempts to balance the traffic selected for transmission from the TCGs such that, when the switch is congested, traffic is selected from the round robin configured TCGs in proportion to their weights.
show classofservice traffic-class-group	Displays the Traffic Class to Traffic Class Group mapping.
show interfaces traffic-class-group	Displays the Traffic Class to Traffic Class Group mapping

ETS Configuration Example

This example configures four classes of traffic:

Best effort traffic	CoS Queue 0 for untagged and VLAN-tagged frames with VPTs 0, 1, and 2
Lossless iSCSI traffic	CoS Queues 1 & 2 for VLAN tagged frames with VPTs 3 & 4 respectively
Expedited traffic	CoS Queue 3 on VLAN tagged frames with VPTs 5, 6, and 7

1. Enable Trust Mode on an Interface

The following command enables the use of the 802.1p priority of the incoming packet. It may be configured on a single interface, a range of interfaces or all interfaces. By default, ports are configured to trust the incoming user priority.

```
console(config-if-Te1/0/1)#classofservice trust dot1p
```

To show the configured trust mode on an interface, use the following command:

```
console#show classofservice trust tengigabitethernet 1/0/1
Class of Service Trust Mode: Dot1P
```

2. Map 802.1p Priority to CoS Queues

This step maps 802.1p user priorities to the CoS queues. This mapping places the incoming packet in the selected CoS queue based on the 802.1p user priority. It may be configured on a single interface, a range of interfaces, or all interfaces.

To ensure lossless behavior, the 802.1p user priority must be mapped one to one to a CoS queue for the lossless priorities. Up to two lossless priorities may be configured on Dell Networking N4000 Series switches. CoS queue 7 is reserved by the system and is not assignable. It is generally recommended that the administrator utilize CoS queues 0-3 as CoS queues 4-6 may be utilized by the system for other types of system traffic, e.g. routing protocol PDU handling. Frames with different user priorities assigned to a single CoS queue receive equal treatment.

This example maps user priorities 0, 1, and 2 to CoS queue 0 (background or best effort traffic), user priorities 3 and 4 to CoS queues 1 and 2 (iSCSI traffic), and all other priorities to CoS queue 2 (low latency and network control traffic).

```
console(config-if-Te1/0/2)#classofservice dot1p-mapping 0 0
console(config-if-Te1/0/2)#classofservice dot1p-mapping 1 0
console(config-if-Te1/0/2)#classofservice dot1p-mapping 2 0
console(config-if-Te1/0/2)#classofservice dot1p-mapping 3 1
console(config-if-Te1/0/2)#classofservice dot1p-mapping 4 2
console(config-if-Te1/0/2)#classofservice dot1p-mapping 5 3
console(config-if-Te1/0/2)#classofservice dot1p-mapping 6 3
console(config-if-Te1/0/2)#classofservice dot1p-mapping 7 3
```

To show the 802.1p priority for an interface, use the following command:

```
console#show classofservice dot1p-mapping tengigabitethernet 1/0/1

 User Priority    Traffic Class
 -------------    -------------
       0                0
       1                0
       2                0
       3                1
       4                2
       5                3
       6                3
       7                3
```

3. Configure Minimum Bandwidth on the CoS Queues (optional)

This step configures the minimum bandwidth guarantee for each CoS queue when the switch is congested. This command affects the primary packet scheduler such that a CoS queue is guaranteed a minimum amount of scheduler capacity. It does not reserve packet buffers to a CoS queue. If the switch is not congested, this setting has no effect. The min-bandwidth setting guarantees that any particular CoS queue is serviced often enough to ensure that the offered load can achieve the minimum transfer rate. The scheduler bandwidth is measured internally as bytes transferred per second. The minimum bandwidth setting is enforced on the egress queue; it does not rate-limit incoming frames. The minimum bandwidth setting is configured as a percentage of the total packet scheduler bandwidth and must be less than or equal to 100%. It may be configured on a single interface, a range of interfaces, or all interfaces.

The minimum bandwidth setting on the CoS queues comes in to effect only when there is congestion among the CoS queues belonging to a single TCG. This is an optional setting and is not generally required, as the secondary scheduler has the capability of guaranteeing minimum bandwidth for a TCG. If this value is set, the sum of the individual minimum bandwidths of the CoS queues belonging to any TCG should be set to the same value as the minimum bandwidth value of the TCG in the secondary scheduler, or no minimum bandwidth value should be configured for the TCG.

CAUTION: Sharing of bandwidth among CoS Queues is disabled if the sum of the minimum bandwidth settings equals 100%.

```
console(config-if-Te1/0/1)#cos-queue min-bandwidth 20 35 35 10 0 0 0
```

4. Configure the Scheduler Mode for the CoS Queues

This step enables strict priority scheduling on one or more CoS queues (traffic classes). Strict priority scheduling ensures that packets assigned to a higher CoS queue number are serviced before packets assigned to lower CoS queue numbers. The strict priority setting does not restrict frame ingress. Use the 'no' command to disable strict priority mode scheduling.

It is recommended that all queues enabled for strict priority scheduling be assigned to a single strict-priority-enabled TCG other than TCG0.

The following example sets CoS queue (traffic class) number 3 to be serviced with strict priority:

```
console(config-if-Te1/0/1)#cos-queue strict 3
```

By default, these are packets tagged with IEEE 802.1p user priority 6 and 7.

To show the minimum bandwidth and scheduler modes for CoS queues, use the following command:

```
console#show interfaces cos-queue tengigabitethernet 1/0/1

Interface....................................... Te1/0/1
Interface Shaping Rate.......................... 0 kbps
WRED Decay Exponent............................. 9

Queue Id   Min. Bandwidth   Scheduler Type   Queue Management Type
--------   --------------   --------------   ---------------------
0          20               Weighted         Tail Drop
1          35               Weighted         Tail Drop
2          35               Weighted         Tail Drop
3          10               Strict           Tail Drop
```

```
4          0                Weighted        Tail Drop
5          0                Weighted        Tail Drop
6          0                Weighted        Tail Drop
```

5. Map the CoS Queues to TCGs

In this step, CoS queues are mapped to Traffic Class Groups (TCGs). Since TCGs are serviced from highest numbered TCG to lowest, higher priority traffic should be assigned to higher numbered TCGs. In general, strict priority traffic (typically control plane or low bandwidth, low latency traffic) is assigned the highest numbered TCG. It is recommended that WDRR queues be assigned to TCG0.

The mapping may be configured on a single interface, a range of interfaces, or all the interfaces.

It is required that TCGs always be assigned in order from 0 to 2. It is further recommended that the operator always utilize consecutive TCGs starting with TCG 0; for example, use TCG 0 and 1 vs. TCG 0 and 2 only.

This example assigns best effort traffic to TCG 0, lossless traffic to TCG 1, and expedited service traffic to TCG 2.

```
console(config-if-Te1/0/1)#classofservice traffic-class-group 0 0
console(config-if-Te1/0/1)#classofservice traffic-class-group 1 1
console(config-if-Te1/0/1)#classofservice traffic-class-group 2 1
console(config-if-Te1/0/1)#classofservice traffic-class-group 3 2
```

To show the CoS queue to TCG mapping, use the following command:

```
console#show classofservice traffic-class-group tengigabitethernet
1/0/1
 Traffic Class    Traffic Class Group
 -------------    ------------------
       0                0
       1                1
       2                1
       3                2
       4                0
       5                0
       6                0
```

6. Set the Weight for Each TCG

This step configures the scheduling weight for each TCG. The weight percentage configures the scheduler to ensure that each TCG is nominally serviced in the ratio specified.

Each WDRR TCG should be assigned a nonzero weight. Weights may be configured on a single interface, a range of interfaces, or all interfaces, and must sum to 100%. It is recommended that strict priority TCGs be assigned a weight of 0%, since they are processed first and ignore the configured TCG weight.

It is recommended that the sum of minimum bandwidth percentages configured on the CoS queues mapped to any TCG be less than or equal to that of the weight percentage configured for the TCG, so that packets are not dropped due to the congestion in the TCG.

In example below, the TCG0 and TCG1 are allocated 30% and 70% of the bandwidth remaining after servicing TCG2 (strict priority) traffic. TCG2 traffic is handled with strict priority but can only consume up to 100% minus the sum of the minimum bandwidths of TCG0 and TCG1 (60%).

```
console(config-if-Te1/0/1)#traffic-class-group weight 30 70 0
```

7. Set the Minimum Bandwidth for Each TCG (optional)

In this step, the minimum bandwidth is configured for each TCG while in congestion. The minimum bandwidth setting ensures that a TCG is serviced during congestion even if a higher priority TCG has an offered load greater than 100% minus the minimum bandwidth assigned to the other TCGs.

The minimum bandwidth is enforced on the egress TCG regardless of the scheduling mode and does not directly affect incoming traffic. The minimum bandwidth for a TCG is configured as a percentage of the total bandwidth, and the configured minimum bandwidths may sum to less than 100%. The sum may not exceed 100%. Minimum bandwidth may be configured on a single interface, a range of interfaces, or all interfaces. It is recommended that the minimum bandwidth for a TCG be configured to be less than or equal to the weight of the TCG.

CAUTION: Sharing of bandwidth among TCGs is disabled if the sum of the minimum bandwidth settings equals 100%.

```
console(config-if-Te1/0/1)#traffic-class-group min-bandwidth 10 50 0
```

8. Set the Maximum Bandwidth for each TCG (optional)

In this step, the maximum bandwidth for each TCG is configured. The bandwidth allowed for the TCG will never exceed the maximum bandwidth configured, even if configured for strict priority. This limit is configured as a percentage of the total bandwidth and is used to shape egress traffic bursts to no greater than the configured value. The maximum bandwidth may be configured on a single interface, range of interfaces or all interfaces. When configured to be 0, unlimited bandwidth is allowed on the TCG.

It is recommended that the maximum bandwidth be configured to be greater than the minimum bandwidth or the weight or be configured to 0 (unlimited burst size).

```
console(config-if-Te1/0/1)#traffic-class-group max-bandwidth 50 90 20
```

9. Set the Scheduler Modes for the TCGs

This step enables strict priority scheduling on TCGs. Strict priority scheduling on multiple TCGs prioritizes traffic from the highest numbered TCG for transmission first. Strict priority scheduling on a single TCG selects that TCG for transmission before the WDRR TCGs. Use the 'no' command to disable strict priority scheduling. It is recommended that all CoS queues enabled for strict priority scheduling be assigned to a single TCG other than TCG0. This scheme allows a larger number of priorities to be configured as strict priorities.

```
console(config-if-Te1/0/1)#traffic-class-group strict 2
```

To show the weight, minimum bandwidth, maximum bandwidth, and scheduler modes of the TCGs, use the following command:

```
console#show interfaces traffic-class-group tengigabitethernet
1/0/1

Interface...................................... Te1/0/1

Traffic Class   Min.       Max
Group          Bandwidth  Bandwidth    Weight   Scheduler Type
------------   ---------  ------------ ------   --------------
0              10         50           30       Weighted Round Robin
1              50         90           70       Weighted Round Robin
2              0          20           0        Strict
```

ETS Theory of Operation

First Level of Scheduling

To understand the first level of scheduling, consider Table 29-1. Assume that we have eight ingress ports, each one receiving line rate traffic with one 802.1p priority each. The table shows the mapping of 802.1p priorities to the cos-queues, the min-bandwidth settings, and scheduler modes.

Table 29-3. First Level of Scheduling

802.1p Priority	CoS Queue	CoS Min-bandwidth	Scheduler	TCG	Input to TCG
0	0	10	WRR	0	100% of pri0
1	1	10	WRR	1	10% each of
2	2	10	Strict	1	pri1, pri3 and
3	3	10	WRR	1	80% of pri2
4	4	15	WRR	2	25% each
5	5	15	WRR	2	of pri4, pri5
6	6	30	WRR	2	and 50% of
7	6	–	–	–	pri6/pri7

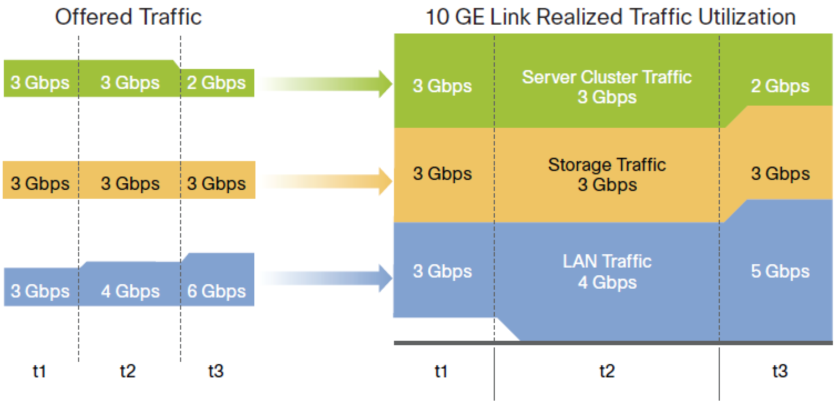
In this scenario, the input to the TCGs is as follows:

- All of the pri0 traffic reaches the TCG0. This is because there is no other priority contending for the same TCG. The minimum bandwidth setting on the CoS queue does not have any effect.
- TCG1 would receive 10% each of pri1 and pri3 and 80% of pri2. Even though strict mode is enabled for pri2, the minimum bandwidth of pri1 and pri3 is first honored before applying strict mode on pri2.
- TCG2 receives 25% each of pri4 and pri5 traffic and the other 50% can be of pri6 or pri7. This is based on the minimum bandwidth ratio, which is 15:15:30, converted to 1:1:2.

Second Level of Scheduling

To consolidate different traffic classes within different traffic types in a typical DCB environment, ETS provides an operational model for prioritization and bandwidth allocation for traffic. Figure 29-3 illustrates a typical example that consolidates three traffic types on a single 10GE link. For consolidation to be effective all traffic types must be serviced according to their requirements. For instance, SAN traffic must be guaranteed to be lossless, and the server cluster traffic (IPC) should be guaranteed a sufficiently high priority to meet the requirement of low latency.

Figure 29-3. Converged Link on the DCB Environment



In this example, to ensure that the server cluster traffic has low latency, it may be assigned to a TCG—say, TCG0—and a strict mode of scheduling is enabled on this group (weight set to 0%). SAN traffic can be assigned to TCG1 and LAN to TCG2. The TCG1 and TCG2 can be set to a weight of 50% each. With these configurations, Figure 29-3 illustrates how the load is managed at different times.

When the offered load is 3 Gbps for each traffic type at time t1, the allocated bandwidth for TCG1 and TCG2 is 3.5 Gbps each. All traffic types, including IPC, SAN, and LAN, are allowed to be transmitted, since the offered load is under the allocated bandwidth.

At time t2, a burst of LAN traffic is incoming at the rate of 4 Gbps, this burst is allowed to borrow the unused 0.5 Gbps bandwidth from SAN TCG and transmitted since the offered load of SAN is only 3 Gbps.

At time t3, when the offered load of IPC falls to 2 Gbps and the bursty LAN traffic is at 6 Gbps, the available bandwidth for SAN and LAN is 4 Gbps each according to the TCG weights, which are set as 50% each. However, ETS allows the LAN traffic to borrow unused bandwidth from SAN traffic, and then 5 Gbps of LAN traffic is transmitted.

Recommendations and Notes

Lossless traffic must map user priorities one-to-one onto CoS queues. Lossless traffic must always be carried within a VLAN-tagged frame to differentiate it from untagged best-effort LAN control traffic, such as ARP and LLDP.

Strict-priority CoS queues may starve other CoS queues of traffic if the offered load of the strict-priority CoS queues equals or exceeds the available capacity of the egress interfaces. It is recommended that either a CoS-queue-level min-bandwidth setting be utilized to ensure a minimum amount of bandwidth is processed on the non-strict-priority queues if there is a possibility that the strict priority traffic is not limited in bandwidth by some other means.

It is recommended that the sum of the minimum bandwidth percentages allocated to the group of CoS queues mapped to a single TCG be less than or equal to the weight percentage, so that packets are not dropped due to the congestion in the TCG.

The maximum bandwidth for a TCG should always be configured to be greater than the minimum bandwidth or the weight of the TCG.

In the case where all TCGs are configured as strict priority, inter-TCG scheduling is Weighted Deficit Round Robin, using the configured weights.

Configuring CoS queues with a total maximum bandwidth that is less than the corresponding TCG weight will result in bandwidth being wasted in the secondary scheduler.

Configuring CoS queues with a total minimum bandwidth that is greater than the corresponding TCG weight will flood the secondary scheduler and result in the minimum bandwidth parameter effectively overriding the TCG weight setting.

Traffic is passed across stacking links using WDRR for all CoS queues. This will affect the observed behavior of ETS on egress ports scheduling traffic from over-subscribed stacking links.

The three supported traffic class groups support an industry standard configuration such that one traffic class group offers lossless service (PFC enabled using WRR); one traffic class group contains priorities which are best effort (PFC disabled using WRR); and one traffic class group offers an expedited service (utilizes strict priority). When auto configuration is enabled on ports using ETS, the traffic classes are combined according to the above. Internally, the DCBX TCG mapping is restricted to using three TCGs regardless of the number of TCGs advertised by the configuration source, with the strict priority traffic mapped onto TCG 2, lossless traffic mapped onto TCG1, and all other traffic mapped onto TCG0.

For ETS to be operational in a manually configured environment, the following minimum steps must be performed:

1 Configure the CoS queue to Traffic Class Group mapping for the egress ports.

2 Enable the appropriate scheduling algorithm for each TCG.

3 Configure the weight percentage for each TCG.

Variation on the Example Configuration

This example configures three classes of traffic and utilizes the secondary (ETS) scheduler only:

- Best effort traffic: CoS Queue 0 for untagged and VLAN tagged frames with VPTs: 0, 1, and 2
- Lossless iSCSI: CoS Queue 1, VLAN tagged frames with VPT equal to 3
- Expedited Traffic: CoS Queue 2, VLAN tagged frames with VPTs: 4, 5, 6, 7

```
console(config-if-Te1/0/2)#classofservice dot1p-mapping 0 0
console(config-if-Te1/0/2)#classofservice dot1p-mapping 1 0
console(config-if-Te1/0/2)#classofservice dot1p-mapping 2 0
console(config-if-Te1/0/2)#classofservice dot1p-mapping 3 1
console(config-if-Te1/0/2)#classofservice dot1p-mapping 4 2
console(config-if-Te1/0/2)#classofservice dot1p-mapping 5 2
console(config-if-Te1/0/2)#classofservice dot1p-mapping 6 2
console(config-if-Te1/0/2)#classofservice dot1p-mapping 7 2
console(config-if-Te1/0/1)#classofservice traffic-class-group 0 0
console(config-if-Te1/0/1)#classofservice traffic-class-group 1 1
```

```
console(config-if-Te1/0/1)#classofservice traffic-class-group 2 2
console(config-if-Te1/0/1)#traffic-class-group weight 30 70 0
console(config-if-Te1/0/1)#traffic-class-group strict 2
```

Dell Networking N4000 Series Operation

When DCBx is enabled on manually configured ports, it is not necessary for the ETS parameters to match, regardless of the version of DCBX negotiated or configured. Configuration mismatches are logged.

In auto configuration mode, ETS parameters from the configuration source are checked (Max TCs 3 and bandwidth equal to 100%) and if the system is capable of performing the configuration, it is accepted and propagated as received to the other auto-configuration ports. The ETS Recommendation TLVs are preferred over the ETS Configuration TLVs. Auto-configuration via DCBX overrides manually configured ETS parameters for auto-configured ports, however, manual configuration is restored should the port be placed back into the manual port role.

The ETS parameters received via DCBX are modified and applied to the system via the DCBX Mapping function as follows (references are to the 802.1Qaz parameters):

- Like traffic classes are combined up to the limits of the system; e.g., no more than 2 lossless CoS queues may be configured.
- The Priority Assignment Table (user priority to CoS queue mapping) is utilized by the system to map user priorities to the traffic classes (CoS queues).
- The TSA Assignment Table is converted to use 3 TCGs internally. Priorities with like characteristics are combined into TCGs, i.e. strict priority traffic is combined into a TCG, lossless traffic is combined into a TCG, etc. Generally, strict priority traffic is mapped onto TCG 2, lossless traffic is mapped onto TCG1 and best effort traffic is mapped onto TCG0.
- The bandwidths from the TC Bandwidth Table are summed based on the internal TCG mapping and are used to set the TCG weights. Other switches may assign bandwidth to strict priority queues. This bandwidth is counted in the sum to ETS 100% validation check, however, internally the weight for strict priority queues is ignored and they are configured for unlimited bandwidth. An implication of this assignment is that the percentage of bandwidth that may be consumed by a WDRR TCG after

processing strict priority traffic is skewed to be the bandwidth of the individual TCG divided by the sum of the weights of all WDRR configured TCGs.

The administrator may configure other parameters to work in conjunction with the received DCBX configuration, e.g. min-bandwidth per CoS queue and minimum or maximum bandwidth per TCG.

30

MAC Addressing and Forwarding

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

Dell Networking N-Series switches implement a MAC Learning Bridge in compliance with IEEE 802.1Q. The N-Series switches implement independent VLAN learning (IVL). Dynamically learned MAC addresses are used to filter the set of ports on which a frame is forwarded within a VLAN; that is, the destination MAC address and ingress VLAN for a frame entering the switch are looked up in the MAC address table and if, a match is found, the frame is forwarded out the matching port(s). If no match is found, the frame is flooded out all ports in the VLAN except for the ingress port.

This chapter describes the layer-2 MAC address table the switch uses to forward L2 frames between ports.

The topics covered in this chapter include:

- MAC Address Table Overview
- Default MAC Address Table Values
- Managing the MAC Address Table (Web)
- Managing the MAC Address Table (CLI)

MAC Address Table Overview

The MAC address table keeps track of the MAC addresses that are associated with each port to allow the switch to forward unicast traffic through the appropriate port. This table is sometimes called the bridge table or the forwarding database.

How Is the Address Table Populated?

The MAC address table can contain two types of addresses:

- Static: The address has been manually configured and does not age out.
- Dynamic: The address has been automatically learned by the switch and will age out if no frames with the learned MAC address (and VLAN) have been forwarded by the switch during the aging time interval.

Static addresses are configured by the administrator and added to the table. Dynamic addresses are learned by examining information in the Ethernet frame.

When a frame arrives on a port, the switch looks at the frame header to learn the source MAC address of the frame, then adds the address, VLAN ID, and the ingress port to the MAC address table. The address table is constantly updated as new addresses are learned, and unused addresses age out.

A frame that has a destination MAC address that matches an entry in the table is forwarded immediately to the associated port(s) within the same VLAN.

What Information Is in the MAC Address Table?

Each entry in the address table, whether it is static or dynamic, includes the MAC address, the VLAN ID associated with the MAC address, and the interface on which the address was learned or configured.

Each port can maintain multiple MAC addresses, and a single MAC address can be associated with multiple VLANs.

How Is the MAC Address Table Maintained Across a Stack?

The MAC address table is synchronized across all stack members. When a member joins the stack, its previous MAC address table is overwritten by the table maintained by the stack.

Default MAC Address Table Values

Table 30-1 summarizes the default values for the MAC address table.

Table 30-1. MAC Address Table Defaults

Parameter	Default Value
Aging time	300 seconds
Dynamic addresses	Enabled (automatically learned)
Static addresses	None configured

Managing the MAC Address Table (Web)

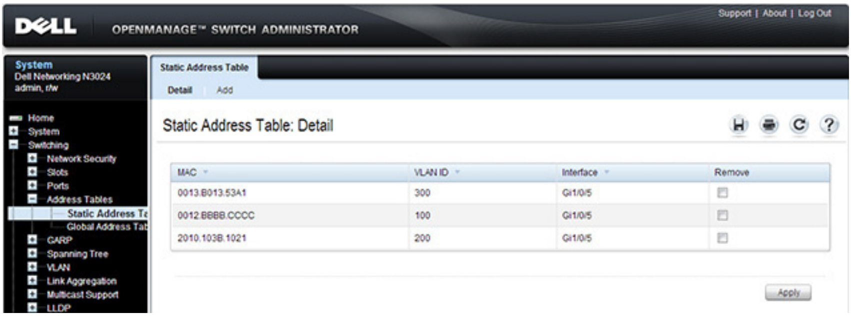
This section provides information about the OpenManage Switch Administrator pages to use to manage the MAC address table on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

Static Address Table

Use the **Static Address Table** page to view MAC addresses that have been manually added to the MAC address table and to configure static MAC addresses.

To display the **Static Address Table** page, click **Switching** → **Address Tables** → **Static Address Table** in the navigation panel.

Figure 30-1. Static MAC Address



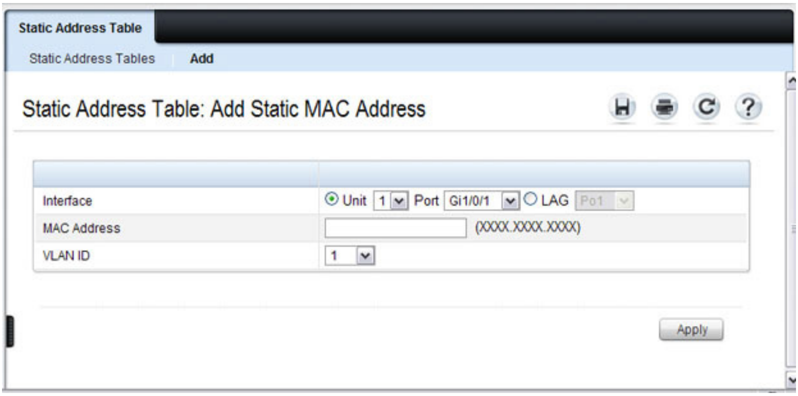
Adding a Static MAC Address

To add a static MAC address:

1. Open the **Static MAC Address** page.
2. Click **Add**.

   The **Add Static MAC Address** page displays.

Figure 30-2. Adding Static MAC Address



3 Select the interface to associate with the static address.

4 Specify the MAC address and an associated VLAN ID.

5 Click **Apply**.

The new static address is added to the **Static MAC Address Table**, and the device is updated.

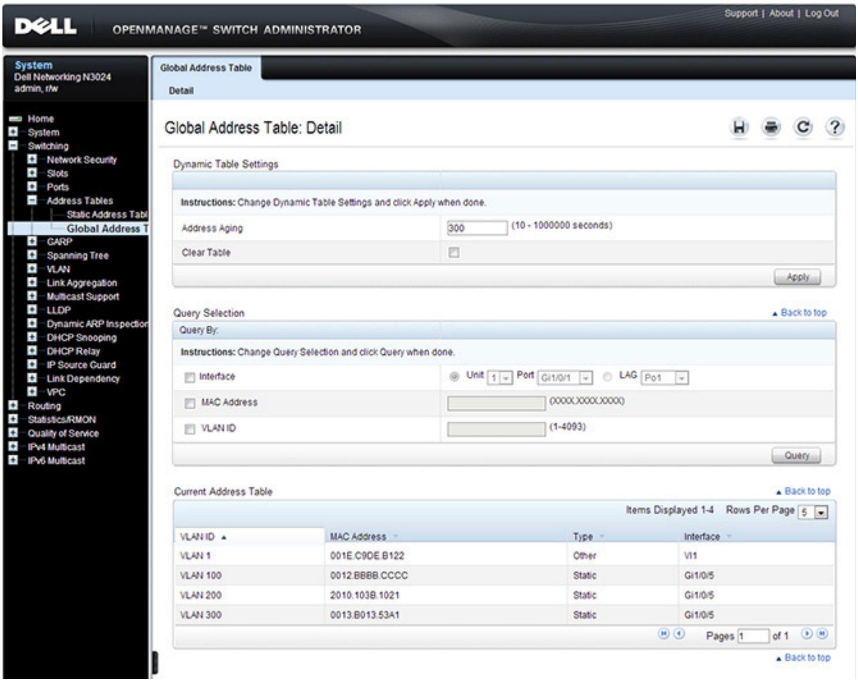
Global Address Table

The **Global Address Table** page contains fields for querying information in the dynamic address table, including the interface type, MAC addresses, VLAN, and table sorting key. Packets forwarded to an address stored in the address table are forwarded directly to those ports.

The **Global Address Table** also contains information about the aging time before a dynamic MAC address is removed from the table.

To display the **Global Address Table**, click **Switching → Address Tables → Global Address Table** in the navigation panel.

Figure 30-3. Global Address Table



Managing the MAC Address Table (CLI)

This section provides information about the commands you use to manage the MAC address table on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Managing the MAC Address Table

Beginning in Privileged EXEC mode, use the following commands to add a static MAC address to the table, control the aging time for dynamic addresses, and view entries in the MAC address table.

Command	Purpose
configure	Enter global configuration mode.
mac address-table static *mac-address* **vlan** *vlan-id* **interface** *interface*	Add a static MAC source address to the MAC address table. • *mac-address* — A valid MAC address in the format xxxx.xxxx.xxxx. • *vlan-id* —A valid VLAN. • *interface* — A valid port or LAG, including the interface type and number.
mac address-table aging-time {0 | 10-1000000}	Specify the number of seconds that must pass before an unused dynamically-learned MAC address is removed from the MAC address table. A value of 0 disables the aging time for the MAC address table.
exit	Exit to Privileged EXEC mode.
show mac address-table [**static | dynamic**]	View information about the entries in the MAC address table. Use the keywords **static** or **dynamic** to specify the address type to view. For dynamic entries, the **clear mac address-table** command can be used to remove entries from the table.
show mac address-table {**vlan** *vlan* | **interface** *interface* [**vlan** *vlan-id*]}	View information about the MAC addresses that have been configured or learned on the switch, a specific VLAN, or an interface (Ethernet port or LAG/port-channel).
show mac address-table count [{**vlan** *vlan-id* |**interface** *interface*}]	View information about the number of addresses that have been configured or learned on the switch, a specific VLAN, or an interface (Ethernet port or LAG/port-channel).

31

DHCP Server and Relay Settings

Dell Networking N2000, N3000, and N4000 Series Switches

This chapter describes how to configure the switch to dynamically assign network information to hosts by using the Dynamic Host Configuration Protocol (DHCP).

The topics covered in this chapter include:

DHCP Overview

DHCP is generally used between clients and servers for the purpose of assigning IP addresses, gateways, and other network settings such as DNS and SNTP server information.

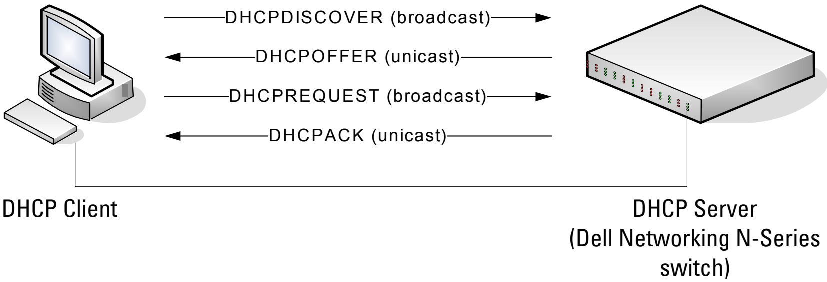
DHCP Snooping is a security feature that monitors DHCP messages between a DHCP client and DHCP server. It filters harmful DHCP messages and builds a bindings database of (MAC address, IP address, VLAN ID, port) tuples that are specified as authorized. DHCP snooping can be enabled globally and on specific VLANs. For information about DHCP Snooping, see "Snooping and Inspecting Traffic " on page 943.

Dell Networking N-Series switches support a DHCP client for obtaining the switch address from the network, an IPv4 DHCP server for serving IPv4 addresses to DHCP clients in the network, layer-2 and layer-3 DHCP relay for relaying IPv4 address assignments from network-based DHCP servers to clients in the same or different subnets, and DHCP snooping for protecting the switch and DHCP clients from certain security risks.

How Does DHCP Work?

When a host connects to the network, the host's DHCP client broadcasts a message requesting information from any DHCP server that receives the broadcast. One or more DHCP servers respond to the request. The response includes the requested information, such as the IP address, subnet mask, and default gateway IP address. The client accepts an offer from one of the servers, and the server sends an acknowledgment to the client to confirm the transaction.

Figure 31-1. Message Exchange Between DHCP Client and Server



The DHCP server maintains one or more set of IP addresses the and other configuration information available, by request, to DHCP clients. Each set of information is known as an address pool.

After a client leases an IP address from the DHCP server, the server adds an entry to its database. The entry is called a binding.

What are DHCP Options?

DHCP options are collections of data with type codes that indicate how the options should be used. Options can specify information that is required for the DHCP protocol, IP stack configuration parameters for the client, information allowing the client to rendezvous with DHCP servers, and so on.

When a client broadcasts a request for information, the request includes the option codes that correspond to the information the client wants the DHCP server to supply. The Web pages and CLI commands to configure DHCP server settings include many predefined options for the information that is most commonly requested by DHCP clients. For example, DHCP client discover requests typically include options for the IP address (option 50), subnet mask (option 1), default gateway (option 3), and DNS server (option 6). These options are predefined.

For options that are not predefined, the option code can be entered and the data type can be specified, along with the data that the switch should include in DHCP offers. RFC2132 specifies many of the DHCP options. Additional options are described in later RFCs.

How is DHCP Option 82 Used?

The Dell Networking Operating System supports insertion of DHCP Option 82 information into DHCP messages relayed to DHCP servers. The Dell Networking N-Series switch can be configured to insert either the Circuit ID or the Remote ID or both. When enabled, the Circuit ID contains the port identifier over which the DHCP request was received. The Remote ID is configurable by the administrator on a per-switch basis.

Consider a network with multiple DHCP servers, where the administrator wishes to serve addresses from a specific server based on the switch and port to which the user station is connected. User traffic is served on VLAN 10 or 20.

The administrator globally enables DHCP relay and configures DHCP relay on the end-user ports of each switch as follows:

```
console(config)#dhcp l2relay
console(config)#interface range gi1/0/1-24
console(config-if)#dhcp l2relay
console(config-if)#exit
```

Then, the administrator configures the remote-id and circuit-id:

```
console(config)#dhcp l2relay circuit-id vlan 10,20
console(config)#dhcp l2relay remote-id “Switch A” vlan 10,20
```

Finally, the administrator configures the uplink for DHCP relay and sets the interface to trust Option 82 information received on the interface:

```
console(config)#dhcp l2relay
console(config)#interface range te1/0/1
console(config-if)#dhcp l2relay
console(config-if)#dhcp l2relay trust
console(config-if)#exit
```

The administrator is using a Microsoft DHCP server. Microsoft DHCP servers do not have native support for DHCP Option 82, but it can be added using the DhcpServerCalloutEntry API to retrieve the information via the DhcpHandleOptionsHook configured on the switches. Adding Option 82 support enables choosing whether or not a particular DHCP server should respond to the DHCP request, and whether it should only respond to requests from a particular switch (as identified by the remote-id) and port (as identified by the circuit-id). For further information and an example, follow this web link:

http://blogs.technet.com/b/teamdhcp/archive/2009/07/06/dhcp-server-callout-api-usage.aspx

For Linux-based systems, which natively support option 82, a configuration to serve two private pools (Pool1 and Pool2) and one public pool of DHCP addresses based upon the remote-id and circuit-id might look like the following:

```
dhcpd.conf file:

class "Pool1" {
  match option agent.remote-id;
  match option agent.circuit-id;
}

subclass "Pool1" "Switch A" "Gi1/0/1";
subclass "Pool1" "Switch A" "Gi1/0/2";
subclass "Pool1" "Switch A" "Gi1/0/3";

class "Pool2" {
  match option agent.remote-id;
  match option agent.circuit-id;
}

subclass "Pool2" "Switch B" "Gi1/0/1";
subclass "Pool2" "Switch B" "Gi1/0/2";
subclass "Pool2" "Switch B" "Gi1/0/3";

shared-network Public {

  subnet 10.1.222.0 netmask 255.255.254.0 {
   pool {
    deny members of "Pool1";
    deny members of "Pool2";
    option routers 10.1.222.1;
```

```
    option subnet-mask 255.255.254.0;
    option domain-name-servers 10.1.218.3, 10.1.219.3;
    range dynamic-bootp 10.1.222.3 10.1.222.254;
    range dynamic-bootp 10.1.223.3 10.1.223.254;
    default-lease-time 21600;
    max-lease-time 43200;
   }
  }
  subnet 10.2.109.192 netmask 255.255.255.224 {
   pool {
    allow members of “Pool1”;
    range 10.2.109.194 10.2.109.222;
    option routers 10.2.109.193;
    option subnet-mask 255.255.255.224;
    option domain-name-servers 10.1.218.3,10.1.219.3;
    default-lease-time 21600;
    max-lease-time 43200;
  }
 }
 subnet 10.2.109.224 netmask 255.255.255.224 {
  pool {
    allow members of “Pool2”;
    range 10.2.109.226 10.2.109.254;
    option routers 10.2.109.225;
    option subnet-mask 255.255.255.224;
    option domain-name-servers 10.1.218.3,10.1.219.3;
    default-lease-time 21600;
    max-lease-time 43200;
   }
  }
 }
}
```

What Additional DHCP Features Does the Switch Support?

The switch software includes a DHCP client that can request network information from a DHCP server on the network during the initial system configuration process. For information about enabling the DHCP client, see "Setting the IP Address and Other Basic Network Information " on page 165.

If the switch is functioning as a layer-3 device, the layer-3 DHCP Relay Agent can relay DHCP messages between DHCP clients and DHCP servers that are located in different IP subnets.

The DHCP L2 relay feature permits L3 relay agent functionality in layer-2 switched networks. The switch supports L2 DHCP relay configuration on individual ports, link aggregation groups (LAGs) and VLANs. For information about layer-2 and layer-3 DHCP Relay, see "Layer-2 and Layer-3 Relay Features " on page 1157.

Default DHCP Server Values

By default, the DHCP server is disabled, and no address pools are configured. You must create at least one address pool and enable the DHCP server to allow the switch to dynamically assign network information to hosts with DHCP clients that broadcast requests.

The DHCP server can lease a maximum of 256 addresses.

The Dell Networking DHCP server does not offer infinite leases. The maximum lease time offered is 60 days, which corresponds to an infinite setting in the UI.

Configuring the DHCP Server (Web)

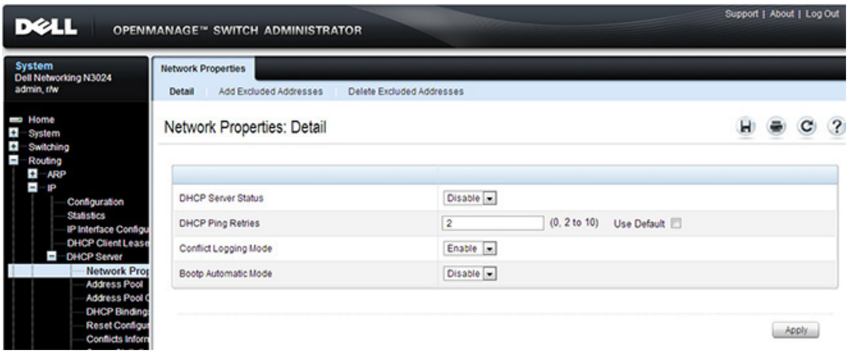
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring the DHCP server on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

DHCP Server Network Properties

Use the **Network Properties** page to define global DHCP server settings and to configure addresses that are not included in any address pools.

To display the **Network Properties** page, click **Routing** → **IP** → **DHCP Server** → **Network Properties** in the navigation panel.

Figure 31-2. DHCP Server Network Properties



Adding Excluded Addresses

To exclude an address:

1 Open the **Network Properties** page.

2 Click **Add Excluded Addresses** to display the **Add Excluded Addresses** page.

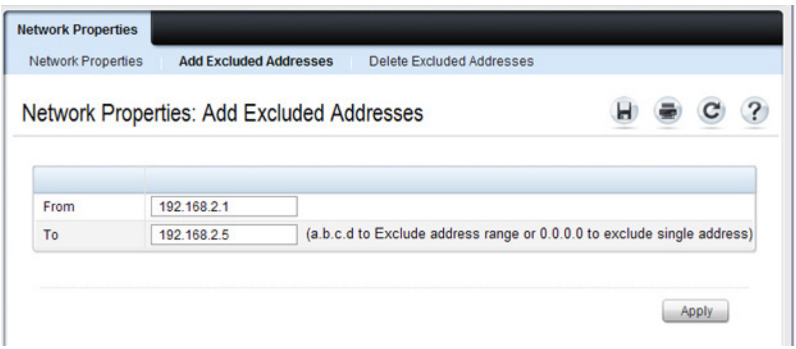
3 In the From field, enter the first IP address to exclude from any configured address pool.

4 If the address in the From field is the only address to exclude, or if the excluded addresses are non-contiguous, leave the To field as the default value of 0.0.0.0. Otherwise, enter the last IP address to excluded from a contiguous range of IP addresses.

   In Figure 31-3, the From field contains the IP address 192.168.2.1, and the To field contains the IP address 192.168.2.5. This means that the following IP addresses are not available for lease:

   - 192.168.2.1
   - 192.168.2.2
   - 192.168.2.3
   - 192.168.2.4
   - 192.168.2.5

Figure 31-3. Add Excluded Addresses



5 Click **Apply**.

Deleting Excluded Addresses

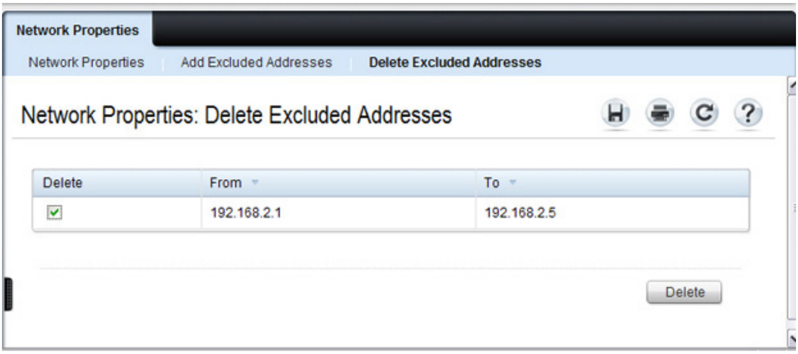
To remove an excluded address:

1 Open the **Network Properties** page.

2 Click **Delete Excluded Addresses** to display the **Delete Excluded Addresses** page.

3 Select the check box next to the address or address range to delete.

Figure 31-4. Delete Excluded Addresses



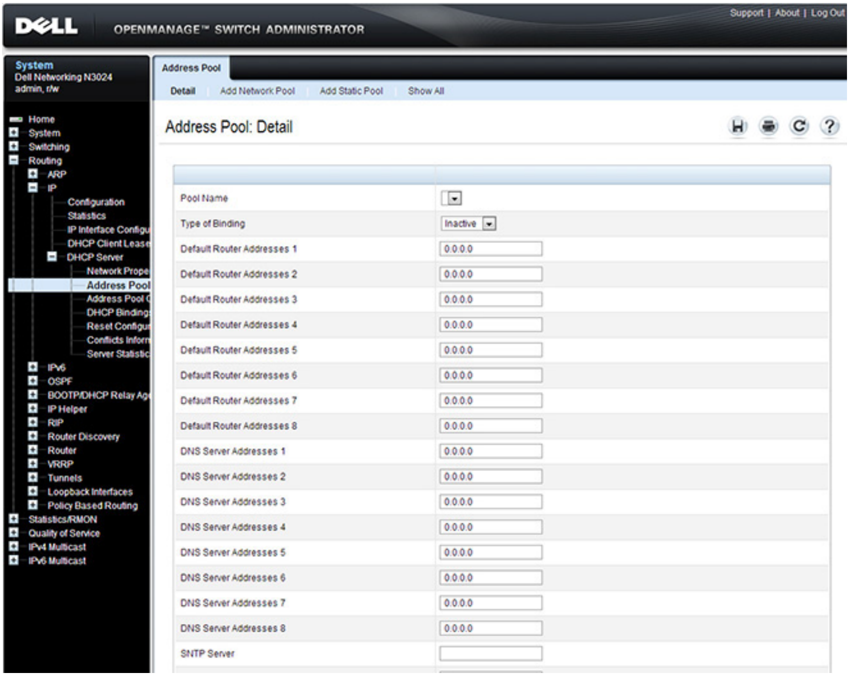
4 Click **Apply**.

Address Pool

Use the **Address Pool** page to create the pools of IP addresses and other network information that can be assigned by the server.

To display the **Address Pool** page, click **Routing** → **IP** → **DHCP Server** → **Address Pool** in the navigation panel.

Figure 31-5. Address Pool



Adding a Network Pool

To create and configure a network pool:

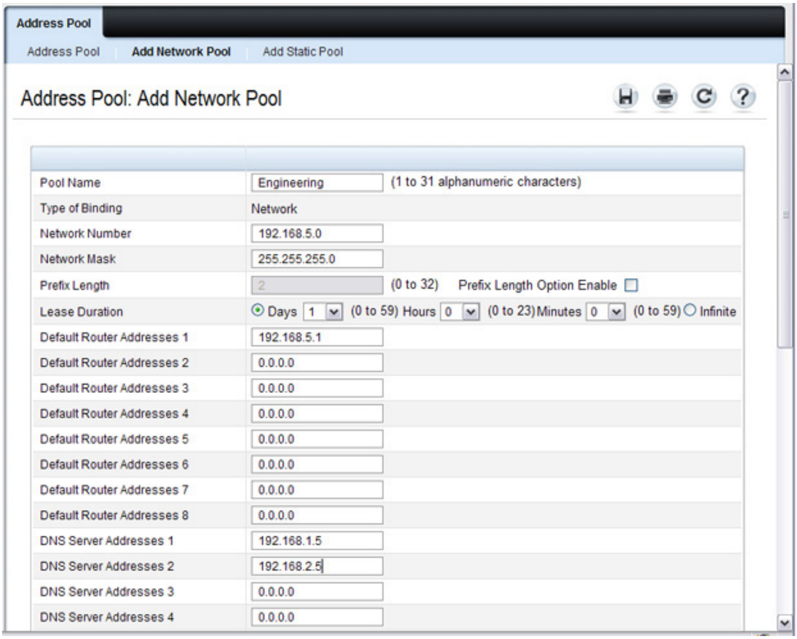
1 Open the **Address Pool** page.

2 Click **Add Network Pool** to display the **Add Network Pool** page.

3 Assign a name to the pool and complete the desired fields.

In Figure 31-6, the network pool name is Engineering, and the address pool contains all IP addresses in the 192.168.5.0 subnet, which means a client that receives an address from the DHCP server might lease an address in the range of 192.168.5.1 to 192.168.5.254.

Figure 31-6. Add Network Pool



The Engineering pool also configures clients to use 192.168.5.1 as the default gateway IP address and 192.168.1.5 and 192.168.2.5 as the primary and secondary DNS servers.

NOTE: The IP address 192.168.5.1 should be added to the global list of excluded addresses so that it is not leased to a client.

4 Click **Apply**.

Adding a Static Pool

To create and configure a static pool of IP addresses:

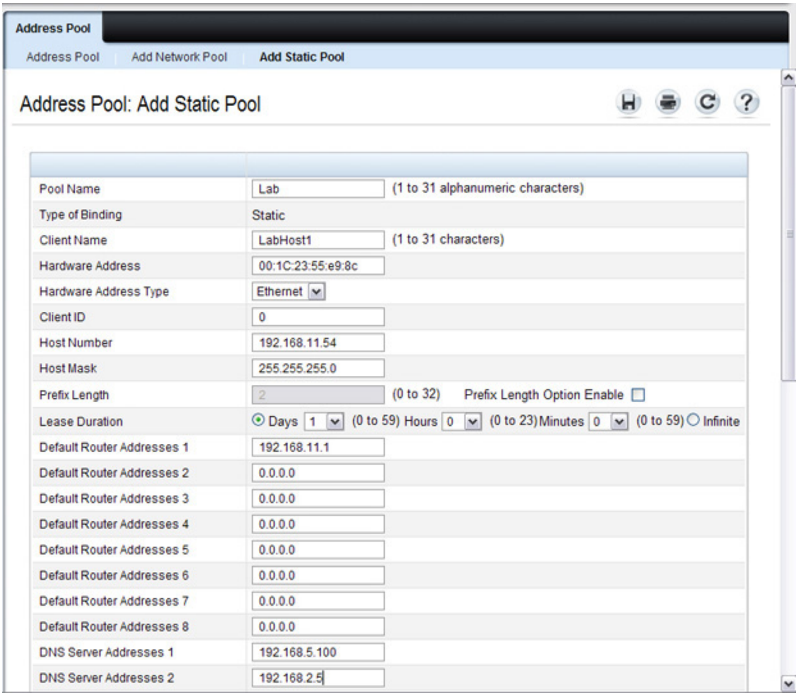
1 Open the **Address Pool** page.

2 Click **Add Static Pool** to display the **Add Static Pool** page.

3 Assign a name to the pool and complete the desired fields.

In Figure 31-7, the Static pool name is Lab, and the name of the client in the pool is LabHost1. The client's MAC address is mapped to the IP address 192.168.11.54, the default gateway is 192.168.11.1, and the DNS servers the client will use have IP addresses of 192.168.5.100 and 192.168.2.5.

Figure 31-7. Add Static Pool



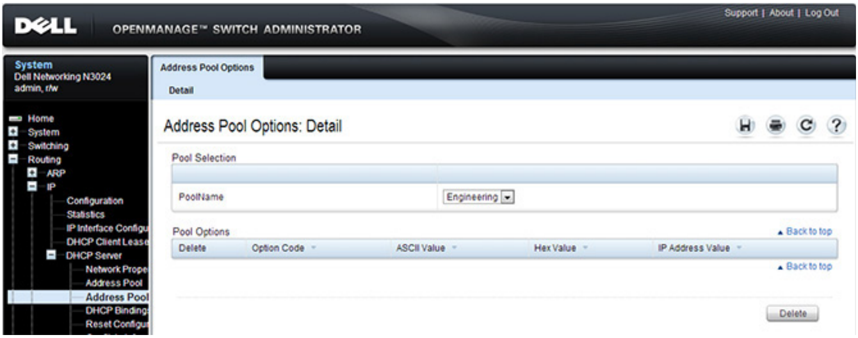
4 Click **Apply**.

Address Pool Options

Use the **Address Pool Options** page to view manually configured options. Options can be defined when an address pool is created or can be added to existing address pools.

To display the **Address Pool Options** page, click **Routing** → **IP** → **DHCP Server** → **Address Pool Options** in the navigation panel.

Figure 31-8. Address Pool Options



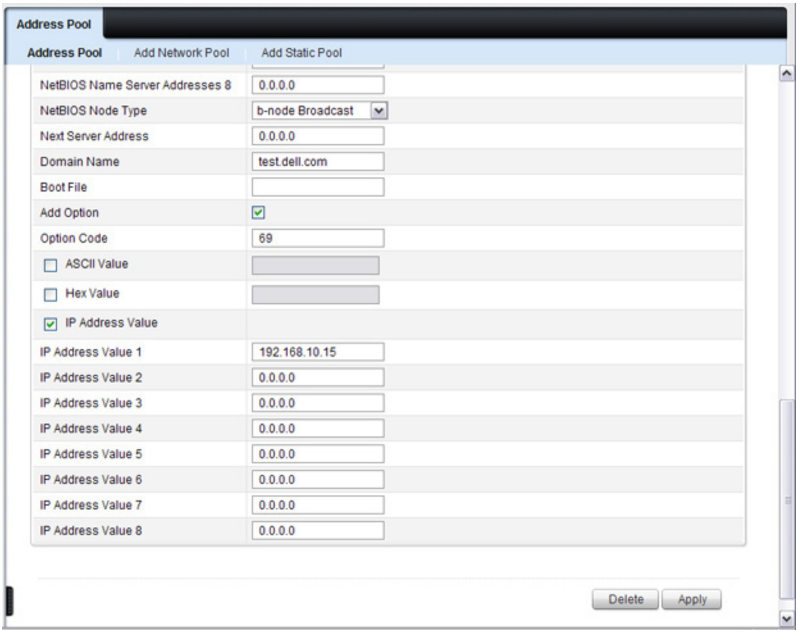
Defining DHCP Options

To configure DHCP options:

1. Open the **Address Pool** page.
2. Select the **Add Options** check box.
3. Select the check box that corresponds to the value type (ASCII, Hexadecimal, or IP address).
4. Specify the value(s) in the corresponding field.

   Figure 31-9 shows an example of adding the SMTP server IP address. The option code for the SMTP server is 69, and the IP address of the SMTP server is 192.168.10.15.

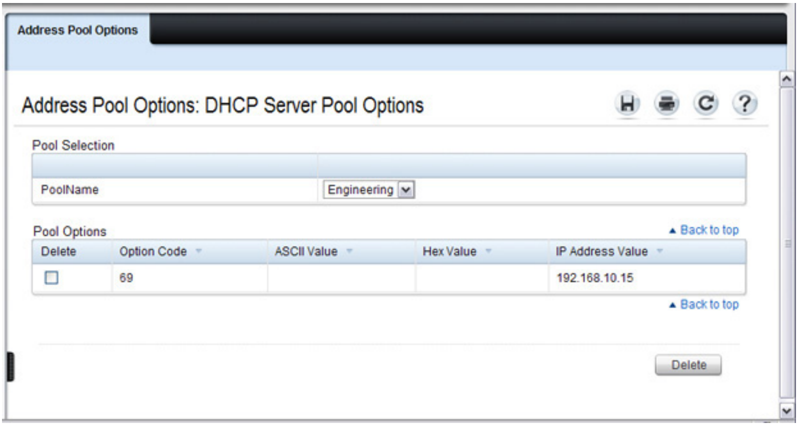
Figure 31-9. Add DHCP Option



5 Click **Apply**.

6 To verify that the option has been added to the address pool, open the **Address Pool Options** page.

Figure 31-10. View Address Pool Options

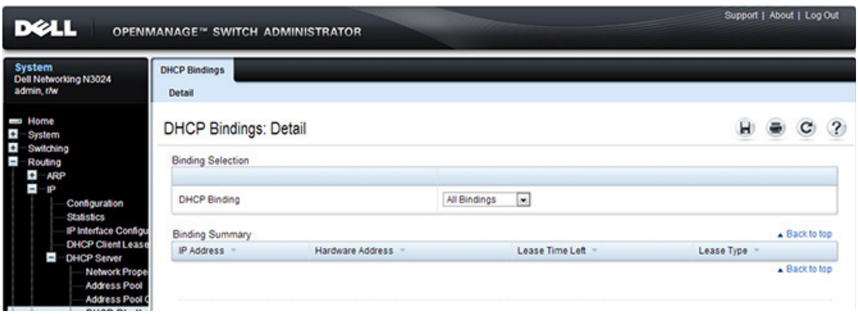


DHCP Bindings

Use the **DHCP Bindings** page to view information about the clients that have leased IP addresses from the DHCP server.

To display the **DHCP Bindings** page, click **Routing** → **IP** → **DHCP Server** → **DHCP Bindings** in the navigation panel.

Figure 31-11. DHCP Bindings

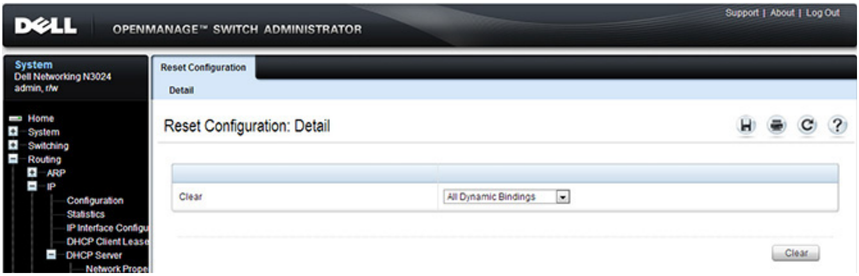


DHCP Server Reset Configuration

Use the **Reset Configuration** page to clear the client bindings for one or more clients. Bindings can also be reset for clients that have leased an IP address that is already in use on the network.

To display the **Reset Configuration** page, click **Routing** → **IP** → **DHCP Server** → **Reset Configuration** in the navigation panel.

Figure 31-12. Reset DHCP Bindings

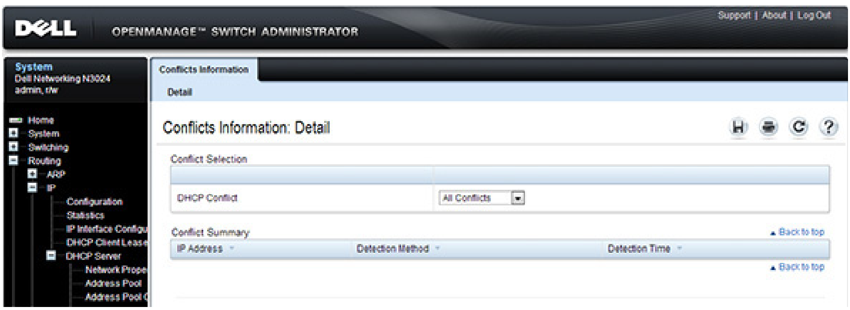


DHCP Server Conflicts Information

Use the **Conflicts Information** page to view information about clients that have leased an IP address that is already in use on the network.

To display the **Conflicts Information** page, click **Routing** → **IP** → **DHCP Server** → **Conflicts Information** in the navigation panel.

Figure 31-13. DHCP Server Conflicts Information

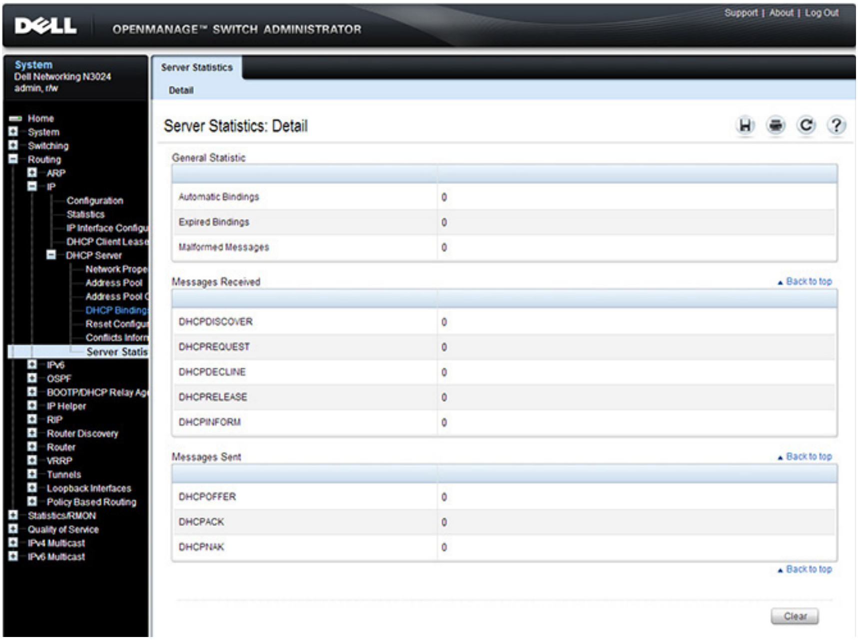


DHCP Server Statistics

Use the **Server Statistics** page to view general DHCP server statistics, messages received from DHCP clients, and messages sent to DHCP clients.

To display the **Server Statistics** page, click **Routing** → **IP** → **DHCP Server** → **Server Statistics** in the navigation panel.

Figure 31-14. DHCP Server Statistics



Configuring the DHCP Server (CLI)

This section provides information about the commands used for configuring and monitoring the DHCP server and address pools. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Global DHCP Server Settings

Beginning in Privileged EXEC mode, use the following commands to configure settings for the DHCP server.

Command	Purpose
configure	Enter Global Configuration mode.
service dhcp	Enable the DHCP server.
ip dhcp ping packets	Specify the number, in a range from 2-10, of packets a DHCP server sends to a pool address as part of a ping operation.
ip dhcp conflict logging	Enable conflict logging on DHCP server
ip dhcp bootp automatic	Enable the allocation of the addresses to the BootP client.
ip dhcp excluded-address *lowaddress* [*highaddress*]	Specify the IP addresses that a DHCP server should not assign to DHCP clients. A single IP address can be specified, or a contiguous range can be specified by using both the low-address and high-address variables.
exit	Exit to Privileged EXEC mode.
show ip dhcp global configuration	Verify the global DHCP server configuration.

Configuring a Dynamic Address Pool

Beginning in Privileged EXEC mode, use the following commands to create an address pool with network information that is dynamically assigned to hosts with DHCP clients that request the information.

Command	Purpose
configure	Enter Global Configuration mode.
ip dhcp pool *name*	Create a DHCP address pool and enters DHCP pool configuration mode.
network *network-ip* [*mask* | *prefixlength*]	Configure the subnet number and mask for a DHCP address pool. Clients requesting an IP address can be assigned any non-excluded IP address within this network.
lease {*days*[*hours*] [*minutes*] | **infinite**}	Specify the duration of the lease for an IP address that is assigned from a DHCP server to a DHCP client. • *days*— Days the lease is valid (Range 0–59, Default is 1). The hours and minutes can optionally be specified after the days. • **infinite** — 60-day lease. The Dell Networking DHCP server does not offer infinite leases. A setting of infinite corresponds to 60 days.
default-router *address1* [*address2....address8*]	Specify the list of default gateway IP addresses to be assigned to the DHCP client.
dns-server *address1* [*address2....address8*]	Specify the list of DNS server IP addresses to be assigned to the DHCP client.
domain-name *domain*	Specify the domain name for a DHCP client.
option *code* {**ascii** *string* | **hex** *string1* [*string2...string8*] | **ip** *address1* [*address2...address8*]}	Manually configure DHCP options.
CTRL + Z	Exit to Privileged EXEC mode.
show ip dhcp pool configuration {*name* | **all**}	View the settings for the specified address pool or for all configured address pools.

Configuring a Static Address Pool

Beginning in Privileged EXEC mode, use the following commands to create a static address pool and specify the network information for the pool. The network information configured in the static address pool is assigned only to the host with the hardware address or client identifier that matches the information configured in the static pool.

Command	Purpose
configure	Enter Global Configuration mode.
ip dhcp pool *name*	Create a DHCP address pool and enters DHCP pool configuration mode.
client-name *name*	Specify the DHCP client name.
hardware-address *mac* [*type*]	Specify the hardware address of the client in the static pool. • *mac*—MAC address of the hardware platform of the client consisting of 6 bytes in dotted hexadecimal format. • *type* — Indicates the protocol of the hardware platform. It is 1 for Ethernet and 6 for IEEE 802.
client-identifier *uniqueidentifier*	Specify the unique identifier for a DHCP client. The unique-identifier is a valid notation in hexadecimal format. In some systems, such as Microsoft DHCP clients, the client identifier is required instead of hardware addresses. The unique-identifier is a concatenation of the media type and the MAC address. For example, the Microsoft client identifier for Ethernet address c819.2488.f177 is 01c8.1924.88f1.77 where 01 represents the Ethernet media type.
host *address* [*mask* | *prefix-length*]	Specify the IP address and (optionally) network mask for a manual binding to a DHCP client.

Command	Purpose
lease {*days*[*hours*] [*minutes*] | **infinite**}	Specify the duration of the lease for an IP address that is assigned from a DHCP server to a DHCP client. • *days*— Days the lease is valid (Range 0–59, Default is 1). The hours and minutes can optionally be specified after the days. • **infinite** — 60 day lease. The Dell Networking DHCP server does not offer infinite leases. A setting of infinite corresponds to 60 days.
default-router *address1* [*address2....address8*]	Specify the list of default gateway IP addresses to be assigned to the DHCP client.
dns-server *address1* [*address2....address8*]	Specify the list of DNS server IP addresses to be assigned to the DHCP client.
domain-name *domain*	Specify the domain name for a DHCP client.
option *code* {**ascii** *string* | **hex** *string1* [*string2...string8*] | **ip** *address1* [*address2...address8*]}	Manually configure DHCP options.
CTRL + Z	Exit to Privileged EXEC mode.
show ip dhcp pool configuration {*name* | **all**}	View the settings for the specified address pool or for all configured address pools.

Monitoring DHCP Server Information

Beginning in Privileged EXEC mode, use the following commands to view bindings, conflicts, and statistics, and to clear the information.

Command	Purpose
show ip dhcp binding [*address*]	View the current binding information in the DHCP server database. Specify the IP address to view a specific binding.
clear ip dhcp binding {*address* | *}	Delete an automatic address binding from the DHCP server database. Use * to clear all bindings.
show ip dhcp conflict [*address*]	View the current binding conflicts in the DHCP server database. Specify the IP address to view a specific conflict.

Command	Purpose
clear ip dhcp conflict {*address* | *}	Clear an address conflict from the DHCP Server database. Use * to clear all conflicts.
show ip dhcp server statistics	View DHCP server statistics.
clear ip dhcp server statistics	Reset all DHCP server statistics to zero.

DHCP Server Configuration Examples

This section contains the following examples:

- Configuring a Dynamic Address Pool
- Configuring a Static Address Pool

Configuring a Dynamic Address Pool

The commands in this example create an address pool that dynamically assigns network information to hosts with DHCP clients that broadcast DHCP messages. The hosts are assigned an IP address from the 192.168.5.0 network. The IP addresses 192.168.5.1–192.168.5.20, and 192.168.5.100 are excluded from the address pool.

To configure the switch:

1. Enable the DHCP service and create an address pool named "Engineering", and then enter into DHCP pool configuration mode for the pool.

```
console#configure
console(config)#service dhcp
console(config)#ip dhcp pool Engineering
```

2. Specify the IP addresses that are available in the pool.

```
console(config-dhcp-pool)#network 192.168.5.0 255.255.255.0
```

3. Specify the IP address to use as the default gateway.

```
console(config-dhcp-pool)#default-router 192.168.5.1
```

4. Specify the primary and secondary DNS servers the hosts will use.

```
console(config-dhcp-pool)#dns-server 192.168.5.10
console(config-dhcp-pool)#dns-server 192.168.5.11
```

5 Specify the domain name to be assigned to clients that lease an address from this pool.

```
console(config-dhcp-pool)#domain-name engineering.dell.com
console(config-dhcp-pool)#exit
```

6 In Global Configuration mode, add the addresses to exclude from the pool. Clients will not be assigned these IP addresses.

```
console(config)#ip dhcp excluded-address 192.168.5.1
192.168.5.20
console(config)#ip dhcp excluded-address 192.168.5.100
```

7 Enable the DHCP server on the switch.

```
console(config)#service dhcp
console(config)#exit
```

8 View DHCP server settings.

```
console#show ip dhcp global configuration

Service DHCP...................Enable
Number of Ping Packets.........2
Excluded Address...............192.168.2.1 to 192.168.2.20
                               1.2.2.2 to 1.5.5.5
                               192.168.5.1 to 192.168.5.20
                               192.168.5.100 to 192.168.5.100
Conflict Logging...............Enable
Bootp Automatic................Disable
```

9 View information about all configured address pools.

```
console#show ip dhcp pool configuration all

Pool: Engineering
Pool Type......................... Network
Network........................... 192.168.5.0 255.255.255.0
Lease Time........................ 1 days 0 hrs 0 mins
DNS Servers....................... 192.168.5.11
Default Routers................... 192.168.5.1
Domain Name....................... engineering.dell.com
```

Configuring a Static Address Pool

The commands in this example create an address pool that assigns the address 192.168.2.10 to the host with a MAC address of 00:1C:23:55:E9:F3. When this hosts sends a DHCP message requesting network information, the switch will offer the information configured in this example, which includes a custom DHCP option to assign the SMTP server IP address.

To configure the switch:

1 Enable the DHCP service and create an address pool named “Tyler PC”, and then enter into DHCP pool configuration mode for the pool.

```
console#configure
console(config)#service dhcp
console(config)#ip dhcp pool “Tyler PC”
```

2 Specify the IP addresses that are available in the pool.

```
console(config-dhcp-pool)#hardware-address 00:1C:23:55:E9:F3
```

3 Specify the IP address and subnet mask to assign to the client.

```
console(config-dhcp-pool)#host 192.168.2.10 255.255.255.0
```

4 Specify the IP address to use as the default gateway.

```
console(config-dhcp-pool)#default-router 192.168.2.1
```

5 Specify the primary and secondary DNS servers the hosts will use.

```
console(config-dhcp-pool)#dns-server 192.168.2.100
console(config-dhcp-pool)#dns-server 192.168.5.101
```

6 Specify the domain name to be assigned to clients that lease an address from this pool.

```
console(config-dhcp-pool)#domain-name executive.dell.com
```

7 Specify the option that configures the SMTP server IP address to the host.

```
console(config-dhcp-pool)#option 69 ip 192.168.1.33
console(config-dhcp-pool)#exit
```

8 View information about the static address pool.

```
console#show ip dhcp pool configuration "Tyler PC"

Pool: Tyler PC
Pool Type.........................Static
Client Name.......................TylerPC
Hardware Address.................. 00:1c:23:55:e9:f3
Hardware Address Type.............ethernet
Host.............................. 192.168.2.10 255.255.255.0
```

```
Lease Time........................ 1 days 0 hrs 0 mins
DNS Servers....................... 192.168.2.101
Default Routers................... 192.168.2.1
Domain Name....................... executive.dell.com
Option............................ 69 ip 192.168.1.33
```

32

IP Routing

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure routing on the switch, including global routing settings, Address Resolution Protocol (ARP), router discovery, and static routes.

The topics covered in this chapter include:

- IP Routing Overview
- Default IP Routing Values
- IP Path MTU and Path MTU Discovery
- ARP Table
- Configuring IP Routing Features (Web)
- Configuring IP Routing Features (CLI)
- IP Routing Configuration Example

IP Routing Overview

The Dell Networking N-Series switches are multilayer switches that support static and dynamic routing. Table 32-1 describes some of the general routing features that can be configured on the switch.

Table 32-1. IP Routing Features

Feature	Description
ICMP message control	The type of ICMP messages that the switch responds to, as well as the rate limit and burst size, are configurable.
Default gateway	The switch supports a single default gateway. A manually configured default gateway is more preferable than a default gateway learned from a DHCP server.

Table 32-1. IP Routing Features (Continued)

Feature	Description
ARP table	The switch maintains an ARP table that maps an IP address to a MAC address. Static ARP entries can be created in the table and various ARP table settings can be managed, such as the aging time of dynamically-learned entries.
ICMP Router Discovery Protocol (IRDP)	Hosts can use IRDP to identify operational routers on the subnet. Routers periodically advertise their IP addresses. Hosts listen for these advertisements and discover the IP addresses of neighboring routers.
Routing table entries	The following route types can be configured in the routing table: • Default: The default route is the route the switch will use to send a packet if the routing table does not contain a longer matching prefix for the packet's destination. • Static: A static route is a route that you manually add to the routing table. • Static Reject: Packets that match a reject route are discarded instead of forwarded. The router may send an ICMP Destination Unreachable message.
Route preferences	The common routing table collects static, local, and dynamic (routing protocol) routes. When there is more than one route to the same destination prefix, the routing table selects the route with the best (lowest) route preference.

Default IP Routing Values

Table 32-2 shows the default values for the IP routing features this chapter describes.

Table 32-2. IP Routing Defaults

Parameter	Default Value
Default Time to Live	64
Routing Mode	Disabled globally and on each interface
ICMP Echo Replies	Enabled
ICMP Redirects	Enabled
ICMP Rate Limit Interval	1000 milliseconds
ICMP Rate Limit Burst Size	100
Maximum Next Hops	4
Global Default Gateway	None
Dynamic ARP Entry Age Time	1200 seconds
Automatic Renewal of Dynamic ARP Entries	Enabled
ARP Response Timeout	1 second
ARP Retries	4
Maximum Static ARP Entries	128
IRDP Advertise Mode	Disabled
IRDP Advertise Address	224.0.0.1
IRDP Maximum Advertise Interval	600 seconds
IRDP Minimum Advertise Interval	450 seconds
IRDP Advertise Lifetime	1800 seconds
IRDP Preference Level	0

Table 32-2. IP Routing Defaults (Continued)

Parameter	Default Value
Route Preference Values	Preference values are as follows: • Local—0 • Static—1 • OSPF Intra—110 • OSPF Inter—110 • OSPF External—110 • RIP—120

IP Path MTU and Path MTU Discovery

The IP stack maintains an IP MTU for each route in its routing table. Conceptually, the route's path MTU defaults to the IP MTU of the outgoing interface. The IP MTU of an interface is set automatically based upon the switch MTU. If the switch receives an ICMPv4 Fragmentation Needed or ICMPv6 Packet Too Big message, the IP stack sets the corresponding route's path MTU to the value in the ICMP message as long as it is less than the switch MTU minus the Ethernet frame header length.

RFC 1191 explains how a router can initiate IPv4 path MTU discovery. The basic idea is that the router sends IPv4 packets with the "don't fragment" bit set. The router initially assumes that the path MTU is equal to the IP MTU of the outgoing interface. If the packet is too big to reach its destination without fragmentation, a router in the path will return a packet too big message. The originator reduces its estimate of the path MTU and continues the process until packets reach the final destination.

Path MTU discovery is required for IPv6. In IPv6, only the originator is allowed to fragment. Any packet too large to reach its destination triggers a packet too big message, updating the IP stack's path MTU for the destination.

ARP Table

The router maintains an ARP table that associates a MAC address and outgoing port with an IP address and VLAN. The ARP table is dynamically updated with the host MAC address and outgoing port information. ARP entries are associated with the VLAN on which the IP address or route is known. The router broadcasts an ARP request in the associated VLAN for any unknown MAC address to which it needs to route packets. The router also refreshes an ARP entry by sending an ARP request before a dynamically learned ARP entry times out and updates the ARP table if a response is received. Host or VM movement within the same VLAN (layer-2 topology change) does not trigger an ARP refresh. Only if the ARP entry is timed out or the port associated with the ARP entry goes down does the ARP entry get refreshed.

If the traffic to a host is bidirectional, it will result in the host ARP entry pointing to the new port. Any gratuitous ARP request sent by a host or VM results in an ARP entry update (including a change in the MAC address and outgoing port).

Configuring IP Routing Features (Web)

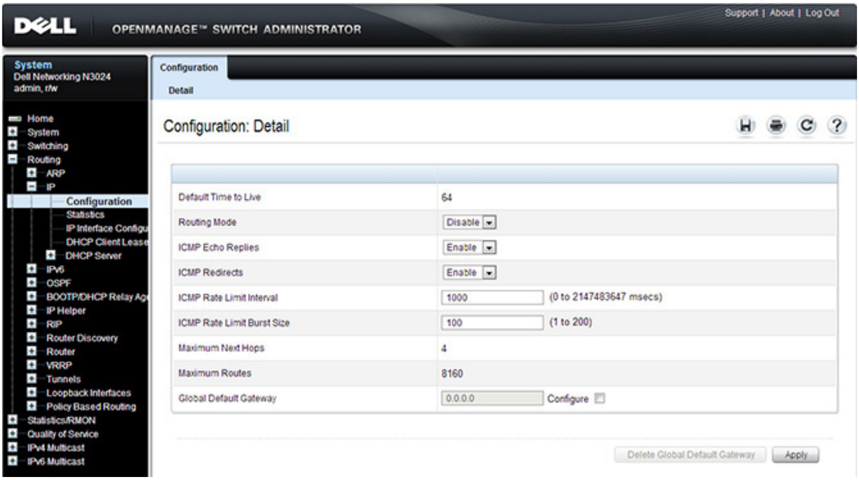
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring IPv4 routing features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

IP Configuration

Use the **Configuration** page to configure routing parameters for the switch as opposed to an interface. The IP configuration settings allow you to enable or disable the generation of various types of ICMP messages.

To display the page, click **Routing** → **IP** → **Configuration** in the navigation panel.

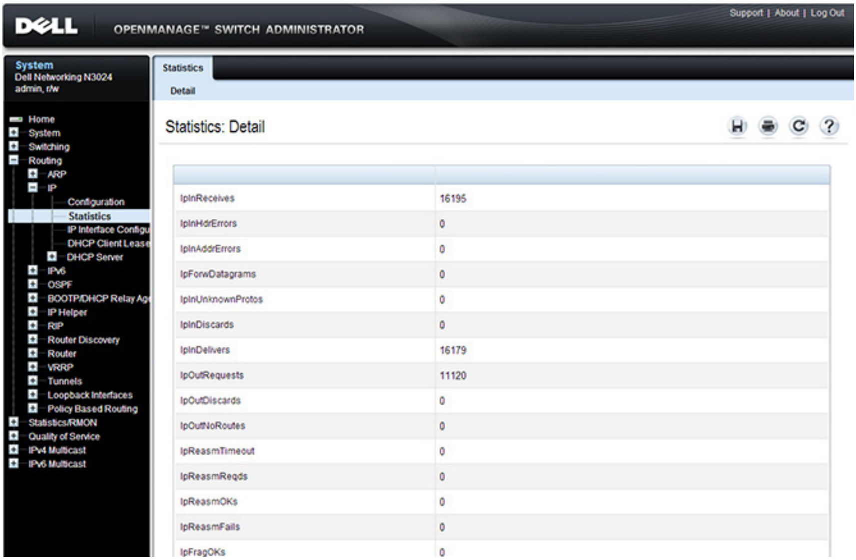
Figure 32-1. IP Configuration



IP Statistics

The IP statistics reported on the **Statistics** page are as specified in RFC 1213. To display the page, click **Routing** → **IP** → **Statistics** in the navigation panel.

Figure 32-2. IP Statistics

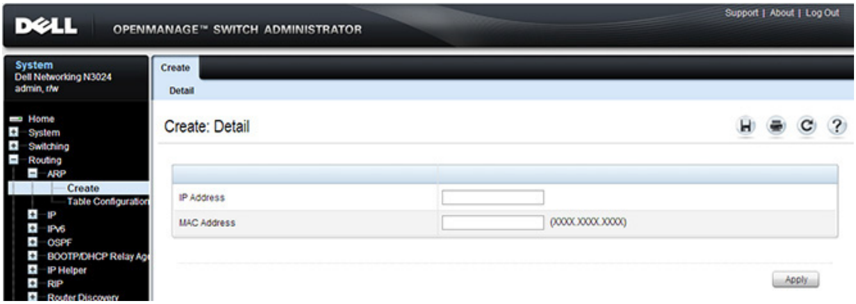


ARP Create

Use the **Create** page to add a static ARP entry to the Address Resolution Protocol table.

To display the page, click **Routing** → **ARP** → **Create** in the navigation panel.

Figure 32-3. ARP Create

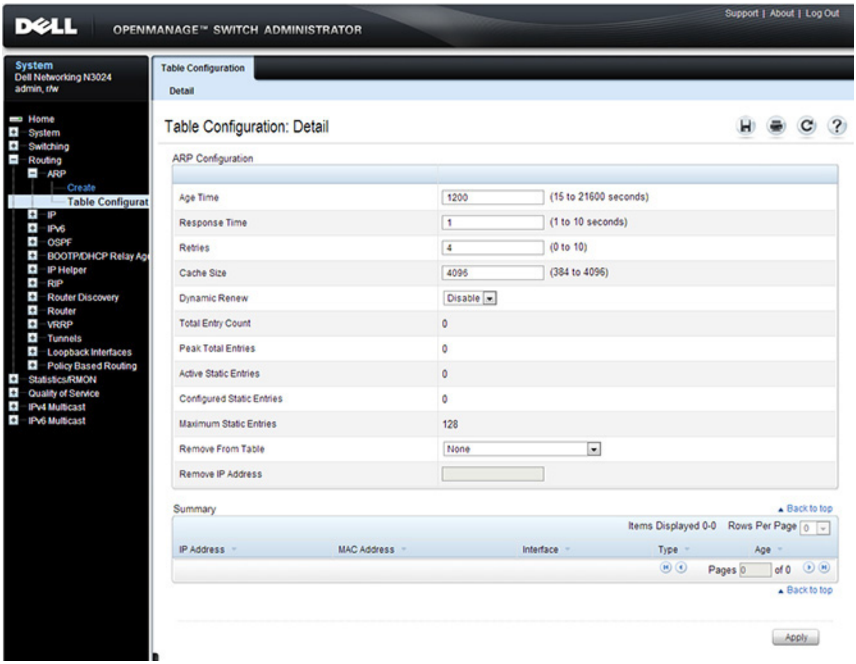


ARP Table Configuration

Use the **Table Configuration** page to change the configuration parameters for the Address Resolution Protocol Table. This page can also display the contents of the table.

To display the page, click **Routing** → **ARP** → **Table Configuration** in the navigation panel.

Figure 32-4. ARP Table Configuration

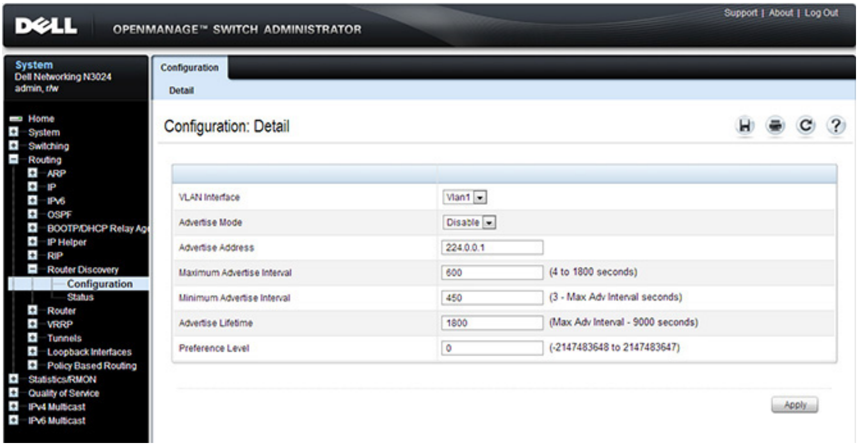


Router Discovery Configuration

Use the **Configuration** page to enter or change router discovery parameters.

To display the page, click **Routing** → **Router Discovery** → **Configuration** in the navigation panel.

Figure 32-5. Router Discovery Configuration

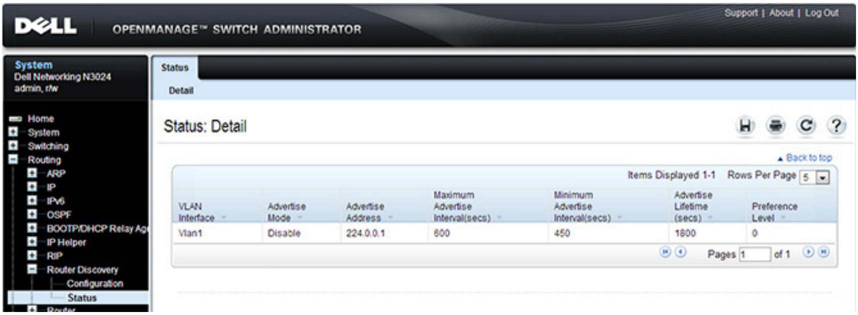


Router Discovery Status

Use the **Status** page to display router discovery data for each interface.

To display the page, click **Routing** → **Router Discovery** → **Status** in the navigation panel.

Figure 32-6. Router Discovery Status

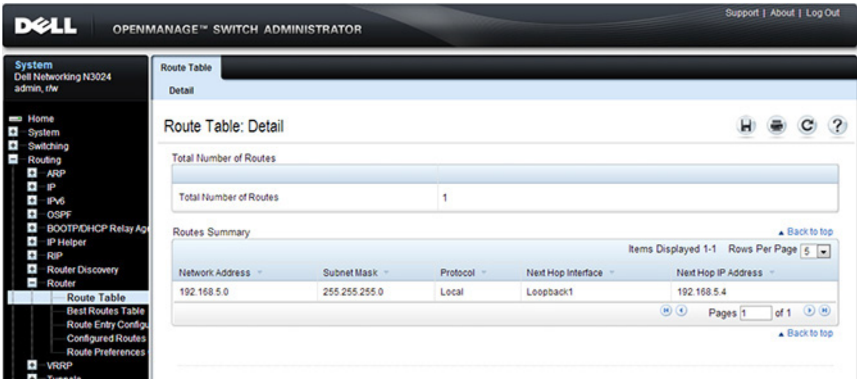


Route Table

Use the **Route Table** page to display the contents of the routing table.

To display the page, click **Routing** → **Router** → **Route Table** in the navigation panel.

Figure 32-7. Route Table

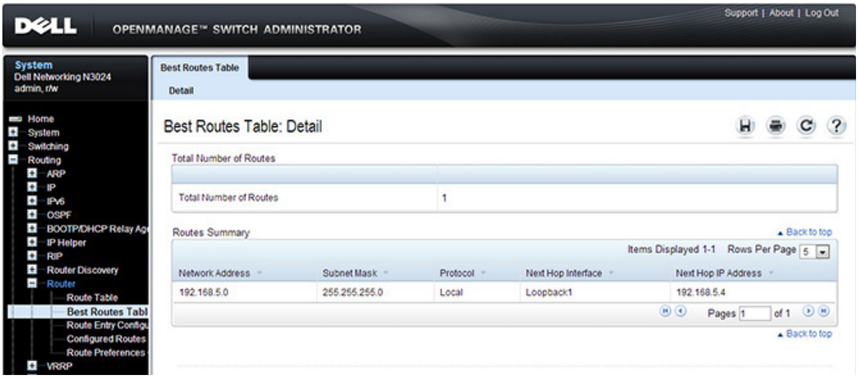


Best Routes Table

Use the **Best Routes Table** page to display the best routes from the routing table.

To display the page, click **Routing** → **Router** → **Best Routes Table** in the navigation panel.

Figure 32-8. Best Routes Table

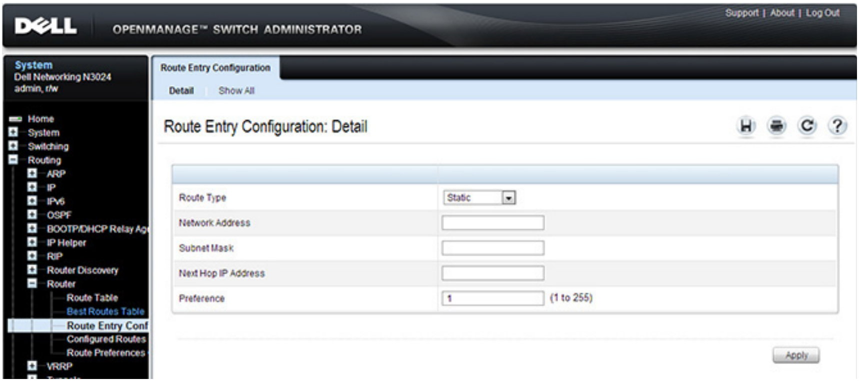


Route Entry Configuration

Use the **Route Entry Configuration** page to add new and configure router routes.

To display the page, click **Routing** → **Router** → **Route Entry Configuration** in the navigation panel.

Figure 32-9. Route Entry Configuration

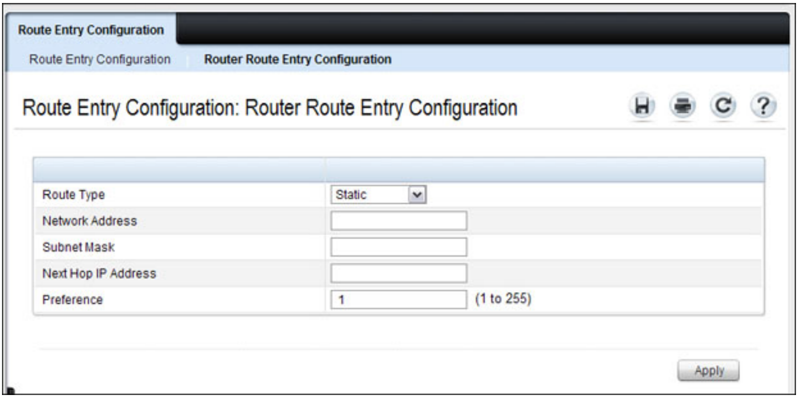


Adding a Route and Configuring Route Preference

To configure routing table entries:

1. Open the **Route Entry Configuration** page.

Figure 32-10. Router Route Entry and Preference Configuration



2 Next to **Route Type,** use the drop-down box to add a **Default, Static, or Static Reject** route.

The fields to configure are different for each route type.

- **Default** — Enter the default gateway address in the **Next Hop IP Address** field.
- **Static** — Enter values for **Network Address**, **Subnet Mask**, **Next Hop IP Address**, and **Preference**.
- **Static Reject** — Enter values for **Network Address**, **Subnet Mask**, and **Preference**.

3 Click **Apply**.

The new route is added to the routing table.

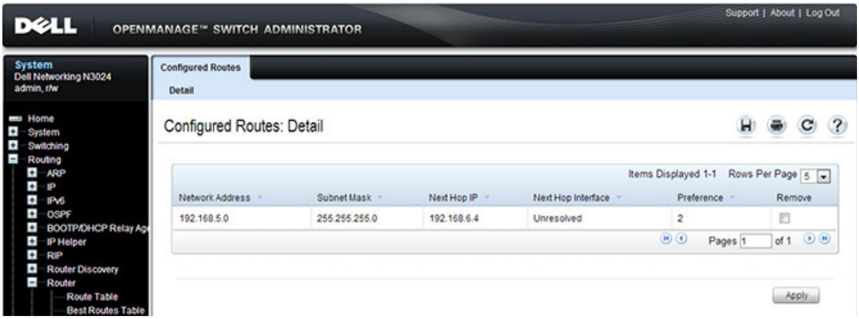
Configured Routes

Use the **Configured Routes** page to display the routes that have been manually configured.

NOTE: For a static reject route, the next hop interface value is Null0. Packets to the network address specified in static reject routes are intentionally dropped.

To display the page, click **Routing** → **Router** → **Configured Routes** in the navigation panel.

Figure 32-11. Configured Routes



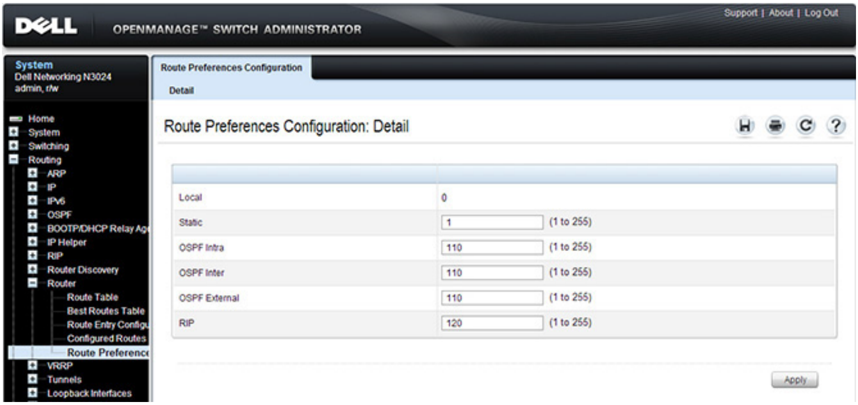
To remove a configured route, select the check box in the **Remove** column of the route to delete, and click **Apply**.

Route Preferences Configuration

Use the **Route Preferences Configuration** page to configure the default preference for each protocol (for example 60 for static routes). These values are arbitrary values that range from 1 to 255, and are independent of route metrics. Most routing protocols use a route metric to determine the shortest path known to the protocol, independent of any other protocol.

To display the page, click **Routing** → **Router** → **Route Preferences Configuration** in the navigation panel.

Figure 32-12. Router Route Preferences Configuration



Configuring IP Routing Features (CLI)

This section provides information about the commands used for configuring IPv4 routing on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Global IP Routing Settings

Beginning in Privileged EXEC mode, use the following commands to configure various global IP routing settings for the switch.

Command	Purpose
configure	Enter global configuration mode.
ip routing	Globally enable IPv4 routing on the switch.
ip icmp echo-reply	Allow the switch to generate ICMP Echo Reply messages.
ip icmp error-interval *burst-interval* [*burst-size*]	Limit the rate at which IPv4 ICMP error messages are sent. • *burst-interval* — How often the token bucket is initialized (Range: 0–2147483647 milliseconds). • *burst-size* — The maximum number of messages that can be sent during a burst interval (Range: 1–200).
ip redirects	Allow the switch to generate ICMP Redirect messages.
ip default-gateway *ip-address*	Configure the global default gateway for the switch. The gateway configured here takes precedence over a default gateway assigned by a network DHCP server.
exit	Exit to Privileged EXEC mode.
show ip brief	View the global IP settings for the switch.

Adding Static ARP Entries and Configuring ARP Table Settings

Beginning in Privileged EXEC mode, use the following commands to configure static ARP entries in the ARP cache and to specify the settings for the ARP cache.

Command	Purpose
configure	Enter global configuration mode.
arp *ip-address hardware-address*	Create a static ARP entry in the ARP table. • *ip-address* — IP address of a device on a subnet attached to an existing routing interface. • *hardware-address* — A unicast MAC address for that device.
arp timeout *seconds*	Configure the ARP entry ageout time.
arp resptime *seconds*	Configure the ARP request response timeout.
arp retries *integer*	Configure the ARP count of maximum requests for retries. The range is 1–10.
arp cachesize *integer*	Configure the maximum number of entries in the ARP cache.
arp dynamicrenew	Allow the ARP component to automatically renew dynamic ARP entries when they age out.
exit	Exit to Privileged EXEC mode.
show arp [brief]	View the user-configured (static) ARP entries. The static entries display regardless of whether they are reachable over an interface. Use the **brief** keyword to view only the ARP table settings.
clear arp-cache [gateway]	Remove all dynamic ARP entries from the ARP cache. Include the keyword **gateway** to remove gateway entries as well.
clear arp-cache management	Remove all dynamic ARP entries from the ARP cache that were learned on the management interface.
arp purge *ip-address*	Remove the specified IP address from the ARP cache. This command removes dynamic and gateway ARP entries only.

Configuring Router Discovery (IRDP)

Beginning in Privileged EXEC mode, use the following commands to configure IRDP settings.

Command	Purpose
configure	Enter global configuration mode.
interface *interface*	Enter interface configuration mode for the specified VLAN routing interface. The *interface* variable includes the interface type (**vlan**) and number, for example **vlan 100**.
ip irdp	Enable IRDP on the interface.
ip irdp address *ip-address*	Configure the address that the interface uses to send the router discovery advertisements. The allowed addresses are 224.0.0.1 (all-hosts IP multicast address) or 255.255.255.255 (limited broadcast address)
ip irdp holdtime *seconds*	Configure the value of the holdtime field of the router advertisement sent from this interface.
ip irdp maxadvertinterval *seconds*	Configure the maximum time allowed between sending router advertisements from the interface.
ip irdp minadvertinterval *seconds*	Configure the minimum time allowed between sending router advertisements from the interface.
ip irdp preference *integer*	Configure the preference of the address as a default router address relative to other router addresses on the same subnet.
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ip irdp [**vlan** *vlan-id*]	View the router discovery information for all interfaces, or for a specified interface.

Configuring Route Table Entries and Route Preferences

Beginning in Privileged EXEC mode, use the following commands to configure IRDP settings.

Command	Purpose
configure	Enter global configuration mode.
ip route default *nextHopRtr* [*preference*]	Configure the default route. • *nextHopRtr* — IP address of the next hop router. • *preference* — Specifies the preference value (administrative distance) of an individual static route. (Range: 1-255)
ip route *ip-addr* {*subnetmask* | *prefix length* } {*nextHopRtr* / **null***}* [*preference*]	Configure a static route. Use the keyword **null** instead of the next hop router IP address to configure a static reject route. • *ip-address* — IP address of destination interface. • *subnet-mask* — Subnet mask of destination interface. • *prefix-length* — Length of prefix. Must be preceded with a forward slash (/). (Range: 0-32 bits) • *nextHopRtr* — IP address of the next hop router. • **null** — Specifies that the route is a static reject route. • *preference* — Specifies the preference value (administrative distance) of an individual static route. (Range: 1-255)
ip route distance *integer*	Set the default distance (preference) for static routes. Lower route preference values are preferred when determining the best route.
exit	Exit to Privileged EXEC mode.

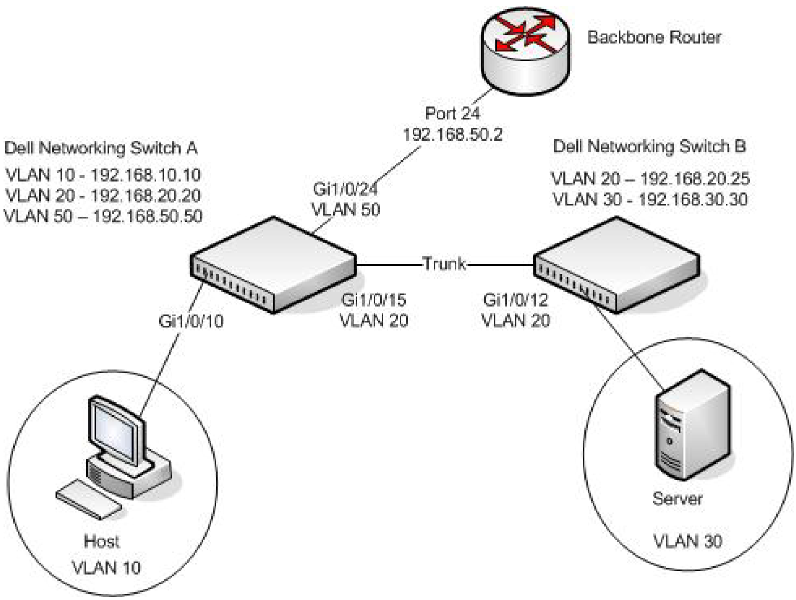
Command	Purpose
show ip route [*ip-address* [*mask* | *prefix-length*] [**longer-prefixes**] | *protocol*]	View the routing table. • *ip-address* — Specifies the network for which the route is to be displayed and displays the best matching best-route for the address. • *mask* — Subnet mask of the IP address. • *prefix-length* — Length of prefix, in bits. Must be preceded with a forward slash ('/'). (Range: 0-32 bits) • **longer-prefixes** — Indicates that the *ip-address* and *subnet-mask* pair becomes the prefix, and the command displays the routes to the addresses that match that prefix. • *protocol* — Specifies the protocol that installed the routes. (Range: **connected, ospf, rip, static**)
show ip route configured	View the configured routes, whether they are reachable or not.
show ip route summary	View summary information about the routing table.
show ip protocols	View the parameters and current state of the active routing protocols.
show ip route preferences	View detailed information about the route preferences.

IP Routing Configuration Example

In this example, the Dell Networking N-Series switches are layer-3 switches with VLAN routing interfaces. VLAN routing is configured on Dell Networking N-Series Switch A and Dell Networking N-Series Switch B. This allows the host in VLAN 10 to communicate with the server in VLAN 30. A static route to the VLAN 30 subnet is configured on Switch A. Additionally, a default route is configured on Switch A so that all traffic with an unknown destination is sent to the backbone router through port 24, which is a member of VLAN 50. A default route is configured on Dell Networking N-Series switch B to use Switch A as the default gateway. The hosts use the IP address of the VLAN routing interface as their default gateway.

This example assumes that all layer-2 VLAN information, such as VLAN creation and port membership, has been configured.

Figure 32-13. IP Routing Example Topology



Configuring Dell Networking N-Series Switch A

To configure Switch A.

1 Enable routing on the switch.

```
console#configure
console(config)#ip routing
```

2 Assign an IP address to VLAN 10. This command also enables IP routing on the VLAN.

```
console(config)#interface vlan 10
console(config-if-vlan10)#ip address 192.168.10.10
255.255.255.0
console(config-if-vlan10)#exit
```

3 Assign an IP address to VLAN 20.

```
console#configure
console(config)#interface vlan 20
console(config-if-vlan20)#ip address 192.168.20.20
255.255.255.0
console(config-if-vlan20)#exit
```

4 Assign an IP address to VLAN 50.

```
console#configure
console(config)#interface vlan 50
console(config-if-vlan50)#ip address 192.168.50.50
255.255.255.0
console(config-if-vlan50)#exit
```

5 Configure a static route to the network that VLAN 30 is in, using the IP address of the VLAN 20 interface on Switch B as the next hop address.

```
console(config)#ip route 192.168.30.0 255.255.255.0
192.168.20.25
```

6 Configure the backbone router interface as the default gateway.

```
console(config)#ip route default 192.168.50.2
```

Configuring Dell Networking N-Series Switch B

To configure Switch B:

1 Enable routing on the switch.

```
console#configure
console(config)#ip routing
```

2 Assign an IP address to VLAN 20. This command also enables IP routing on the VLAN.

```
console#configure
console(config)#interface vlan 20
console(config-if-vlan20)#ip address 192.168.20.25
255.255.255.0
console(config-if-vlan20)#exit
```

3 Assign an IP address to VLAN 30. This command also enables IP routing on the VLAN.

```
console#configure
console(config)#interface vlan 30
console(config-if-vlan30)#ip address 192.168.30.30
255.255.255.0
console(config-if-vlan30)#exit
```

4 Configure the VLAN 20 routing interface on Switch A as the default gateway so that any traffic with an unknown destination is sent to Switch A for forwarding.

```
console(config)#ip route default 192.168.20.20
```

33

Routing Interfaces

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes the routing (layer-3) interfaces the Dell Networking N-Series switches support, which includes VLAN routing interfaces, loopback interfaces, and tunnel interfaces.

The topics covered in this chapter are:

- Routing Interface Overview
- Default Routing Interface Values
- Configuring Routing Interfaces (Web)
- Configuring Routing Interfaces (CLI)

For information about configuring IPv6 characteristics on routing interfaces, see "IPv6 Routing " on page 1397.

For configuration examples that configure VLAN routing interfaces, see "IP Routing Configuration Example " on page 1137 in the IP Routing chapter. For a configuration example that includes tunnel and loopback interface creation, see "Interconnecting an IPv4 Backbone and Local IPv6 Network " on page 1256.

Routing Interface Overview

Routing interfaces are logical interfaces that can be configured with an IP address. Routing interfaces provide a means of transmitting IP packets between subnets on the network.

What Are VLAN Routing Interfaces?

VLANs divide a single physical network (broadcast domain) into separate logical networks. To forward traffic across VLAN boundaries, a layer-3 device, such as router, is required. Dell Networking N-Series switches can act as layer-3 devices when you configure VLAN routing interfaces. VLAN routing interfaces make it possible to transmit traffic between VLANs while still containing broadcast traffic within VLAN boundaries. The configuration of VLAN routing interfaces makes inter-VLAN routing possible.

For each VLAN routing interface a static IP address can be assigned, or a network DHCP server can assign a dynamic IP address.

When a port is enabled for bridging (layer-2 switching) rather than routing, which is the default, all normal bridge processing is performed for an inbound packet, which is then associated with a VLAN. Its MAC Destination Address (MAC DA) and VLAN ID are used to search the MAC address table. If routing is enabled for the VLAN, and the MAC DA of an inbound unicast packet is that of the internal router interface, the packet is routed. An inbound multicast packet is forwarded to all ports in the VLAN, plus the internal bridge-router interface, if it was received on a routed VLAN.

Since a port can be configured to belong to more than one VLAN, VLAN routing might be enabled for all of the VLANs on the port or for only some of the VLANs on the port. VLAN Routing can be used to allow more than one physical port to reside on the same subnet. It could also be used when a VLAN spans multiple physical networks, or when additional segmentation or security is required.

What Are Loopback Interfaces?

A loopback interface is a logical interface that is always up and, because it cannot go down, allows the switch to have a stable IP address that other network devices and protocols can use to reach the switch. The loopback can provide the source address for sent packets.

NOTE: In this context, loopback interfaces should not be confused with the loopback IP address, usually 127.0.0.1, assigned to a host for handling self-routed packets.

The loopback interface does not behave like a network switching port. Specifically, there are no neighbors on a loopback interface; it is a pseudo-device for assigning local addresses so that the other layer-3 devices can communicate with the switch by using the loopback IP address. The loopback interface is always up and can receive traffic from any of the existing active interfaces. Thus, given reachability from a remote client, the address of the loopback can be used to communicate with the switch through various services such as Telnet and SSH. In this way, the IP address on a loopback behaves identically to any of the local addresses of the VLAN routing interfaces in terms of the processing of incoming packets.

What Are Tunnel Interfaces?

Tunnels are a mechanism for transporting a packet across a network so that it can be evaluated at a remote location or *tunnel endpoint*. The tunnel, effectively, hides the packet from the network used to transport the packet to the endpoint. This allows for the transmission of packets that the transport network cannot process directly, such as in one of the following cases:

- The packet protocol is not supported.
- The packet is in an incompatible addressing space.
- The packet is encrypted.

Dell Networking N-Series switches support tunnels to encapsulate IPv6 traffic in IPv4 tunnels to provide functionality to facilitate the transition of IPv4 networks to IPv6 networks.

The switch supports two types of tunnels: configured (6-in-4) and automatic (6-to-4). Configured tunnels have an explicit configured endpoint and are considered to be point-to-point interfaces. Automatic tunnels determine the endpoint of the tunnel from the destination address of packets routed into the tunnel. These tunnels correspond to Non-Broadcast Multi-Access (NBMA) interfaces. A configured tunnel interface has a single tunnel associated with it, while an automatic tunnel interface has an infinite number of tunnels (limited only by the address encoding scheme).

Because tunnels are used as logical interfaces, static routes can be defined that reference the tunnels. Additionally, dynamic routing can be configured to use the tunnels.

Why Are Routing Interfaces Needed?

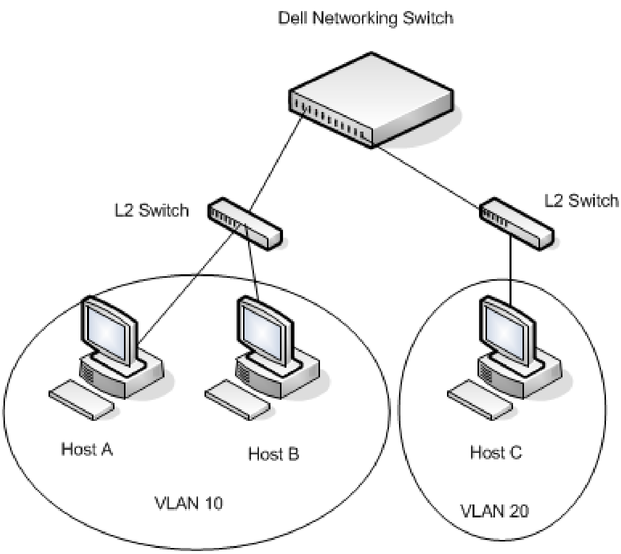
The routing interfaces this chapter describes have very different applications and uses, as this section describes. If you use the switch as a layer-2 device that handles switching only, routing interface configuration is not required. When the switch is used as a layer-2 device, it typically connects to an external layer-3 device that handles the routing functions.

VLAN Routing

VLAN routing is required when the switch is used as a layer-3 device. VLAN routing must be configured to allow the switch to forward IP traffic between subnets and allow hosts in different networks to communicate.

In Figure 33-1 the Dell Networking N-Series switch is configured as a layer-3 device and performs the routing functions for hosts connected to the layer-2 switches. For Host A to communicate with Host B, no routing is necessary. These hosts are in the same VLAN. However, for Host A in VLAN 10 to communicate with Host C in VLAN 20, the Dell Networking N-Series switch must perform inter-VLAN routing.

Figure 33-1. Inter-VLAN Routing



Loopback Interfaces

When packets are sent to the loopback IP address, the network should be able to deliver the packets as long as any physical interface on the switch is up. There are many cases where you need to send traffic to a switch, such as in switch management. The loopback interface IP address is a good choice for communicating with the switch in these cases because the loopback interface cannot go down when the switch is powered on and operational.

Tunnel Interface

Tunnels can be used in networks that support both IPv6 and IPv4. The tunnel allows non-contiguous IPv6 networks to be connected over an IPv4 infrastructure.

Default Routing Interface Values

By default, no routing interfaces are configured.

When you create a VLAN, no IP address is configured, and DHCP is disabled. After you configure an IP address on a VLAN or loopback interface, the VLAN interface is available for layer-3 routing (if enabled) and is capable of resolved ARPs and responding to pings, and the interface has the default configuration shown in Table 33-1.

Most interface configuration parameters are not applicable to loopback interfaces, so the default values cannot be changed. However, when a loopback interface is created, the default values are similar to those of VLAN routing interfaces, as Table 33-1 shows.

Table 33-1. VLAN Routing Interface and Loopback Interface Defaults

Parameter	Default Value
Forward Net Directed Broadcasts	Disabled
Encapsulation Type	Ethernet (N/A for loopbacks)
Proxy Arp	Enabled
Local Proxy Arp	Disabled
IP MTU	1500
Bandwidth	Not configured.
Destination Unreachables	Enabled
ICMP Redirects	Enabled

When you create a tunnel, it has the default values shown in Table 33-2

Table 33-2. Tunnel Interface Defaults

Parameter	Default Value
Tunnel mode	6-in-4 configured
Link Local Only Mode	Disabled
Source address	None
Destination address	0.0.0.0

Configuring Routing Interfaces (Web)

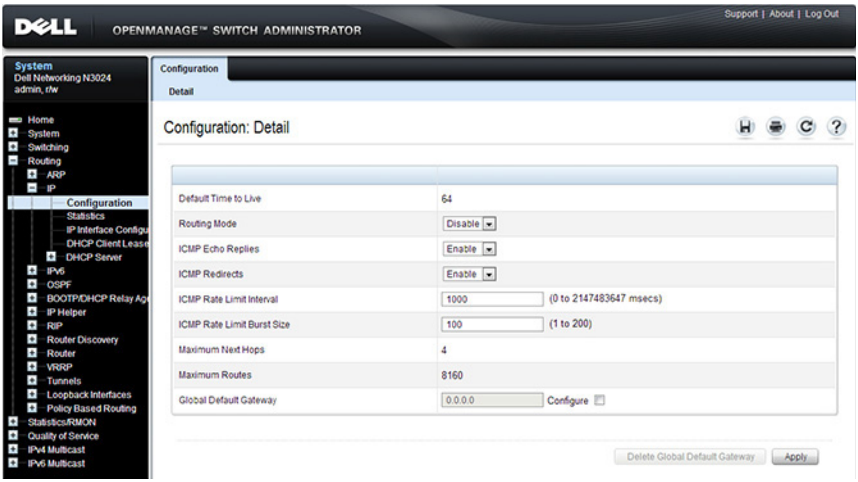
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring VLAN routing interfaces, loopback interfaces, and tunnels on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

IP Interface Configuration

Use the **IP Interface Configuration** page to update IP interface data for this switch. The IP interface configuration includes the ability to configure the bandwidth, Destination Unreachable messages, and ICMP Redirect messages.

To display the page, click **Routing** → **IP** → **IP Interface Configuration** in the navigation panel.

Figure 33-2. IP Interface Configuration

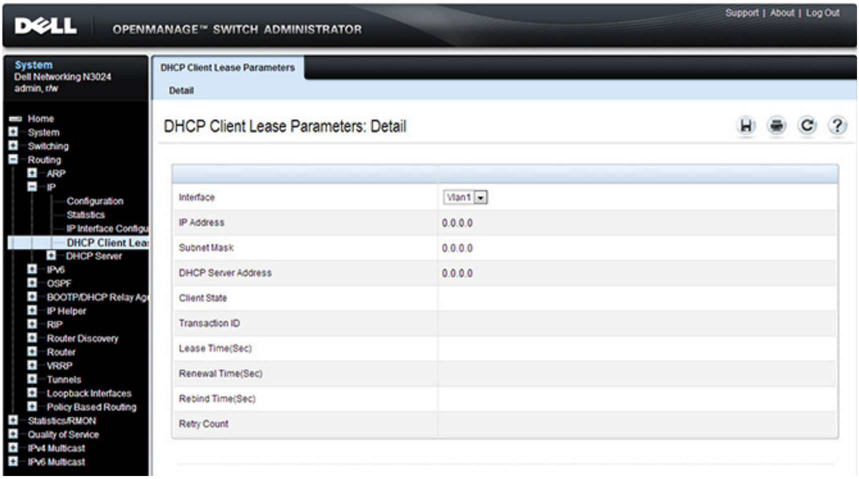


DHCP Lease Parameters

Use the **DHCP Lease Parameters** page to view information about the network information automatically assigned to an interface by the DHCP server.

To display the page, click **Routing** → **IP** → **DHCP Lease Parameters** in the navigation panel.

Figure 33-3. DHCP Lease Parameters

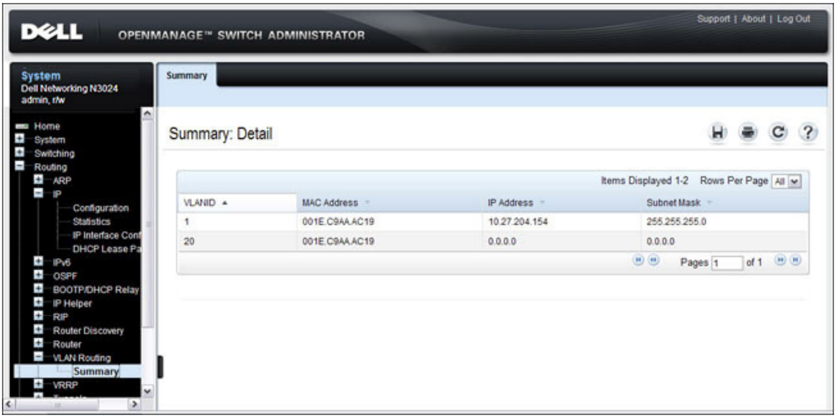


VLAN Routing Summary

Use the **VLAN Routing Summary** page to view summary information about VLAN routing interfaces configured on the switch.

To display the page, click **Routing** → **VLAN Routing** → **Summary** in the navigation panel.

Figure 33-4.   VLAN Routing Summary

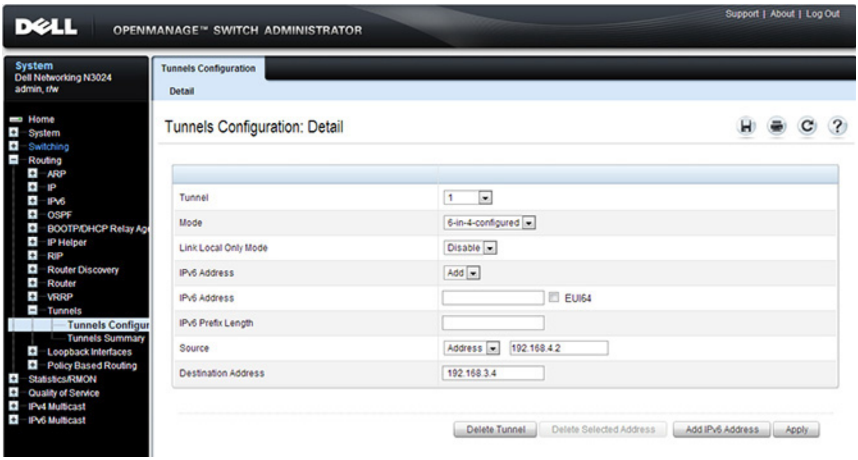


Tunnel Configuration

Use the **Tunnels Configuration** page to create, configure, or delete a tunnel.

To display the page, click **Routing** → **Tunnels** → **Configuration** in the navigation panel.

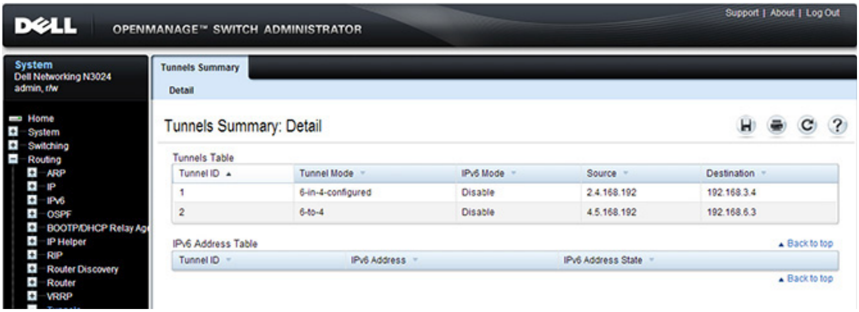
Figure 33-5.   Tunnel Configuration



Tunnels Summary

Use the **Tunnels Summary** page to display a summary of configured tunnels. To display the page, click **Routing** → **Tunnels** → **Summary** in the navigation panel.

Figure 33-6. Tunnels Summary

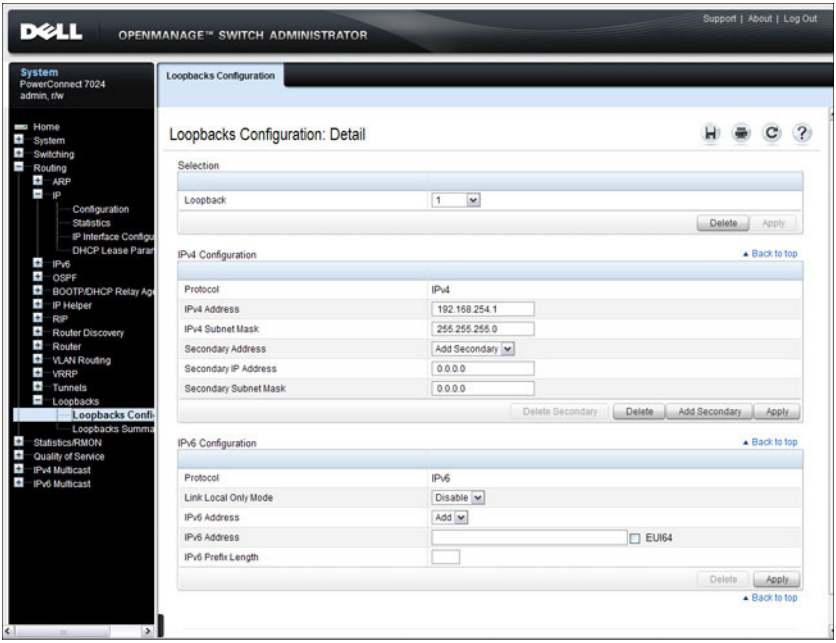


Loopbacks Configuration

Use the **Loopbacks Configuration** page to create, configure, or remove loopback interfaces. A secondary address for a loopback can also be set up or deleted.

To display the page, click **Routing** → **Loopback Interfaces** → **Loopback Interfaces Configuration** in the navigation panel.

Figure 33-7. Loopback Configuration

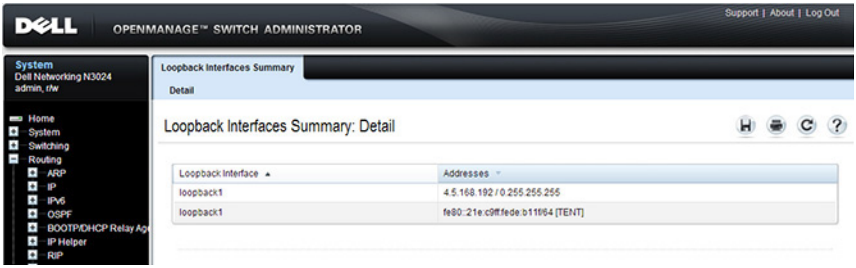


Loopbacks Summary

Use the **Loopbacks Summary** page to display a summary of configured loopback interfaces on the switch.

To display the page, click **Routing** → **Loopback Interfaces** → **Loopback Interfaces Summary** in the navigation panel.

Figure 33-8. Loopbacks Summary



Configuring Routing Interfaces (CLI)

This section provides information about the commands used for configuring VLAN routing interfaces, loopbacks, and tunnels on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring VLAN Routing Interfaces (IPv4)

Beginning in Privileged EXEC mode, use the following commands to configure a VLAN as a routing interface and set the IP configuration parameters.

Command	Purpose
configure	Enter Global Configuration mode.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ip address {dhcp | none | *ip_address subnet_mask* **[secondary]}**	Configure the IP address. Use the **dhcp** keyword to enable the DHCP client and obtain an IP address from a network DHCP server. Use **none** to release the address obtained from the DHCP server. Use *ip_address* and *subnet_mask* to assign a static IP address. For a static address, use the **secondary** keyword to specify that the address is a secondary IP address.
ip netdirbcast	Enable the forwarding of network-directed broadcasts.
encapsulation {ethernet | snap}	Configure the link-layer encapsulation type for the packet. Routed frames are always ethernet encapsulated when a frame is routed to a VLAN.
ip proxy-arp	Enable proxy ARP on the interface. Without proxy ARP, the switch responds to an ARP request only if the target IP address is an address configured on the interface where the ARP request arrived. This command is not available in interface range mode.

Command	Purpose
ip local-proxy-arp	Enable local proxy ARP on the interface to allow the switch to respond to ARP requests for hosts on the same subnet as the ARP source.
bandwidth *size*	Set the configured bandwidth on this interface to communicate the speed of the interface to higher level protocols. OSPF uses the bandwidth value to compute link cost. The range is 1–10000000.
ip unreachables	Allow the switch to send ICMP Destination Unreachable messages in response to packets received on the interface.
ip redirects	Allow the switch to send ICMP Redirect messages in response to packets received on the interface.
exit	Exit to Global Config mode.
ip default-gateway *ip_address*	Configure the default gateway. All switch interfaces use the same default gateway.
exit	Exit to Privileged EXEC mode.
show dhcp lease [**interface** *interface*]	View information about the DHCP leases acquired for all interfaces or for the specified interface. For a VLAN, the *interface_string* parameter is **vlan** followed by the VLAN ID, with no space, for example vlan10.
show ip interface vlan *vlan-id*	View the IP interface configuration information for the specified routing VLAN.

Configuring Loopback Interfaces

Beginning in Privileged EXEC mode, use the following commands to configure a loopback interface.

Command	Purpose
configure	Enter Global Configuration mode.
interface loopback *loopback-id*	Create the loopback interface and enter Interface Configuration mode for the specified loopback interface.
ip address *ip_address subnet_mask* [**secondary**]	Configure a static IP address and subnet mask. Use the **secondary** keyword to specify that the address is a secondary IP address.
CTRL + Z	Exit to Privileged EXEC mode.
show ip interface loopback *loopback-id*	View interface configuration information for the specified loopback interface.

Configuring Tunnels

Beginning in Privileged EXEC mode, use the following commands to configure a loopback interface.

NOTE: For information about configuring the IPv6 interface characteristics for a tunnel, see "IPv6 Routing " on page 1397.

Command	Purpose
configure	Enter Global Configuration mode.
interface tunnel *tunnel-id*	Create the tunnel interface and enter Interface Configuration mode for the specified tunnel.
tunnel mode ipv6ip [6to4]	Specify the mode of the tunnel. If you use the **6to4** keyword, the tunnel is an automatic tunnel. If you omit the keyword, the tunnel is a point-to-point (configured) tunnel.
ipv6 enable	Enable IPv6 on this interface using the Link Local address.
tunnel source {*ipv4addr* | **vlan** *vlan-id*}	Specify the source transport address of the tunnel, either, which can be an IPv4 address or a VLAN routing interface.
tunnel destination *ipv4addr*	Specify the destination transport IPv4 address of the tunnel.
CTRL + Z	Exit to Privileged EXEC mode.
show interfaces tunnel [*tunnel-id*]	View configuration information for all tunnels or for the specified tunnel.

34

Layer-2 and Layer-3 Relay Features

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure the layer-2 (L2) DHCP relay, layer-3 (L3) DHCP relay, and IP Helper features on Dell Networking N-Series switches.

The topics covered in this chapter include:

- L2 and L3 Relay Overview
- Default L2/L3 Relay Values
- Configuring L2 and L3 Relay Features (Web)
- Configuring L2 and L3 Relay Features (CLI)
- Relay Agent Configuration Example

L2 and L3 Relay Overview

When a DHCP client and server are in the same IP subnet, they can directly connect to exchange IP address requests and replies. However, buying and maintaining a DHCP server on each subnet can be expensive and is often impractical. The IP Helper/DHCP Relay features on the Dell Networking N-Series switches can help enable communication between DHCP clients and DHCP servers that reside in different subnets. Configuring L3 DHCP relay also enables the bootstrap protocol (BOOTP) relay.

What Is L3 DHCP Relay?

Network infrastructure devices can be used to relay packets between a DHCP client and server on different subnets. Such a device, a layer-3 relay agent, is often a router or L3 switch. The L3 relay agent must have an IP interface on the client subnets and, if it does not have an IP interface on the server's subnet, it should be able to route traffic toward the server's subnet.

The Dell Networking DHCP Relay Agent enables DHCP clients and servers to exchange DHCP messages across different subnets. The relay agent receives the requests from the clients, and checks the valid hops and *giaddr* fields in the DHCP request. If the number of hops is greater than the configured number, the agent discards the packet. If the *giaddr* field is zero, the agent must fill in this field with the IP address of the interface on which the request was received. The agent unicasts the valid packets to all configured DHCP servers. Each server responds with a unicast BOOTREPLY addressed to the relay agent closest to the client as indicated by *giaddr* field. Upon reception of the BOOTREPLY from the server, the agent forwards this reply as broadcast or unicast on the interface where the BOOTREQUEST arrived. This interface can be identified by the *giaddr* field or option 82.

The Dell Networking N1500, N2000, N3000, and N4000 Series switches DHCP component also supports DHCP relay agent options to identify the client interface. If configured, the relay agent inserts these options when forwarding the request to the server and removes them when sending the reply to the clients.

If an interface has more than one IP address, the relay agent uses the primary IP address configured as its relay agent IP address.

What Is L2 DHCP Relay?

In layer-2 switched networks, hosts (DHCP clients) may be connected directly to a switch which is connected to a router configured as a DHCP relay agent, or they may be connected directly to a DHCP server (unusual). In this instance, some of the client device information required by the DHCP server may not be included in the DHCP packets sent by the DHCP client. An L2 relay agent can be used to add the information that the DHCP server needs to perform its role in address and configuration and assignment. The information added by the L2 relay agent can include location and identification information that can assist the DHCP server in applying policies such as service offerings or address assignment.

Before it relays DHCP requests from clients, the switch can add a Circuit ID and a Remote ID. These IDs provide information about the circuit and port number connected to the client. This information is added as suboptions in the DHCP Option 82 packets as defined in sections 3.1 and 3.2 of RFC3046. The switch removes this option from packets that it relays from L3 Relay agents/DHCP servers to clients.

These sub-options may be used by the DHCP server to affect how it treats the client and also may be used by the relay agent to limit broadcast replies to the specific circuit or attachment point of the client.

Enabling L2 Relay on VLANs

L2 DHCP relay can be enabled on a particular VLAN. The VLAN is identified by a service VLAN ID (S-VID), which a service provider uses to identify a customer's traffic while traversing the provider network to multiple remote sites. The switch uses the VLAN membership of the switch port client (the customer VLAN ID, or C-VID) to perform a lookup a corresponding S-VID.

If the S-VID is enabled for DHCP Relay, then the packet can be forwarded. If the C-VID does not correspond to an S-VID that is enabled for DHCP Relay, then the switch will not relay the DHCP request packet.

What Is the IP Helper Feature?

The IP Helper feature provides the ability for a router to unicast-forward configured UDP broadcast packets to a particular IP address (including DHCP packets). This allows applications to reach servers on non-local subnets. This is possible even when the application is designed to assume a server is always on a local subnet or when the application uses broadcast packets to reach the server (with the limited broadcast address 255.255.255.255, or a network directed broadcast address).

Relay entries may be configured globally and on routing interfaces. Each relay entry maps an ingress interface and destination UDP port number to a single IPv4 address (the helper address). Multiple relay entries may be configured for the same interface and UDP port, in which case the relay agent relays matching packets to each server address. Interface configuration takes priority over global configuration. If the destination UDP port for a packet matches any entry on the ingress interface, the packet is handled according to the interface configuration. If the packet does not match any entry on the ingress interface, the packet is handled according to the global IP helper configuration.

Discard relay entries may also be configured. Discard entries are used to discard packets received on a specific interface when those packets would otherwise be relayed according to a global relay entry. Discard relay entries may be configured on interfaces, but are not configured globally.

Additionally, the administrator can configure which UDP ports are forwarded. Certain UDP port numbers can be selected from the web interface or specified by name in the CLI, but a relay entry can also be configured with any UDP port number. It is possible to configure relay entries that do not specify a destination UDP port. The relay agent assumes that these entries match packets with the UDP destination ports listed in Table 34-1 (the list of default ports).

Table 34-1. Default Ports - UDP Port Numbers Implied By Wildcard

Protocol	UDP Port Number
IEN-116 Name Service	42
DNS	53
NetBIOS Name Server	137
NetBIOS Datagram Server	138
TACACS Server	49
Time Service	37
DHCP	67
Trivial File Transfer Protocol	69

The system limits the total number of relay entries to four times the maximum number of routing interfaces (512 relay entries). There is no individual limit to the number of relay entries on an individual interface, and no individual limit to the number of servers for a given {interface, UDP port} pair. The system limit applies in these cases.

Certain configurable DHCP relay options do not apply to relay of other protocols. You may optionally set a maximum hop count or minimum wait time using the **bootpdhcprelay maxhopcount** and **bootpdhcprelay minwaittime** commands.

The relay agent relays DHCP packets in both directions. It relays broadcast packets from the client to one or more DHCP servers, and relays packets to the client that the DHCP server unicasts back to the relay agent. For other protocols, the relay agent only relays broadcast packets from the client to the server. Packets from the server back to the client are assumed to be unicast directly to the client. Because there is no relay in the return direction for protocols other than DHCP, the relay agent retains the source IP address from the original client packet. The relay agent uses a local IP address as the source IP address of relayed DHCP client packets.

When a switch receives a broadcast UDP packet on a routing interface, the relay agent verifies that the interface is configured to relay to the destination UDP port. If so, the relay agent unicasts the packet to the configured server IP addresses. Otherwise, the relay agent verifies that there is a global configuration for the destination UDP port. If so, the relay agent unicasts the packet to the configured server IP addresses. Otherwise the packet is not relayed.

NOTE: If the packet matches a discard relay entry on the ingress interface, the packet is not forwarded, regardless of the global configuration.

The relay agent relays packets that meet only the following conditions:

- The destination MAC address must be the all-ones broadcast address (FF:FF:FF:FF:FF:FF).
- The destination IP address must be the limited broadcast address (255.255.255.255) or a directed broadcast address for the receive interface.
- The IP time-to-live (TTL) must be greater than 1.
- The protocol field in the IP header must be UDP (17).
- The destination UDP port must match a configured relay entry.

Table 34-2 shows the most common protocols and their UDP port numbers and names that are relayed.

Table 34-2. UDP Port Allocations

UDP Port Number	Acronym	Application
7	Echo	Echo
11	SysStat	Active User
15	NetStat	NetStat
17	Quote	Quote of the day
19	CHARGEN	Character Generator
20	FTP-data	FTP Data
21	FTP	FTP
37	Time	Time
42	NAMESERVER	Host Name Server
43	NICNAME	Who is
53	DOMAIN	Domain Name Server
69	TFTP	Trivial File Transfer
111	SUNRPC	Sun Microsystems Rpc
123	NTP	Network Time
137	NetBiosNameService	NT Server to Station Connections
138	NetBiosDatagramService	NT Server to Station Connections
139	NetBios	SessionServiceNT Server to Station Connections
161	SNMP	Simple Network Management
162	SNMP-trap	Simple Network Management Traps
513	who	Unix Rwho Daemon
514	syslog	System Log
525	timed	Time Daemon

Default L2/L3 Relay Values

By default L2 DHCP relay is disabled. L3 relay (UDP) is enabled, but no UDP destination ports or server addresses are defined on the switch or on any interfaces.

Table 34-3. L2/L3 Relay Defaults

Parameter	Default Value
L2 DHCP Relay	
Admin Mode	Disabled globally and on all interfaces and VLANs
Trust Mode	Disabled on all interfaces
Circuit ID	Disabled on all VLANs
Remote ID	None configured
L3 DHCP Relay	
UDP Relay Mode (IP Helper)	Enabled
Hop Count	4
Minimum Wait Time	0 seconds
Circuit ID Option Mode	Disabled
Circuit ID Check Mode	Enabled
Information Option-Insert	Disabled on all VLAN interfaces
Information Check-Reply	Enabled on all VLAN interfaces

Configuring L2 and L3 Relay Features (Web)

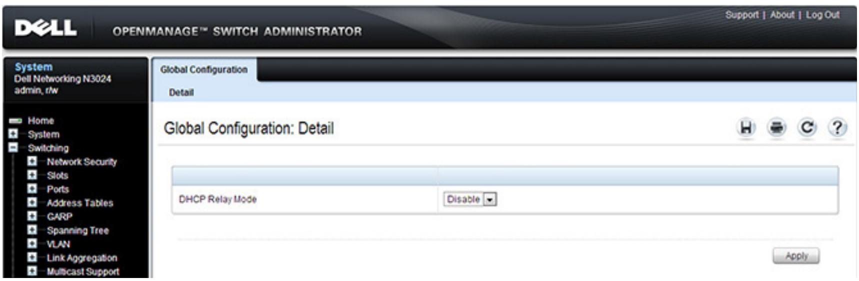
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring L2 and L3 relay features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

DHCP Relay Global Configuration

Use this page to enable or disable the switch to act as a DHCP Relay agent. This functionality must also be enabled on each port you want this service to operate on (see "DHCP Relay Interface Configuration " on page 1165). The switch can also be configured to relay requests only when the VLAN of the requesting client corresponds to a service provider's VLAN ID that has been enabled with the L2 DHCP relay functionality (see "DHCP Relay VLAN Configuration " on page 1168).

To access this page, click **Switching → DHCP Relay → Global Configuration** in the navigation panel.

Figure 34-1. DHCP Relay Global Configuration



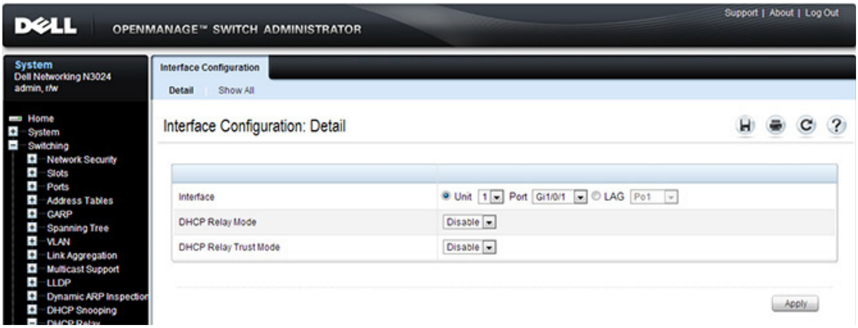
DHCP Relay Interface Configuration

Use this page to enable L2 DHCP relay on individual ports.

NOTE: L2 DHCP relay must also be enabled globally on the switch.

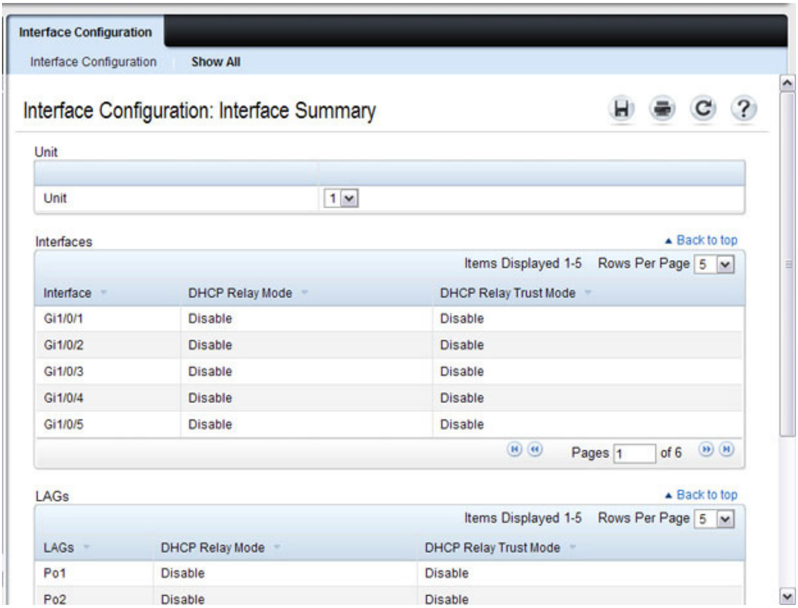
To access this page, click **Switching** → **DHCP Relay** → **Interface Configuration** in the navigation panel.

Figure 34-2. DHCP Relay Interface Configuration



To view a summary of the L2 DHCP relay configuration on all ports and LAGS, click **Show All**.

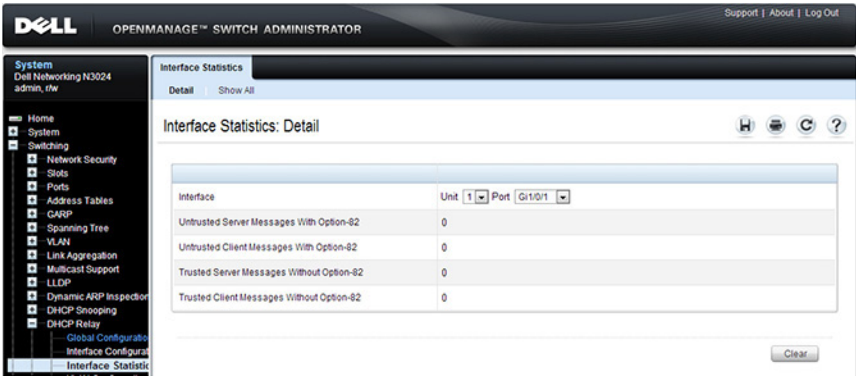
Figure 34-3. DHCP Relay Interface Summary



DHCP Relay Interface Statistics

Use this page to display statistics on DHCP Relay requests received on a selected port. To access this page, click **Switching** → **DHCP Relay** → **Interface Statistics** in the navigation panel.

Figure 34-4. DHCP Relay Interface Statistics

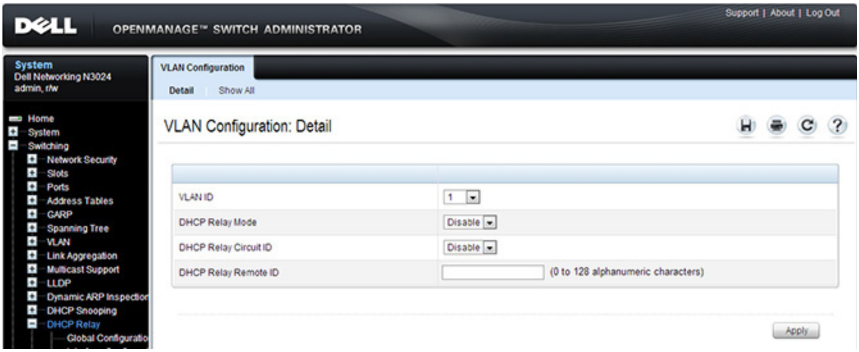


DHCP Relay VLAN Configuration

Use this page to enable and configure DHCP Relay on specific VLANs.

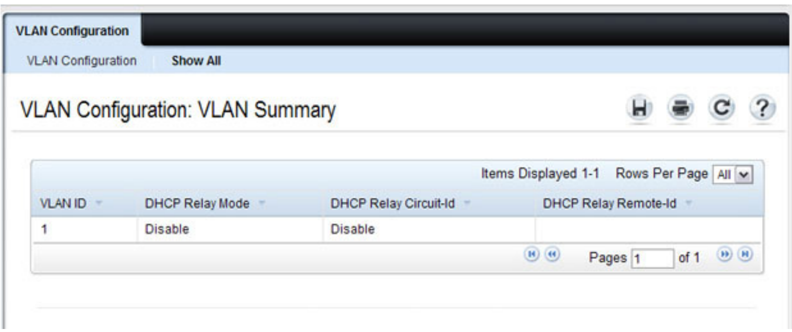
To access this page, click **Switching** → **DHCP Relay** → **VLAN Configuration** in the navigation panel.

Figure 34-5. DHCP Relay VLAN Configuration



To view a summary of the L2 DHCP relay configuration on all VLANs, click **Show All**.

Figure 34-6. DHCP Relay VLAN Summary

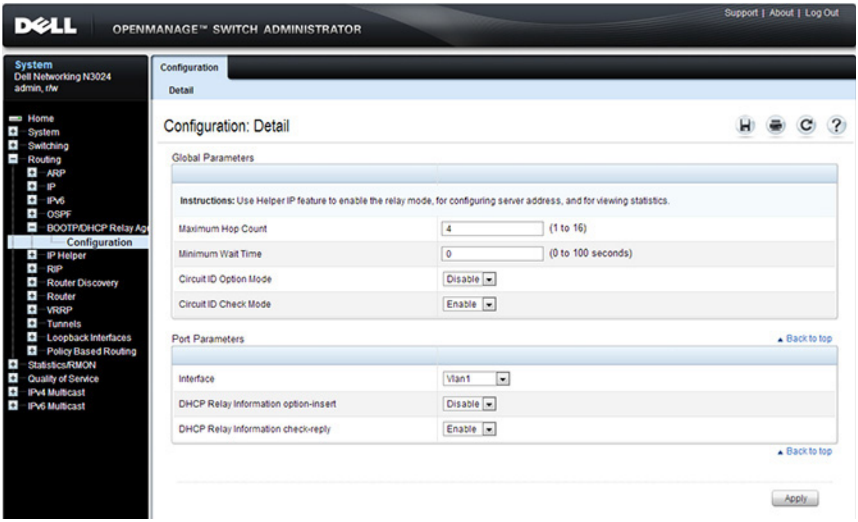


DHCP Relay Agent Configuration

Use the **Configuration** page to configure and display a DHCP relay agent.

To display the page, click **Routing** → **BOOTP/DHCP Relay Agent** → **Configuration** in the navigation panel.

Figure 34-7. DHCP Relay Agent Configuration

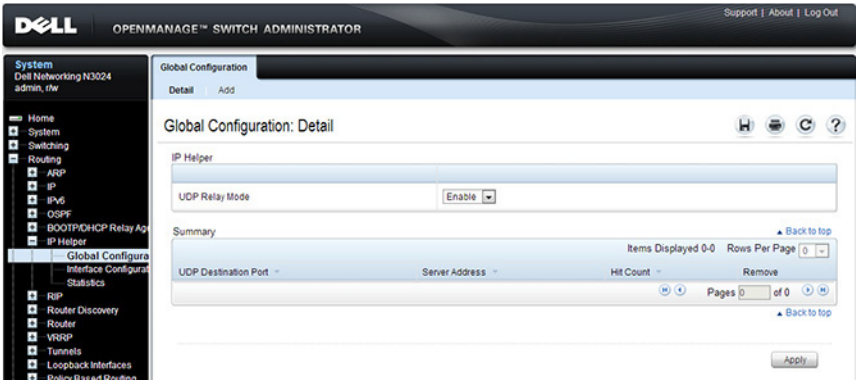


IP Helper Global Configuration

Use the **Global Configuration** page to add, show, or delete UDP Relay and Helper IP configuration

To display the page, click **Routing** → **IP Helper** → **Global Configuration** in the navigation panel.

Figure 34-8. IP Helper Global Configuration

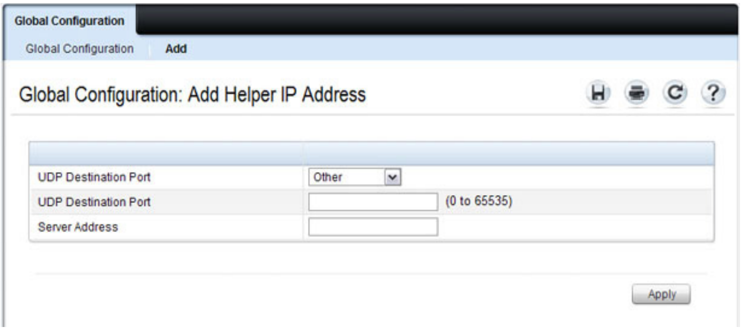


Adding an IP Helper Entry

To configure an IP helper entry:

1. Open the IP Helper **Global Configuration** page.
2. Click **Add** to display the **Add Helper IP Address** page:

Figure 34-9. Add Helper IP Address



3. Select a UDP Destination port name from the menu or enter the UDP Destination Port ID. Select the Default Set to configure for the relay entry for the default set of protocols.

   NOTE: If the DefaultSet option is specified, the device by default forwards UDP Broadcast packets for the following services: IEN-116 Name Service (port 42), DNS (port 53), NetBIOS Name Server (port 137), NetBIOS Datagram Server (port 138), TACACS Server (Port 49), and Time Service (port 37).

4. Enter the IP address of the server to which the packets with the given UDP Destination Port will be relayed.

5. Click **Apply**.

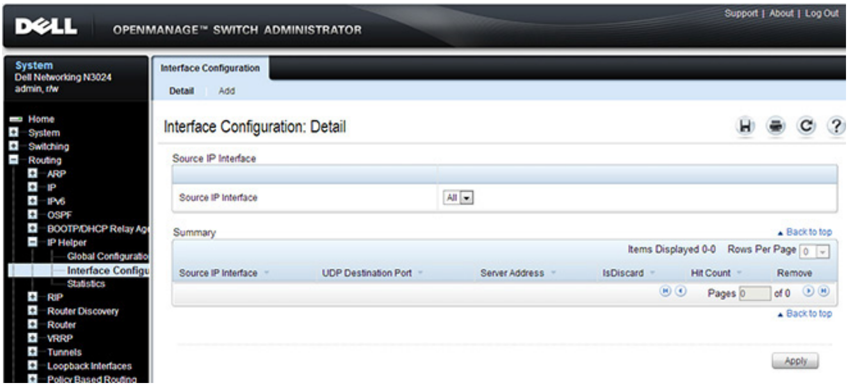
   The UDP/Helper Relay is added and the device is updated.

IP Helper Interface Configuration

Use the **Interface Configuration** page to add, show, or delete UDP Relay and Helper IP configuration for a specific interface.

To display the page, click **Routing** → **IP Helper** → **Interface Configuration** in the navigation panel.

Figure 34-10. IP Helper Interface Configuration

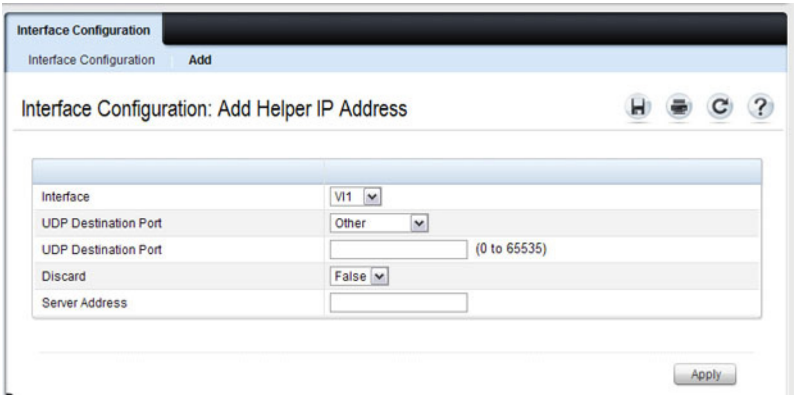


Adding an IP Helper Entry to an Interface

To add an IP helper entry to an interface:

1. Open the IP Helper **Interface Configuration** page.
2. Click **Add** to display the **Add IP Helper Address** page:

Figure 34-11. Add Helper IP Address



3. Select the interface to use for the relay.
4. Select a UDP Destination port name from the menu or enter the UDP Destination Port ID. Select the Default Set to configure for the relay entry for the default set of protocols.

   NOTE: If the DefaultSet option is specified, the device by default forwards UDP Broadcast packets for the following services: IEN-116 Name Service (port 42), DNS (port 53), NetBIOS Name Server (port 137), NetBIOS Datagram Server (port 138), TACACS Server (Port 49), and Time Service (port 37).

5. Choose whether to discard (True) or keep (False) packets arriving on the given interface with the given destination UDP port.
6. Enter the IP address of the server to which the packets with the given UDP Destination Port will be relayed.
7. Click **Apply**.

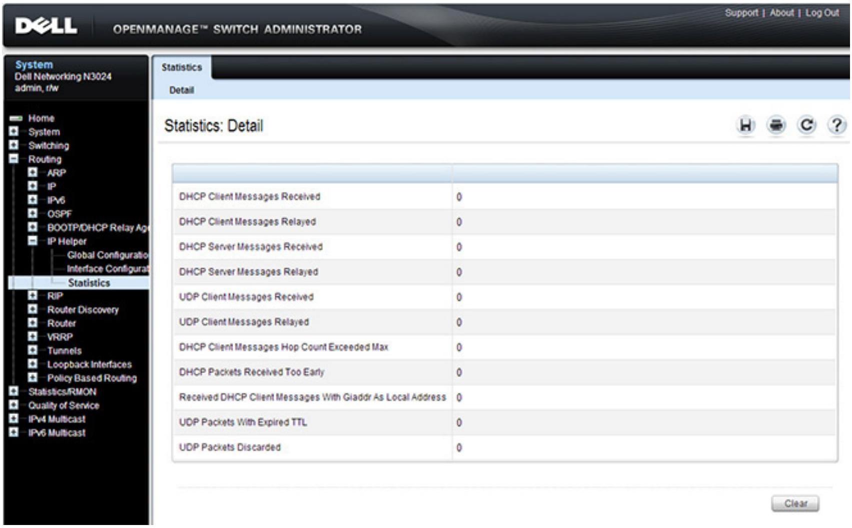
   The UDP/Helper Relay is added to the interface and the device is updated.

IP Helper Statistics

Use the **Statistics** page to view UDP Relay Statistics for the switch.

To display the page, click **Routing** → **IP Helper** → **Statistics** in the navigation panel.

Figure 34-12. IP Helper Statistics



Configuring L2 and L3 Relay Features (CLI)

This section provides information about the commands used for configuring L2 and L3 relay features on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring L2 DHCP Relay

Beginning in Privileged EXEC mode, use the following commands to configure switch and interface L2 DHCP relay settings.

Command	Purpose
configure	Enter global configuration mode.
dhcp l2relay	Globally enable L2 DHCP relay on the switch
interface *interface*	Enter interface configuration mode for the specified port or LAG. The *interface* variable includes the interface type and number, for example **tengigabitethernet 1/0/3**. For a LAG, the interface type is **port-channel**. A range of ports can be specified using the **interface range** command. For example, **interface range tengigabitethernet 1/0/8-12** configures interfaces 8, 9, 10, 11, and 12.
dhcp l2relay	Enable L2 DHCP relay on the port(s) or LAG(s).
dhcp l2relay trust	Configure the interface(s) to mandate Option-82 on receiving DHCP packets.
exit	Exit to Global Configuration mode.
dhcp l2relay vlan *vlan-range*	Enable the L2 DHCP Relay agent for a set of VLANs. All DHCP packets which arrive on interfaces in the configured VLAN are subject to L2 Relay processing.
dhcp l2relay circuit-id vlan *vlan-range*	Enable setting the DHCP Option 82 Circuit ID for a VLAN. When enabled, the interface number is added as the Circuit ID in DHCP option 82.

Command	Purpose
dhcp l2relay remote-id *remoteId* **vlan** *vlan-range*	Enable setting the DHCP Option 82 Remote ID for a VLAN. When enabled, the supplied string is used for the Remote ID in DHCP Option 82. The *remoteId* variable is a string to be used as the remote ID in the Option 82 (Range: 1 - 128 characters).
exit	Exit to Privileged EXEC mode.
show dhcp l2relay all	View L2 DHCP relay settings on the switch.
show dhcp l2relay interface [**all** | *interface*]	View L2 DHCP relay settings for all interfaces or for the specified interface.
show dhcp l2relay vlan *vlan-range*	View L2 DHCP relay settings for the specified VLAN
show dhcp l2relay stats interface [**all** | *interface*]	View the number of DHCP packets processed and relayed by the L2 relay agent. To reset the statistics to 0, use the **clear dhcp l2relay statistics interface** [**all** | *interface*] **command.**
show dhcp l2relay agent-option vlan *vlan-id*	View the DHCP L2 Relay Option-82 configuration for the specified VLAN.
show dhcp l2relay circuit-id vlan *vlan-id*	View the DHCP L2 Relay circuit ID configuration for the specified VLAN.
show dhcp l2relay remote-id vlan *vlan-id*	View the DHCP L2 Relay remote ID configuration for the specified VLAN.

Configuring L3 Relay (IP Helper) Settings

Beginning in Privileged EXEC mode, use the following commands to configure switch and interface L3 DHCP relay and IP helper settings.

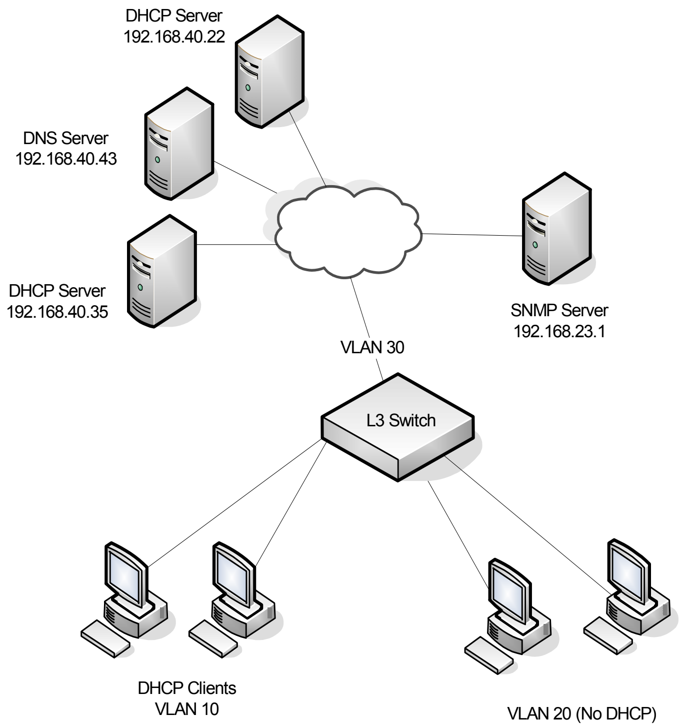
Command	Purpose
configure	Enter global configuration mode.
ip helper enable	Use this command to enable the IP helper feature. It is enabled by default.
ip helper-address *server-address* [*dest-udp-port* | **dhcp** | **domain** | **isakmp** | **mobile-ip** | **nameserver** | **netbios-dgm** | **netbios-ns** | **ntp** | **pim-auto-rp** | **rip** | **tacacs** | **tftp** | **time**]	Configure the relay of certain UDP broadcast packets received on any interface. Specify the one of the protocols defined in the command or the UDP port number. • *server-address* — The IPv4 unicast or directed broadcast address to which relayed UDP broadcast packets are sent. The server address cannot be an IP address configured on any interface of the local router. • *dest-udp-port* — A destination UDP port number from 0 to 65535.
interface vlan *vlan-id*	Enter interface configuration mode for the specified VLAN routing interface. A range of VLAN routing interfaces can be specified using the **interface range vlan** command. For example, **interface range vlan 10,20,30** configures VLAN interfaces 10, 20, and 30. **NOTE:** All VLANs must be configured as VLAN routing interfaces.

Command	Purpose
ip helper-address {*server-address* | **discard**} [*dest-udp-port* | **dhcp** | **domain** | **isakmp** | **mobile-ip** | **nameserver** | **netbios-dgm** | **netbios-ns** | **ntp** | **pim-auto-rp** | **rip** | **tacacs** | **tftp** | **time**]	Configure the relay of certain UDP broadcast packets received on the VLAN routing interface(s). This command takes precedence over an ip helper-address command given in global configuration mode. Specify the one of the protocols defined in the command or the UDP port number. • *server-address* — The IPv4 unicast or directed broadcast address to which relayed UDP broadcast packets are sent. The server address cannot be an IP address configured on any interface of the local router. • *dest-udp-port* — A destination UDP port number from 0 to 65535.
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ip helper-address [**vlan** *vlan-id*]	View IP helper (L3 relay) settings for all interfaces or for the specified VLAN routing interface.
show ip helper statistics	View the number of DHCP and other UDP packets processed and relayed by the UDP relay agent. To reset the statistics to 0, use the **clear ip helper statistics** command.

Relay Agent Configuration Example

The example in this section shows how to configure the L3 relay agent (IP helper) to relay and discard various protocols.

Figure 34-13. L3 Relay Network Diagram



This example assumes that multiple VLAN routing interfaces have been created, and configured with IP addresses.

To configure the switch:

1 Relay DHCP packets received on VLAN 10 to 192.168.40.35

```
console#config
console(config)#interface vlan 10
console(config-if-vlan10)#ip helper-address 192.168.40.35 dhcp
```

2 Relay DNS packets received on VLAN 10 to 192.168.40.43

```
console(config-if-vlan10)#ip helper-address 192.168.40.35
domain
console(config-if-vlan10)#exit
```

3 Relay SNMP traps (port 162) received on VLAN 20 to 192.168.23.1

```
console(config)#interface vlan 20
console(config-if-vlan20)#ip helper-address 192.168.23.1 162
```

4 The clients on VLAN 20 have statically-configured network information, so the switch is configured to drop DHCP packets received on VLAN 20

```
console(config-if-vlan20)#ip helper-address discard dhcp
console(config-if-vlan20)#exit
```

5 DHCP packets received from clients in any VLAN other than VLAN 10 and VLAN 20 are relayed to 192.168.40.22.

NOTE: The following command is issued in Global Configuration mode, so it applies to all interfaces except VLAN 10 and VLAN 20. IP helper commands issued in Interface Configuration mode override the commands issued in Global Configuration Mode.

```
console(config)#ip helper-address 192.168.40.22 dhcp
```

6 Verify the configuration.

```
console#show ip helper-address

IP helper is enabled

I/F      UDP Port     Discard  Hit Count    Server Address
----     --------     -----    ---------    ---------------
Vl10     domain       No       0            192.168.40.43
Vl10     dhcp         No       0            192.168.40.35
Vl20     dhcp         Yes      0
Vl20     162          No       0            192.168.23.1
Any      dhcp         No       0            192.168.40.22
```

35

OSPF and OSPFv3

Dell Networking N2000, N3000, and N4000 Series Switches

This chapter describes how to configure Open Shortest Path First (OSPF) and OSPFv3. OSPF is a dynamic routing protocol for IPv4 networks, and OSPFv3 is used to route traffic in IPv6 networks. The protocols are configured separately within the software, but their functionality is largely similar for IPv4 and IPv6 networks.

NOTE: In this chapter references to OSPF apply to OSPFv2 and OSPFv3 unless otherwise noted.

NOTE: Dell Networking N1500 Series switches do not support OSPF.

NOTE: Dell Networking N2000 Series switches do not support OSPFv3.

The topics covered in this chapter include:

- OSPF Overview
- OSPF Feature Details
- Default OSPF Values
- Configuring OSPF Features (Web)
- Configuring OSPFv3 Features (Web)
- Configuring OSPF Features (CLI)
- Configuring OSPFv3 Features (CLI)
- OSPF Configuration Examples
- Configuring OSPF VRFs

OSPF Overview

OSPF is an Interior Gateway Protocol (IGP) that performs dynamic routing within a network. Dell Networking N-Series switches support two dynamic routing protocols: OSPF and Routing Information Protocol (RIP).

Unlike RIP, OSPF is a link-state protocol. Larger networks typically use the OSPF protocol instead of RIP.

What Are OSPF Areas and Other OSPF Topology Features?

The top level of the hierarchy of an OSPF network is known as an OSPF domain. The domain can be divided into areas. Routers within an area must share detailed information on the topology of their area, but require less detailed information about the topology of other areas. Segregating a network into areas enables limiting the amount of route information communicated throughout the network.

Areas are identified by a numeric ID in IP address format n.n.n.n (note, however, that these are not used as actual IP addresses). For simplicity, the area can be configured and referred to in normal integer notation. For example, Area 20 is identified as 0.0.0.20 and Area 256 as 0.0.1.0. The area identified as 0.0.0.0 is referred to as Area 0 and is considered the OSPF backbone. All other OSPF areas in the network must connect to Area 0 directly or through a virtual link. The backbone area is responsible for distributing routing information between non-backbone areas.

A virtual link can be used to connect an area to Area 0 when a direct link is not possible. A virtual link traverses an area between the remote area and Area 0.

A stub area is an area that does not accept external LSAs (LSAs generated by redistributing routes) that were learned from a protocol other than OSPF or were statically configured. These routes typically send traffic outside the AS. Therefore, routes from a stub area to locations outside the AS use the default gateway. A virtual link cannot be configured across a stub area. A Not So Stubby Area can import limited external routes only from a connected ASBR.

What Are OSPF Routers and LSAs?

When a Dell Networking N-Series switch is configured to use OSPF for dynamic routing, it is considered to be an OSPF router. OSPF routers keep track of the state of the various links they send data to. Routers exchange OSPF link state advertisements (LSAs) with other routers. External LSAs provide information on static routes or routes learned from other routing protocols.

OSPF defines various router types:

- Backbone routers have an interface in Area 0.
- Area border routers (ABRs) have interfaces in multiple areas.
- Internal routers have all their interfaces in a single OSPF area.
- Autonomous system boundary routers (ASBRs) redistribute routes from other protocols and originate external LSAs.

How Are Routes Selected?

OSPF determines the best route using the route metric and the type of the OSPF route. The following order is used for choosing a route if more than one type of route exists:

1. Intra-area (the destination prefix is in the same area as the router computing the route)
2. Inter-area (the destination is not in the same area as the router computing the route
3. External Type 1
4. External Type 2

How Are OSPF and OSPFv3 Different?

OSPFv3 is the Open Shortest Path First routing protocol for IPv6. It is similar to OSPFv2 in its concept of a link state database, intra/inter area, and AS external routes and virtual links. It differs from its IPv4 counterpart in a number of respects. Peering is done through link-local addresses, and the protocol is link rather than network centric; and addressing semantics have been moved to leaf LSAs.

OSPF Feature Details

This section provides details on the following OSPF features:

- Max Metric
- Static Area Range Cost
- LSA Pacing
- LSA Pacing

Max Metric

RFC 3137 introduced stub router behavior to OSPFv2. As a stub, a router can inform other routers that it is not available to forward data packets. This can be useful if OSPF has run out of resources (for example, memory) to compute a complete routing table, or to avoid routing transients as OSPF learns its neighbors and a complete set of routes at startup. Thus, OSPF can enter stub router mode either automatically (as a result of a resource condition) or by configuration.

When OSPF enters stub router mode, it re-originates its router LSAs and sets the metric on each of its non-stub links to the maximum value, 0xFFFF. Whenever OSPF originates a router LSA while in stub router mode, it sets the metrics in this way. Stub router mode is global and applies to router LSAs for all areas. Other routers prefer alternate paths that avoid the stub router; however, if no alternate path is available, another router may compute a transit route through a stub router. Because the stub router does not adjust the metric for stub links in its router LSA, routes to destinations on these networks are unaffected. Thus, stub router mode does not affect management connections to the router, even if the router and management station depend on OSPF routes to communicate with each other.

The feature supports two modes of operation. The network administrator can put OSPF in stub router mode. OSPF remains in stub router mode until the network administrator takes OSPF out of stub router mode. Alternatively, the network administrator can configure OSPF to start in stub router mode for a configurable period of time after the router boots up. On a stack, the startup period also applies when a unit takes over as the management unit. The **clear configuration** command also restarts OSPF in stub router mode. OSPF does not begin in stub router mode when OSPF is globally enabled. If the operator wants to avoid routing transients when he enables or configures OSPF, he can manually set OSPF in stub router mode.

If OSPF is in startup stub router mode and encounters a resource limitation that would normally cause OSPF to become a stub router, OSPF cancels the timer to exit startup stub router and remains in stub router mode until the network administrator takes action.

The network administrator can optionally configure OSPF to override the metric in summary LSAs while in stub router mode. The option applies to both type 3 and type 4 summary LSAs.

When a router is in stub router mode, all its virtual links are down. This is because the cost to the virtual neighbor is guaranteed to be greater than or equal to 0xFFFF. RFC 2328 section 15 states that:

> "...a virtual link whose underlying path has cost greater than hexadecimal 0xffff (the maximum size of an interface cost in a router-LSA) should be considered non-operational."

To configure a router for stub router mode, use the **max-metric router-lsa** command in Global Router Configuration mode. The following example sets the router to start in stub router mode on a restart and remain in stub router mode for 5 minutes:

```
ABR-R0(config)#router ospf
ABR-R0(config-router)#max-metric router-lsa on-startup 300
```

The following example sets the router to advertise the metric in type 3 and type 4 summary LSAs as 32768 for 5 minutes after a restart, after which time the router will exit stub router mode and advertise the full set of LSAs:

```
ABR-R0(config)#router ospf
ABR-R0(config-router)#max-metric router-lsa on-startup 300 summary-
lsa 32768
```

The following example causes the router to exit stub router mode, whether entered automatically due to resource constraints or due to configuration by the operator. Virtual links are enabled when the router exits stub router mode.

```
ABR-R0(config)#router ospf
ABR-R0(config-router)#no max-metric router-lsa
```

Static Area Range Cost

This feature allows a network operator to configure a fixed OSPF cost that is always advertised when an area range is active. This feature applies to both OSPFv2 and OSPFv3.

An OSPF domain can be divided into areas to limit the processing required on each router. Area Border Routers (ABRs) advertise reachability across area boundaries. It is common to summarize the set of prefixes that an ABR advertises across an area boundary. RFC 2328 specifies that when an ABR originates a type 3 LSA for an active area range, the cost in the LSA is set to "the largest cost of any of the component networks." Thus, when an area's topology changes in a way that increases the largest cost, the type 3 LSA must be re-originated. In some cases, advertising the change in cost may be less important than preventing the topology change from propagating outside the area (thus causing routers in other areas to process and flood a changed LSA and rerun their routing table calculations). For this reason, it is common to give the network administrator the option of configuring the cost for an area range. When a static cost is configured, the cost advertised in the type 3 LSA does not depend on the cost of the component networks. Thus, topology changes within an area do not propagate outside the area, resulting in greater stability within the OSPF domain.

Dell Networking N-Series switches also use area ranges to summarize type 7 LSAs when they are translated to type 5 LSAs. The cost option may be configured on area ranges used for type 7 to type 5 translation.

If an area range is configured for type 3 summarization and the static cost is set to the maximum value, 16,777,215, the range is not advertised. Setting this static cost is equivalent to configuring a range with the not-advertise option. A summary LSA with this metric (LSInfinity) cannot be advertised, according to RFC 2328 section 12.4.3. This behavior is consistent with the industry standard.

If an area range is configured for type 7 to type 5 translation, a type 5 LSA is sent if the metric is set to 16,777,215; however, other routers will not compute a route from a type 5 LSA with this metric.

See "Configuring the Static Area Range Cost " on page 1259 for a configuration example.

LSA Pacing

OSPF refreshes each self-originated LSA every 30 minutes. Because a router tends to originate many LSAs at the same time, either at startup or when adjacencies are formed or when routes are first learned, LSA refreshes tend to be grouped. Further, Area Border Routers (ABRs) attached to the same area tend to originate summary LSAs into the area at the same time. This behavior leads to periodic bursts of LS Update packets. Update bursts can lead to high CPU utilization, packet loss, and retransmission, if a receiver cannot absorb all packets in a burst. These losses occur primarily in two places: 1) at the Class of Service (CoS) queue where the hardware queues packets to the CPU, and 2) when a message buffer is allocated for an incoming packet.

This feature makes changes to OSPFv2 to improve the efficiency of LSA flooding, with the expectation that the improvements will greatly reduce or eliminate the packet drops caused by bursts in OSPF control packets. The changes are as follows:

- Introduce LSA transmit pacing, limiting the rate of LS Update packets that OSPF can send
- Introduce LSA refresh groups, so that OSPF efficiently bundles LSAs into LS Update packets when periodically refreshing self-originated LSAs

To configure LSA transmit pacing, use the timers pacing flood command in router config mode:

```
ABR-R0(config)#router ospf
ABR-R0(config-router)#timers pacing flood 50
```

This will cause LSA Update packets to be sent at no less than a 50 millisecond interval.

When OSPF refreshes LSAs, it considers all self-originated LSAs whose age is from 1800 to 1800 plus the pacing group size. Grouping LSAs for refresh allows OSPF to combine refreshed LSAs into a minimal number of LS Update packets. Minimizing the number of Update packets makes LSA distribution more efficient. To configure an LSA Refresh window, use the timers pacing lsa-group command in router-config mode:

```
ABR-R0(config)#router ospf
ABR-R0(config-router)#timers pacing lsa-group 300
```

This sets the LSA Refresh window to 2100 seconds or about 35 minutes.

Flood Blocking

OSPF is a link state routing protocol. Routers describe their local environment in Link State Advertisements (LSAs), which are distributed throughout an area or OSPF domain. Through this process, each router learns enough information to compute a set of routes consistent with the routes computed by all other routers.

Normally, OSPF floods an LSA on all interfaces within the LSA's flooding scope. Flooding ensures that all routers receive all LSAs. A router normally receives a duplicate copy of each LSA once on each interface in the LSA's flooding scope. The duplicate deliveries make OSPF LSA distribution robust, but in highly interconnected networks, can cause a lot of buffer and CPU usage. Buffer and CPU use can be reduced by selectively blocking LSA flooding on some interfaces, while ensuring that LSAs are flooded on enough interfaces to guarantee delivery of all LSAs to all routers. When enabling flood blocking, the network administrator must ensure there is sufficient LSA flooding even when there are router and link failures.

This feature enables a network administrator to disable LSA flooding on an interface. Flood blocking only affects flooding of LSAs with area or AS (i.e., domain-wide) scope. Such LSAs are expected to be flooded to neighbors on other, unblocked interfaces, and eventually reach neighbors on blocked interfaces. An LSA with interface flooding scope cannot be blocked; there is no other way for interface-scope LSAs to reach neighbors on the blocked interface. Allowing interface-scope LSAs on blocked interfaces allows graceful restart to work, even if the restarting router has neighbors on flood blocked interfaces.

When an interface is blocked, LSAs with area or AS scope are not sent to any neighbor on that interface. When flood blocking is enabled, OSPF does not advertise any LSAs with area or AS scope in its database description packets sent to neighbors on a blocked interface. When OSPF receives an LSA from a neighbor and the local database copy is newer than the received LSA, OSPF normally sends the newer LSA directly to the neighbor. If the neighbor is on a blocked interface, OSPF neither acknowledges the LSA nor sends the newer LSA. Instead, OSPF expects that the neighbor will receive the newer LSA indirectly.

Flooding is enabled by default.

Flood blocking cannot be enabled on virtual interfaces. While the feature could be allowed on virtual interfaces, it is less likely to be used on a virtual interface, since virtual interfaces are created specifically to allow flooding between two backbone routers. So the option of flood blocking on virtual interfaces is not supported.

See "Configuring Flood Blocking " on page 1264 for a configuration example.

MTU

OSPF database description packets announce the IP MTU of the interface where they are transmitted. Two routers form an OSPF adjacency only if their IP MTUs are the same. If OSPF receives a database description packet whose IP MTU is larger than the local IP MTU, it drops the packet. Adjacencies in this situation remain in Exchange Start state. A log message identifies the IP MTU mismatch:

```
<11> JAN 01 00:00:51 192.168.75.1-1 OSPF[175099648]:
spnbo.c(672) 12 %% Dropping a DD packet received on interface
0/1. DD MTU is 2000. Local MTU is 1500.
```

The administrator can configure OSPF to ignore MTU mismatches using the **ip ospf mtu-ignore** command.

Default OSPF Values

OSPF is globally enabled by default. To make it operational on the router, you must configure a router ID and enable OSPF on at least one interface.

Table 35-1 shows the global default values for OSPF and OSPFv3.

Table 35-1. OSPF/OSPFv3 Global Defaults

Parameter	Default Value
Router ID	None
Admin Mode	Enabled
RFC 1583 Compatibility	Enabled (OSPFv2 only)
ABR Status	Enabled
Opaque LSA Status	Enabled (OSPFv2 only)
Exit Overflow Interval	Not configured
SPF Delay Time	5 (OSPFv2 only)
SPF Hold Time	10 (OSPFv2 only)
External LSDB Limit	None
Default Metric	Not configured
Maximum Paths	4
AutoCost Reference Bandwidth	100 Mbps
Default Passive Setting	Disabled
Default Information Originate	Disabled
Non-Stop Forwarding (NSF) Support	Disabled

Table 35-2 shows the per-interface default values for OSPF and OSPFv3.

Table 35-2. OSPF Per-Interface Defaults

Parameter	Default Value
Admin Mode	Disabled
Advertise Secondaries	Enabled (OSPFv2 only)
Router Priority	1
Retransmit Interval	5 seconds
Hello Interval	10 seconds
Dead Interval	40 seconds
LSA Ack Interval	1 second
Interface Delay Interval	1 second
MTU Ignore	Disabled
Passive Mode	Disabled
Network Type	Broadcast
Authentication Type	None (OSPFv2 only)
Metric Cost	Not configured

Configuring OSPF Features (Web)

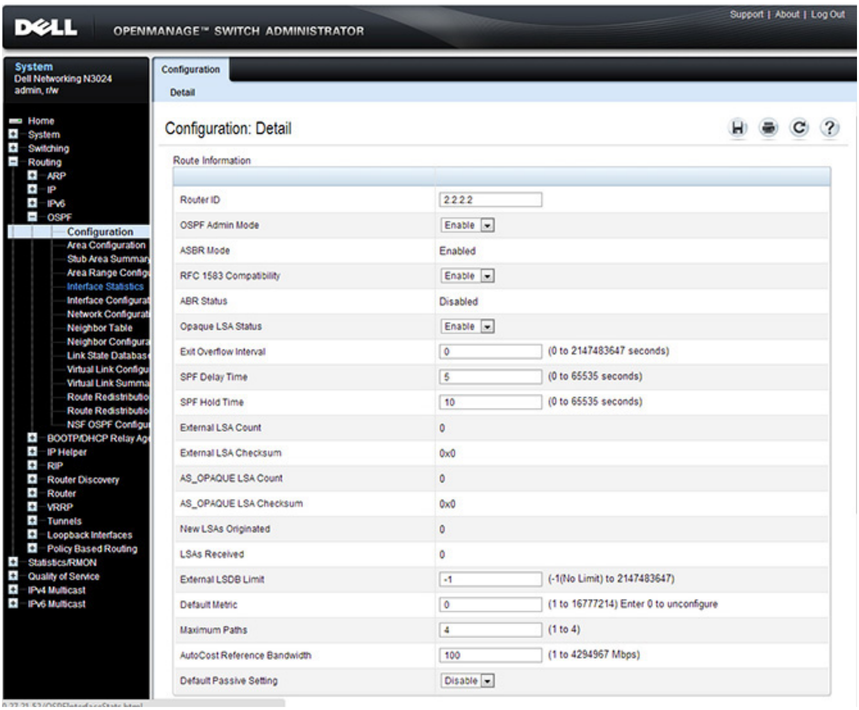
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring OSPF features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click at the top of the page.

OSPF Configuration

Use the **Configuration** page to enable OSPF on a router and to configure the related OSPF settings.

To display the page, click **Routing** → **OSPF** → **Configuration** in the navigation panel.

Figure 35-1. OSPF Configuration

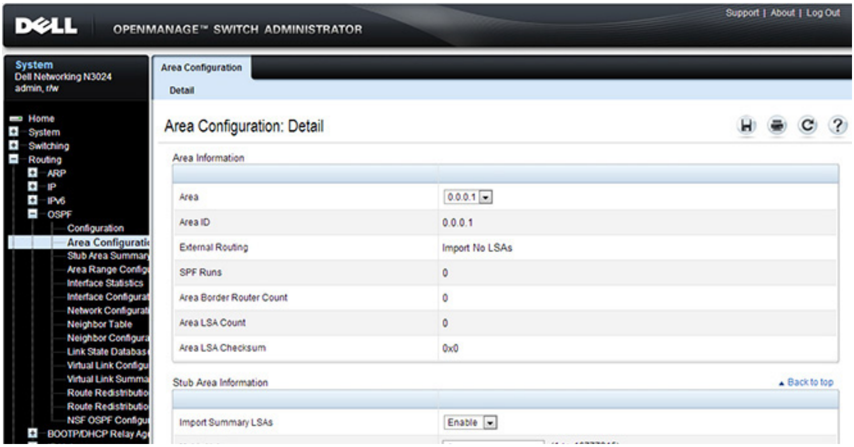


OSPF Area Configuration

The **Area Configuration** page lets you create a Stub area configuration and NSSA once you've enabled OSPF on an interface through **Routing** → **OSPF** → **Interface Configuration**. At least one router must have OSPF enabled for this web page to display.

To display the page, click **Routing** → **OSPF** → **Area Configuration** in the navigation panel. If a Stub Area has been created, the fields in the Stub Area Information are available. If a NSSA has been created, the fields in the NSSA Area Information are available.

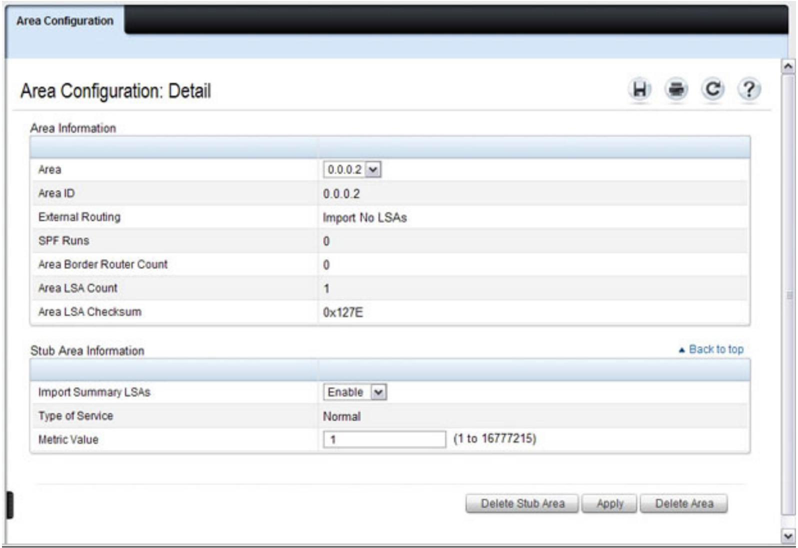
Figure 35-2. OSPF Area Configuration



Configuring an OSPF Stub Area

To configure the area as an OSPF stub area, click **Create Stub Area**. The pages refreshes, and displays additional fields that are specific to the stub area.

Figure 35-3. OSPF Stub Area Configuration

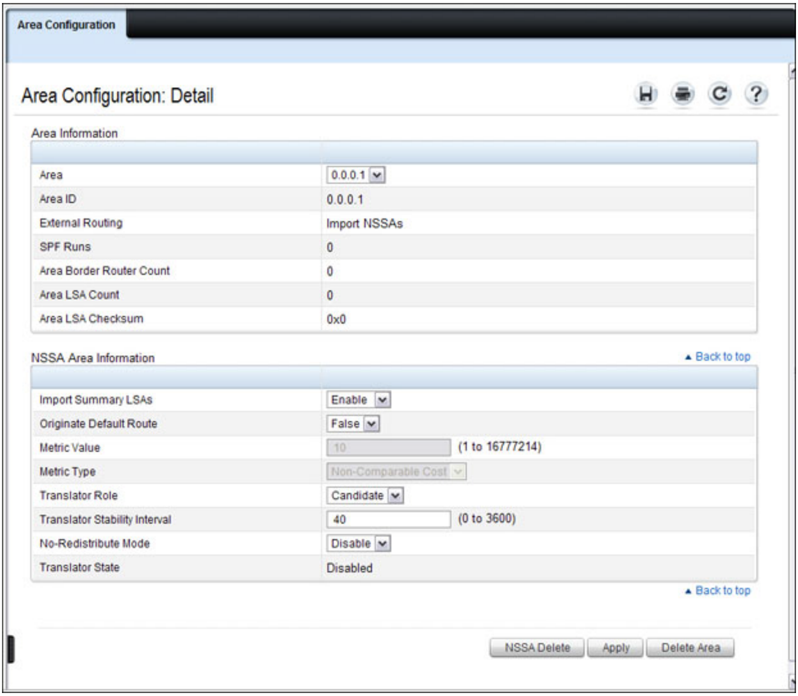


Use the **Delete Stub Area** button to remove the stub area.

Configuring an OSPF Not-So-Stubby Area

To configure the area as an OSPF not-so-stubby area (NSSA), click **NSSA Create**. The pages refreshes, and displays additional fields that are specific to the NSSA.

Figure 35-4. OSPF NSSA Configuration



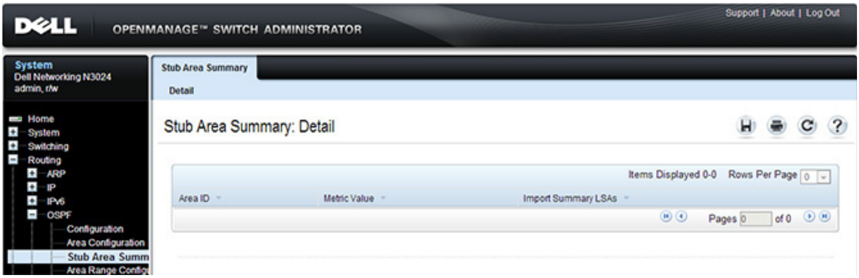
Use the **NSSA Delete** button to remove the NSSA area.

OSPF Stub Area Summary

The **Stub Area Summary** page displays OSPF stub area detail.

To display the page, click **Routing** → **OSPF** → **Stub Area Summary** in the navigation panel.

Figure 35-5. OSPF Stub Area Summary

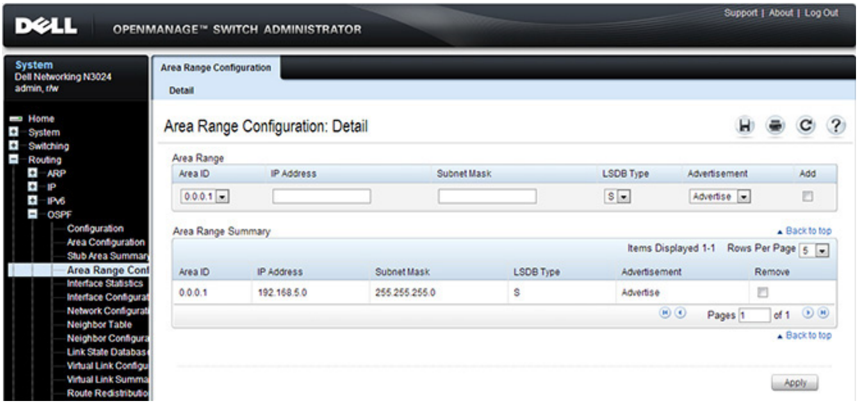


OSPF Area Range Configuration

Use the **Area Range Configuration** page to configure and display an area range for a specified NSSA.

To display the page, click **Routing** → **OSPF** → **Area Range Configuration** in the navigation panel.

Figure 35-6. OSPF Area Range Configuration

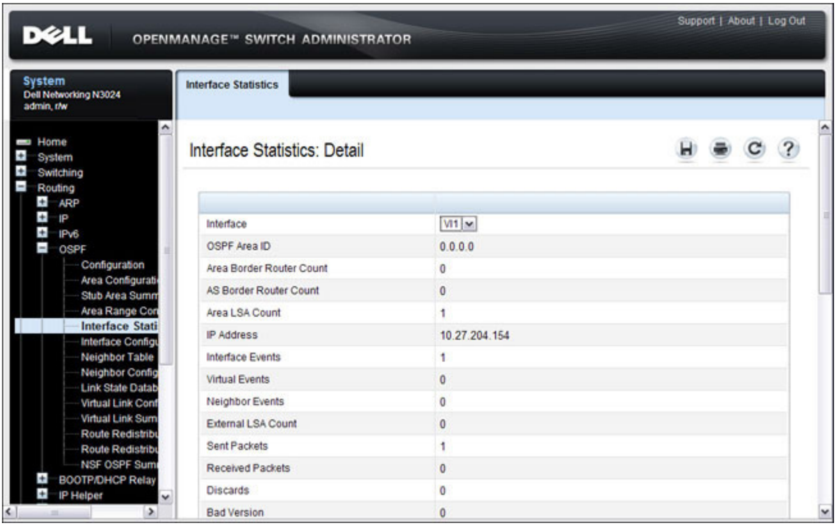


OSPF Interface Statistics

Use the **Interface Statistics** page to display statistics for the selected interface. The information is displayed only if OSPF is enabled.

To display the page, click **Routing** → **OSPF** → **Interface Statistics** in the navigation panel.

Figure 35-7. OSPF Interface Statistics

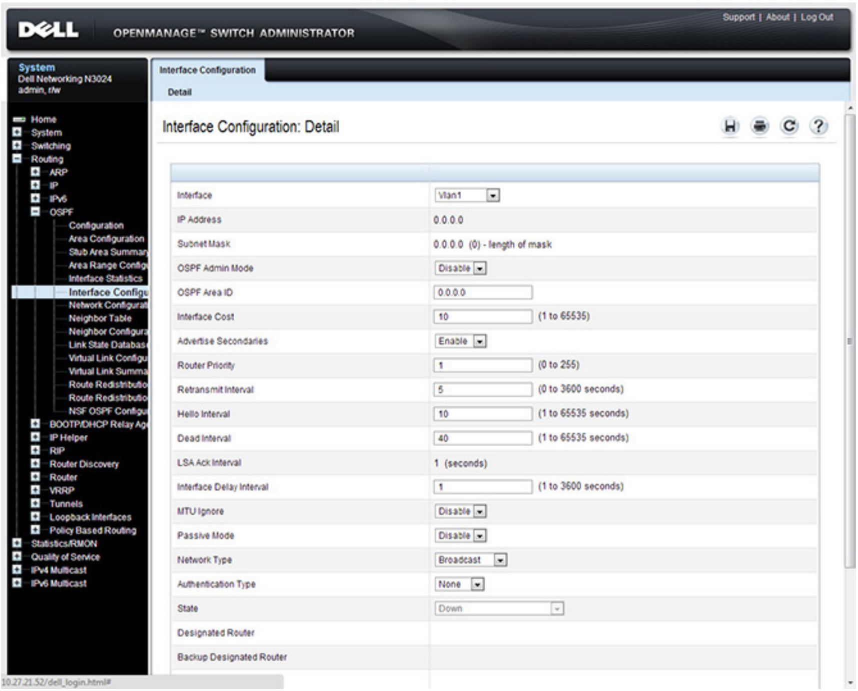


OSPF Interface Configuration

Use the **Interface Configuration** page to configure an OSPF interface.

To display the page, click **Routing** → **OSPF** → **Interface Configuration** in the navigation panel.

Figure 35-8. OSPF Interface Configuration

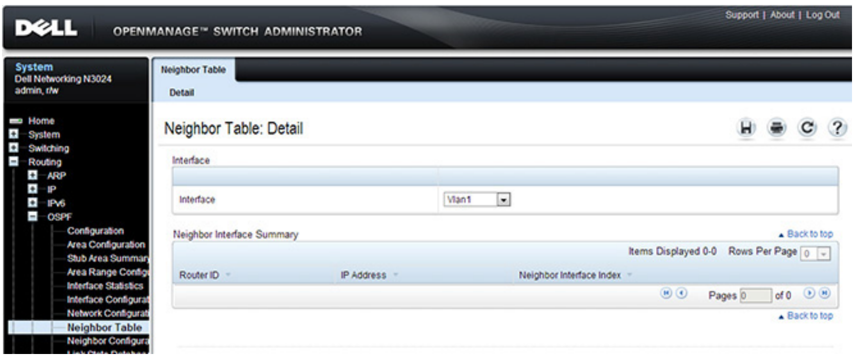


OSPF Neighbor Table

Use the **Neighbor Table** page to display the OSPF neighbor table list. When a particular neighbor ID is specified, detailed information about a neighbor is given. The information below is only displayed if OSPF is enabled.

To display the page, click **Routing** → **OSPF** → **Neighbor Table** in the navigation panel.

Figure 35-9. OSPF Neighbor Table

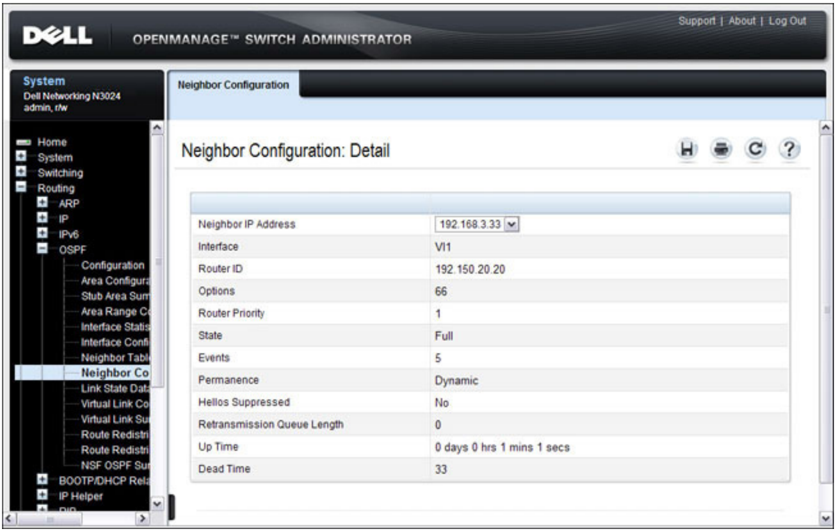


OSPF Neighbor Configuration

Use the **Neighbor Configuration** page to display the OSPF neighbor configuration for a selected neighbor ID. When a particular neighbor ID is specified, detailed information about a neighbor is given. The information below is only displayed if OSPF is enabled and the interface has a neighbor. The IP address is the IP address of the neighbor.

To display the page, click **Routing** → **OSPF** → **Neighbor Configuration** in the navigation panel.

Figure 35-10. OSPF Neighbor Configuration

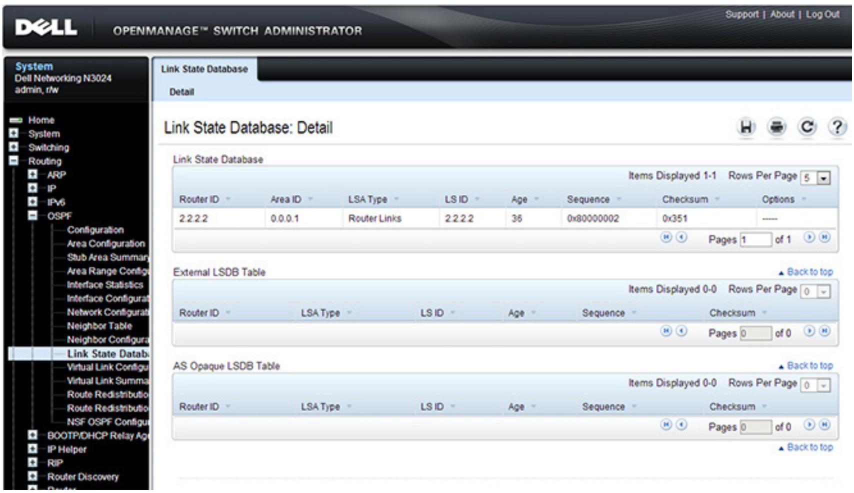


OSPF Link State Database

Use the **Link State Database** page to display OSPF link state, external LSDB table, and AS opaque LSDB table information.

To display the page, click **Routing** → **OSPF** → **Link State Database** in the navigation panel.

Figure 35-11. OSPF Link State Database

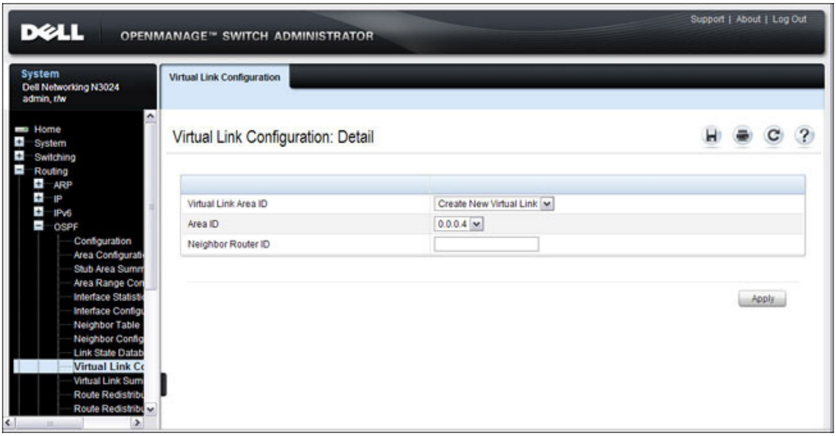


OSPF Virtual Link Configuration

Use the **Virtual Link Configuration** page to create or configure virtual interface information for a specific area and neighbor. A valid OSPF area must be configured before this page can be displayed.

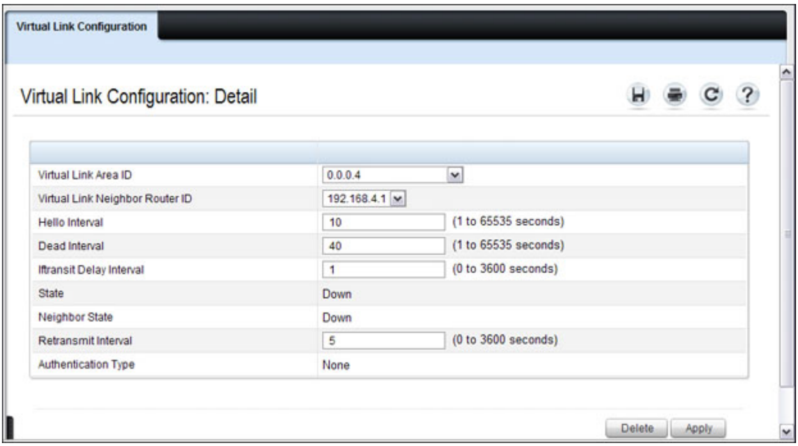
To display the page, click **Routing** → **OSPF** → **Virtual Link Configuration** in the navigation panel.

Figure 35-12. OSPF Virtual Link Creation



After you create a virtual link, additional fields display, as the Figure 35-13 shows.

Figure 35-13. OSPF Virtual Link Configuration

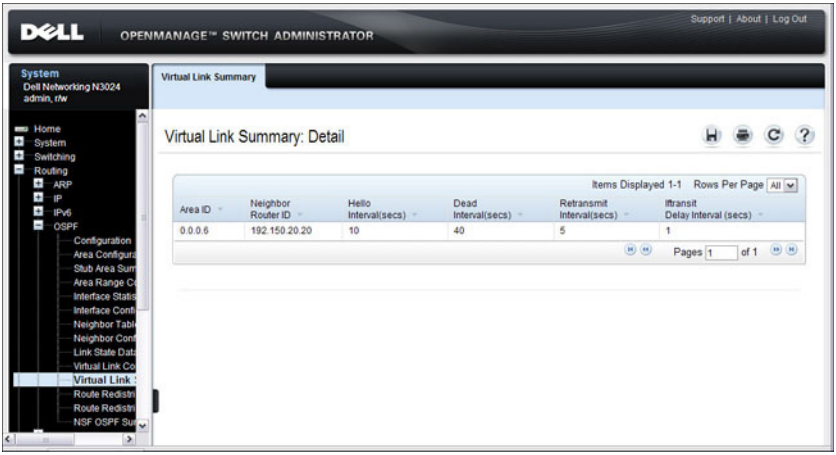


OSPF Virtual Link Summary

Use the **Virtual Link Summary** page to display all of the configured virtual links.

To display the page, click **Routing** → **OSPF** → **Virtual Link Summary** in the navigation panel.

Figure 35-14. OSPF Virtual Link Summary

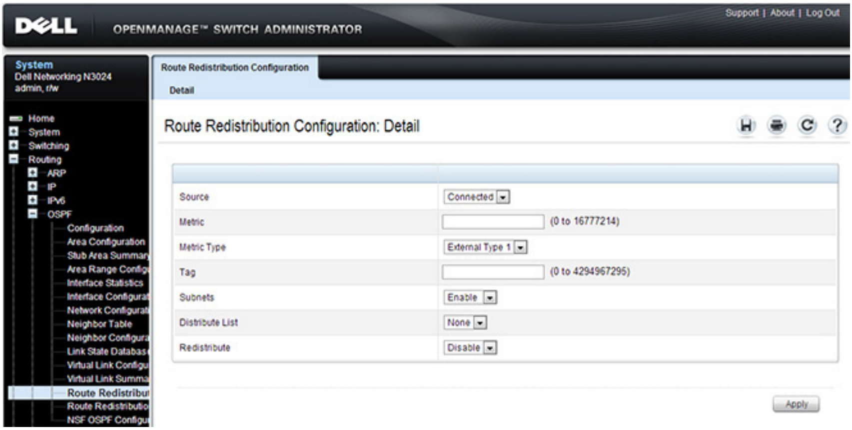


OSPF Route Redistribution Configuration

Use the **Route Redistribution Configuration** page to configure redistribution in OSPF for routes learned through various protocols. Routes learned from all available protocols, or from selected protocols, can be redistributed.

To display the page, click **Routing** → **OSPF** → **Route Redistribution Configuration** in the navigation panel.

Figure 35-15. OSPF Route Redistribution Configuration

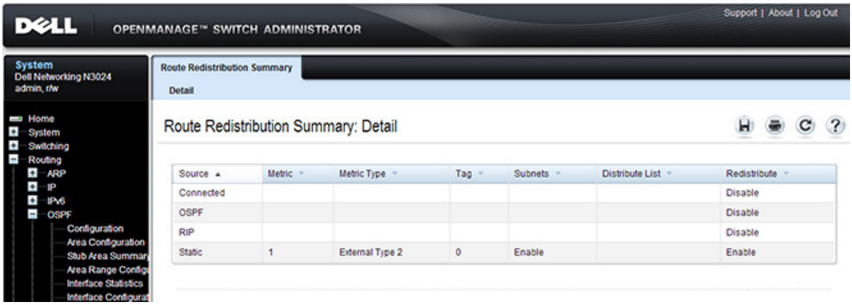


OSPF Route Redistribution Summary

Use the **Route Redistribution Summary** page to display OSPF Route Redistribution configurations.

To display the page, click **Routing** → **OSPF** → **Route Redistribution Summary** in the navigation panel.

Figure 35-16. OSPF Route Redistribution Summary

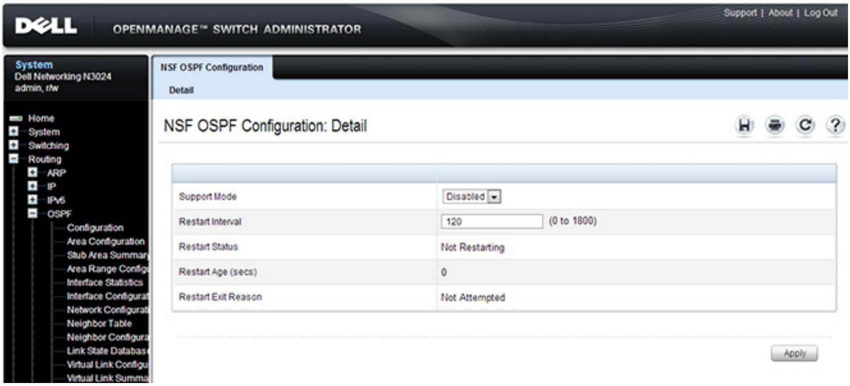


NSF OSPF Configuration

Use the **NSF OSPF Configuration** page to configure the non-stop forwarding (NSF) support mode and to view NSF summary information for the OSPF feature. NSF is a feature used in switch stacks to maintain switching and routing functions in the event of a stack unit failure. For information about NSF, see "What is Nonstop Forwarding? " on page 201 in the Stacking chapter.

To display the page, click **Routing** → **OSPF** → **NSF OSPF Configuration** in the navigation panel.

Figure 35-17. NSF OSPF Configuration



Configuring OSPFv3 Features (Web)

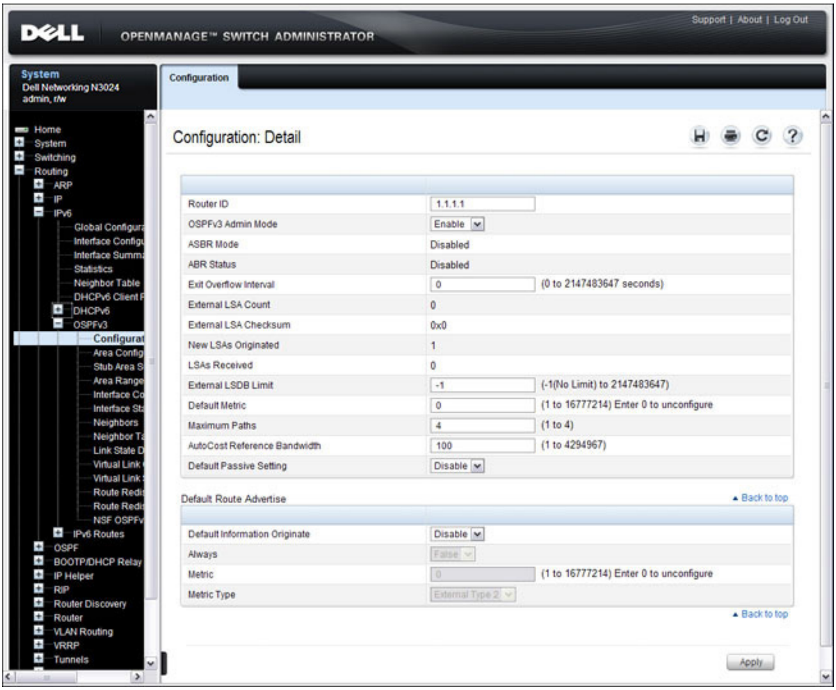
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring OSPFv3 features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

OSPFv3 Configuration

Use the **Configuration** page to activate and configure OSPFv3 for a switch.

To display the page, click **IPv6** → **OSPFv3** → **Configuration** in the navigation panel.

Figure 35-18. OSPFv3 Configuration

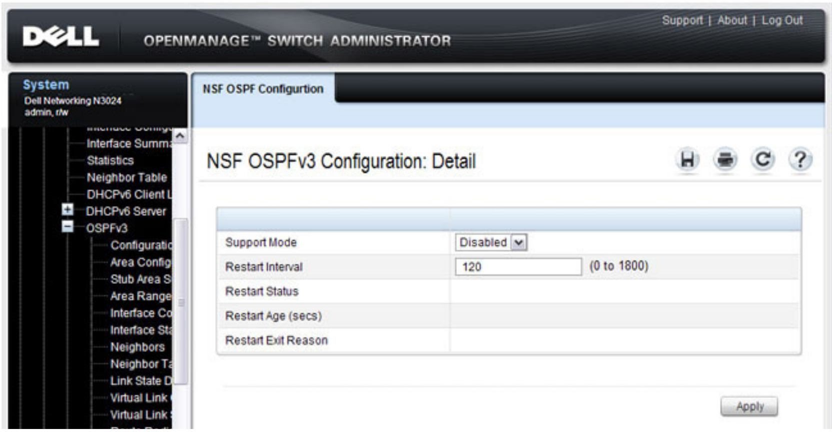


OSPFv3 Area Configuration

Use the **Area Configuration** page to create and configure an OSPFv3 area.

To display the page, click **IPv6** → **OSPFv3** → **Area Configuration** in the navigation panel.

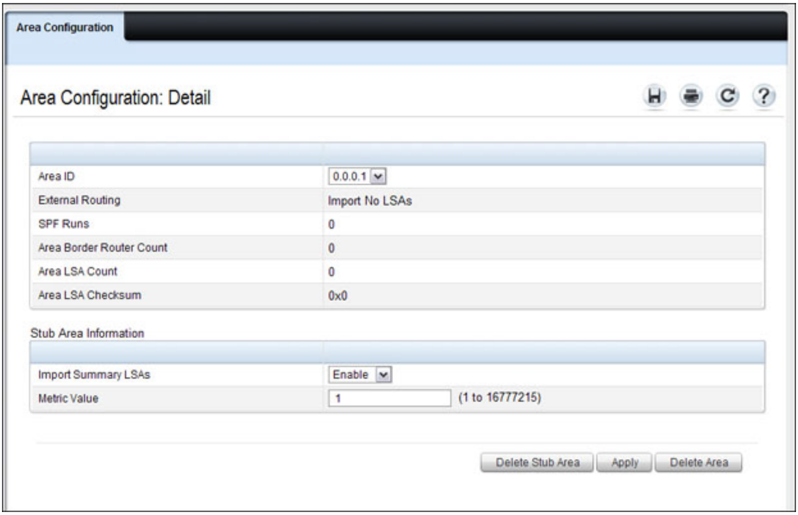
Figure 35-19. OSPFv3 Area Configuration



Configuring an OSPFv3 Stub Area

To configure the area as an OSPFv3 stub area, click **Create Stub Area**. The pages refreshes, and displays additional fields that are specific to the stub area.

Figure 35-20. OSPFv3 Stub Area Configuration

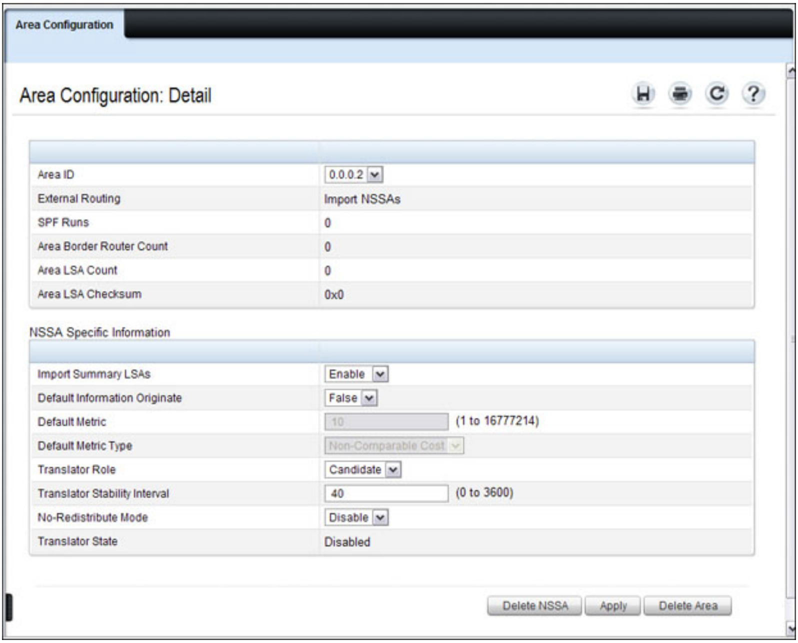


Use the **Delete Stub Area** button to remove the stub area.

Configuring an OSPFv3 Not-So-Stubby Area

To configure the area as an OSPFv3 not-so-stubby area (NSSA), click **Create NSSA**. The pages refreshes, and displays additional fields that are specific to the NSSA.

Figure 35-21. OSPFv3 NSSA Configuration



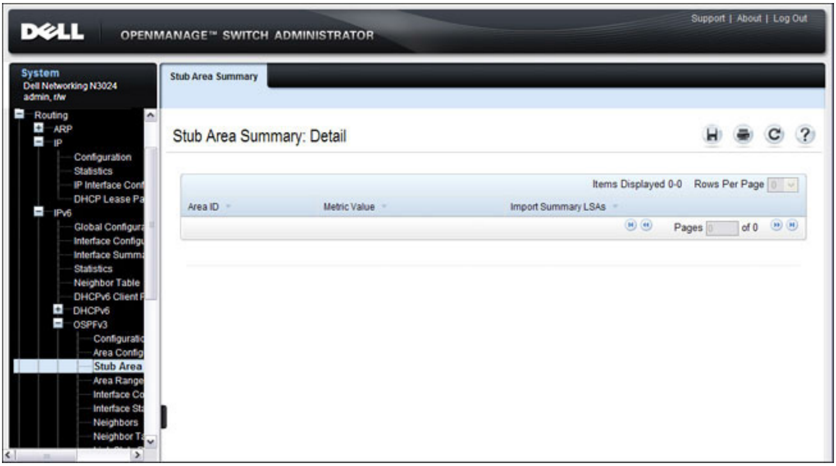
Use the **Delete NSSA** button to remove the NSSA area.

OSPFv3 Stub Area Summary

Use the **Stub Area Summary** page to display OSPFv3 stub area detail.

To display the page, click **IPv6** → **OSPFv3** → **Stub Area Summary** in the navigation panel.

Figure 35-22. OSPFv3 Stub Area Summary

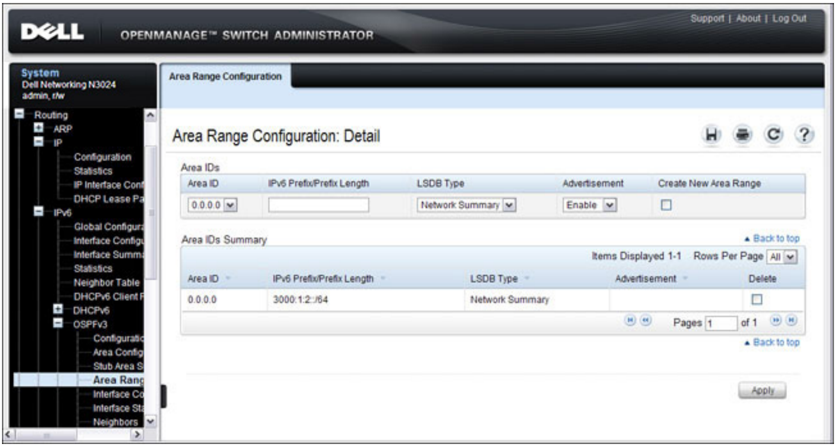


OSPFv3 Area Range Configuration

Use the **Area Range Configuration** page to configure OSPFv3 area ranges.

To display the page, click **IPv6** → **OSPFv3** → **Area Range Configuration** in the navigation panel.

Figure 35-23. OSPFv3 Area Range Configuration

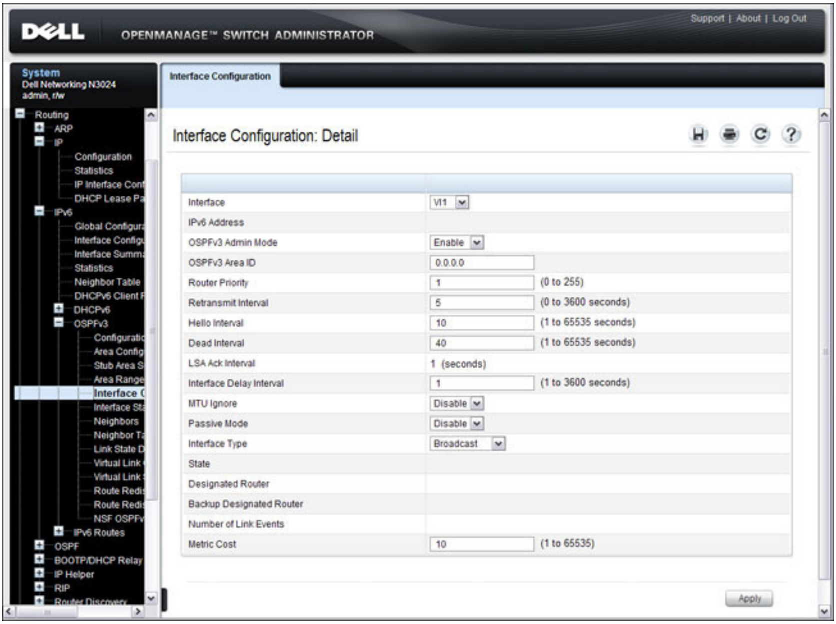


OSPFv3 Interface Configuration

Use the **Interface Configuration** page to create and configure OSPFv3 interfaces.

To display the page, click **IPv6** → **OSPFv3** → **Interface Configuration** in the navigation panel.

Figure 35-24. OSPFv3 Interface Configuration



OSPFv3 Interface Statistics

Use the **Interface Statistics** page to display OSPFv3 interface statistics. Information is only displayed if OSPF is enabled.

To display the page, click **IPv6** → **OSPFv3** → **Interface Statistics** in the navigation panel.

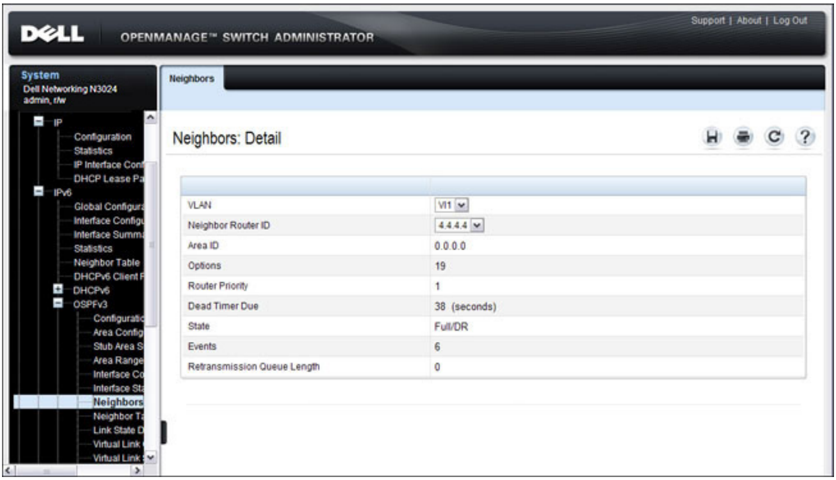
Figure 35-25. OSPFv3 Interface Statistics

OSPFv3 Neighbors

Use the **Neighbors** page to display the OSPF neighbor configuration for a selected neighbor ID. When a particular neighbor ID is specified, detailed information about that neighbor is given. Neighbor information only displays if OSPF is enabled and the interface has a neighbor. The IP address is the IP address of the neighbor.

To display the page, click **IPv6** → **OSPFv3** → **Neighbors** in the navigation panel.

Figure 35-26. OSPFv3 Neighbors

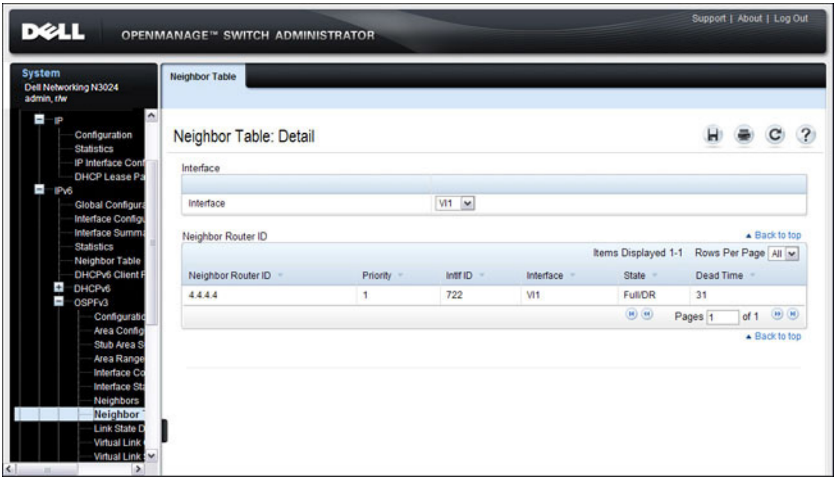


OSPFv3 Neighbor Table

Use the **Neighbor Table** page to display the OSPF neighbor table list. When a particular neighbor ID is specified, detailed information about a neighbor is given. The neighbor table is only displayed if OSPF is enabled.

To display the page, click **IPv6** → **OSPFv3** → **Neighbor Table** in the navigation panel.

Figure 35-27. OSPFv3 Neighbor Table

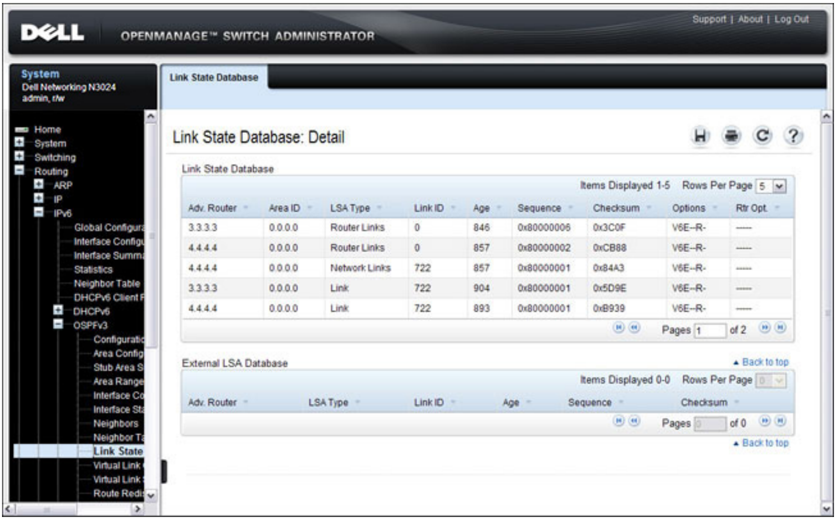


OSPFv3 Link State Database

Use the **Link State Database** page to display the link state and external LSA databases. The OSPFv3 **Link State Database** page has been updated to display external LSDB table information in addition to OSPFv3 link state information.

To display the page, click **IPv6 → OSPFv3 → Link State Database** in the navigation panel.

Figure 35-28. OSPFv3 Link State Database

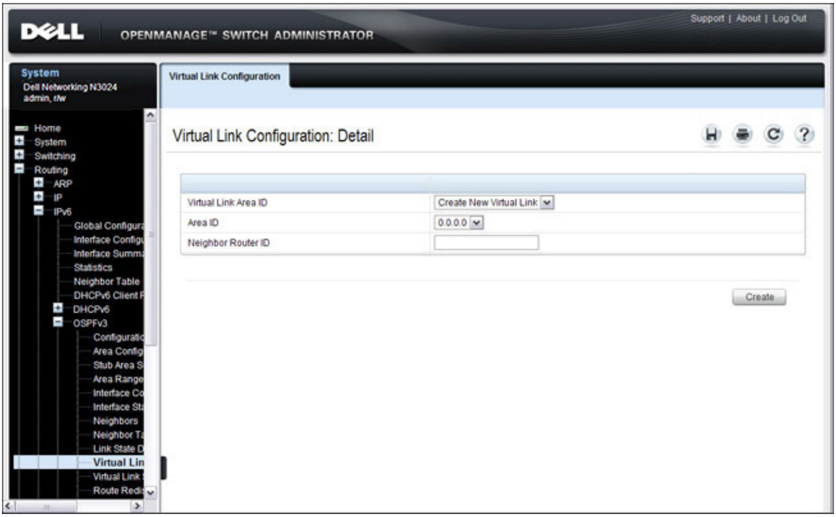


OSPFv3 Virtual Link Configuration

Use the **Virtual Link Configuration** page to define a new or configure an existing virtual link. To display this page, a valid OSPFv3 area must be defined through the OSPFv3 Area Configuration page.

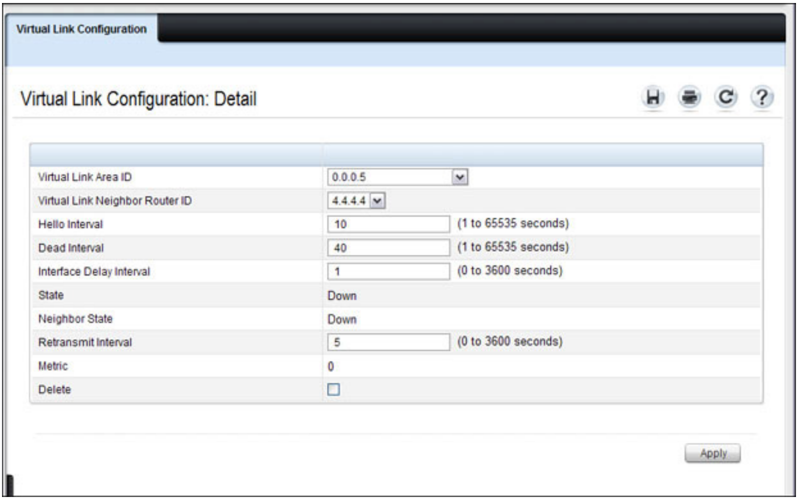
To display the page, click **IPv6** → **OSPFv3** → **Virtual Link Configuration** in the navigation panel.

Figure 35-29. OSPFv3 Virtual Link Configuration



After you create a virtual link, additional fields display, as the Figure 35-30 shows.

Figure 35-30. OSPFv3 Virtual Link Configuration

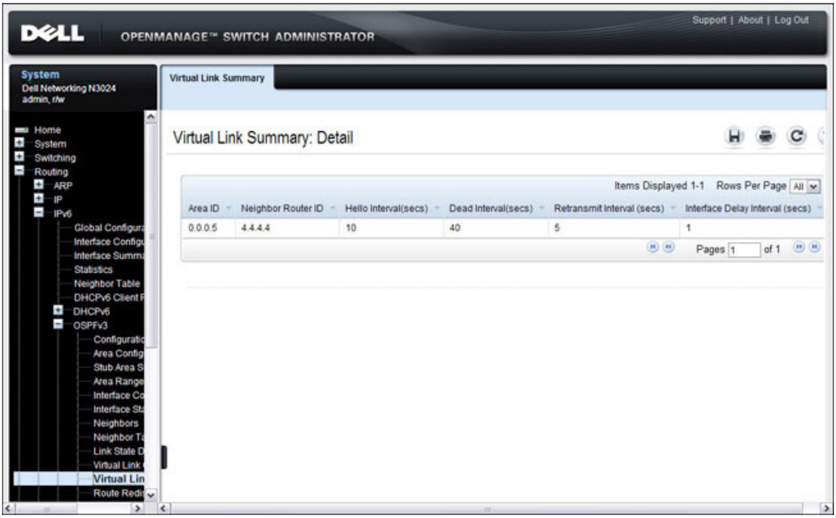


OSPFv3 Virtual Link Summary

Use the **Virtual Link Summary** page to display virtual link data by Area ID and Neighbor Router ID.

To display the page, click **IPv6** → **OSPFv3** → **Virtual Link Summary** in the navigation panel.

Figure 35-31. OSPFv3 Virtual Link Summary

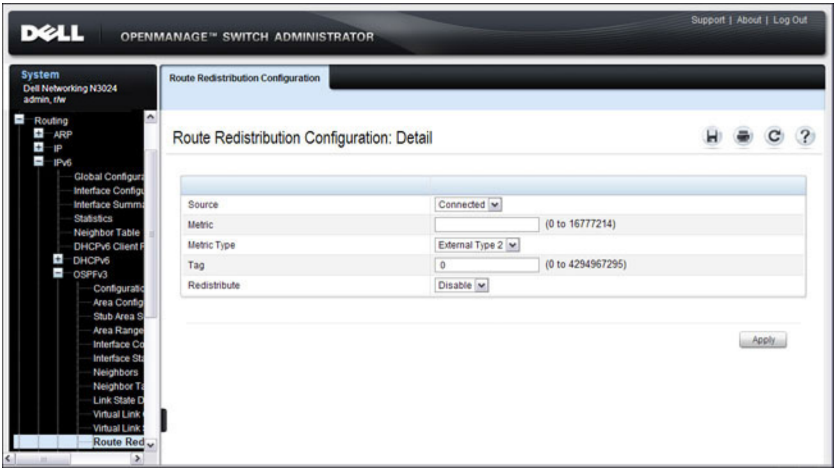


OSPFv3 Route Redistribution Configuration

Use the **Route Redistribution Configuration** page to configure route redistribution.

To display the page, click **IPv6 → OSPFv3 → Route Redistribution Configuration** in the navigation panel.

Figure 35-32. OSPFv3 Route Redistribution Configuration

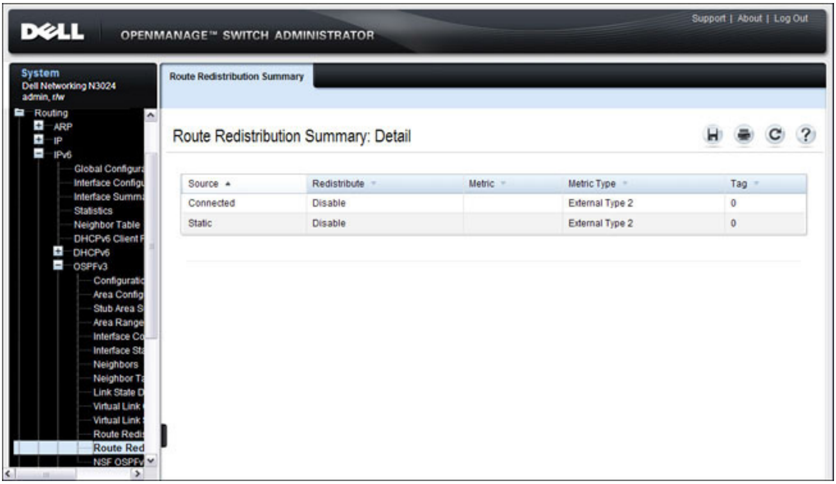


OSPFv3 Route Redistribution Summary

Use the **Route Redistribution Summary** page to display route redistribution settings by source.

To display the page, click **IPv6** → **OSPFv3** → **Route Redistribution Summary** in the navigation panel.

Figure 35-33. OSPFv3 Route Redistribution Summary

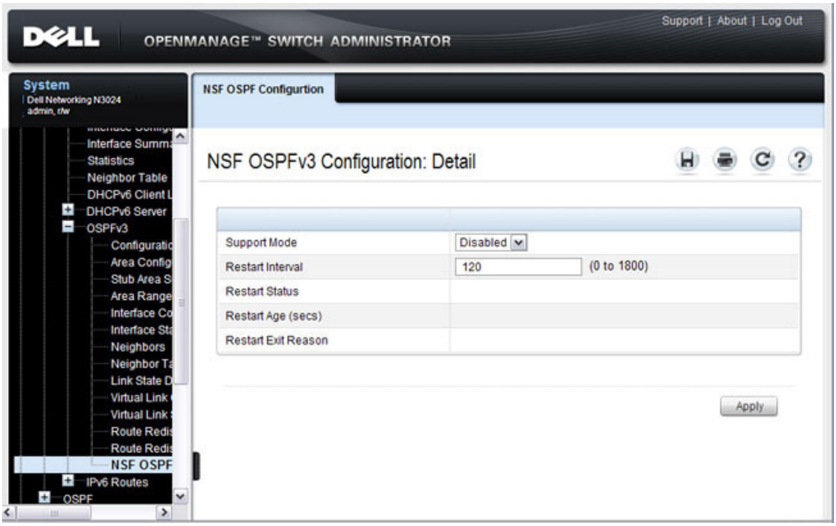


NSF OSPFv3 Configuration

Use the **NSF OSPFv3 Configuration** page to configure the non-stop forwarding (NSF) support mode and to view NSF summary information for the OSPFv3 feature. NSF is a feature used in switch stacks to maintain switching and routing functions in the event of a stack unit failure. For information about NSF, see "What is Nonstop Forwarding? " on page 201 in the Stacking chapter.

To display the page, click **Routing** → **OSPFv3** → **NSF OSPFv3 Configuration** in the navigation panel.

Figure 35-34. NSF OSPFv3 Configuration



Configuring OSPF Features (CLI)

This section provides information about the commands used for configuring and viewing OSPF settings on the switch. This section does not describe all available **show** commands. For more information about all available OSPF commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Global OSPF Settings

Beginning in Privileged EXEC mode, use the following commands to configure various global OSPF settings for the switch.

Command	Purpose
configure	Enter global configuration mode.
router ospf	Enter OSPF configuration mode.
router-id *ip-address*	Set the 4-digit dotted-decimal number that uniquely identifies the router.
auto-cost reference-bandwidth *ref_bw*	Set the reference bandwidth used in the formula to compute link cost for an interface: link cost = ref_bw÷interface bandwidth The *ref_bw* variable is the reference bandwidth in Mbps (Range: 1–4294967).
capability opaque	Allow OSPF to store and flood opaque LSAs. An opaque LSA is used for flooding user defined information within an OSPF router domain.
compatible rfc1583	(Optional) Enable compatibility with RFC 1583. If all OSPF routers in the routing domain are capable of operating according to RFC 2328, OSPF 1583 compatibility mode should be disabled.

Command	Purpose
default-information originate [**always**] [**metric** *metric-value*] [**metric-type** *type-value*]	Control the advertisement of default routes. • **always** — Normally, OSPF originates a default route only if a default route is redistributed into OSPF (and default-information originate is configured). When the always option is configured, OSPF originates a default route, even if no default route is redistributed. • *metric-value* — The metric (or preference) value of the default route. (Range: 1–16777214) • *type-value* — The value is either 1 or 2: External type-1 route or External type-2 route.
default-metric *metric-value*	Set a default for the metric of distributed routes (Range: 1–16777214).
distance ospf {**external** | **inter-area** | **intra-area** } *distance*	Set the preference values of OSPF route types in the router. The range for the *distance* variable is 1–255. Lower route preference values are preferred when determining the best route.
enable	Enable OSPF.
exit-overflow-interval *seconds*	Specify the exit overflow interval for OSPF as defined in RFC 1765. The interval is the number of seconds after entering overflow state that a router will wait before attempting to leave the overflow state. (Range: 0–2147483647)
external-lsdb-limit *limit*	Configure the external LSDB limit for OSPF as defined in RFC 1765. If the value is –1, then there is no limit. The *limit* variable is the maximum number of non-default AS external-LSAs allowed in the router's link-state database. (Range: 1 to 2147483647)
maximum-paths *integer*	Set the number of paths that OSPF can report for a given destination (Range: 1–4). **Note**: The upper limit of this command depends on the selected SDM template. Use **show sdm prefer** command to display the upper limit of ECMP next hops.

Command	Purpose
passive-interface default	Configure OSPF interfaces as passive by default. This command overrides any interface-level passive mode settings.OSPF does not form adjacencies on passive interfaces but does advertise attached networks as stub networks.
timers spf *delay-time hold-time*	Specify the SPF delay and hold time. • *delay-time* — SPF delay time. (Range: 0–65535 seconds) • *hold-time* — SPF hold time. (Range: 0–65535 seconds)
exit	Exit to Global Configuration mode.
exit	Exit to Privileged EXEC mode.
show ip ospf	View OSPF global configuration and status.
show ip ospf statistics	View OSPF routing table calculation statistics.
clear ip ospf [{**configuration** | **redistribution** | **counters** | **neighbor** [**interface vlan** *vlan-id* [*neighbor-id*]]}]	Reset specific OSPF states. If no parameters are specified, OSPF is disabled and then re-enabled.

Configuring OSPF Interface Settings

Beginning in Privileged EXEC mode, use the following commands to configure per-interface OSPF settings.

Command	Purpose
configure	Enter global configuration mode.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ip ospf area *area-id* [**secondaries none**]	Enables OSPFv2 on the interface and sets the area ID of an interface. This command supersedes the effects of network area command. The *area-id* variable is the ID of the area (Range: IP address or decimal from 0 –4294967295) Use the **secondaries none** keyword to prevent the interface from advertising its secondary addresses into the OSPFv2 domain.
ip ospf priority *number-value*	Set the OSPF priority for the interface. The *number-value* variable specifies the priority of an interface (Range: 0 to 255). The default priority is 1, which is the highest router priority. A value of 0 indicates that the router is not eligible to become the designated router on this network.
ip ospf retransmit-interval *seconds*	Set the OSPF retransmit interval for the interface. The *seconds* variable is the number of seconds between link-state advertisements for adjacencies belonging to this router interface. This value is also used when retransmitting database descriptions and link-state request packets. Valid values range from 0 to 3600 seconds (1 hour).
ip ospf hello-interval *seconds*	Set the OSPF hello interval for the interface. This parameter must be the same for all routers attached to a network. The *seconds* variable indicates the number of seconds to wait before sending Hello packets from the interface. (Range: 1–65535).

Command	Purpose
ip ospf dead-interval *seconds*	Set the OSPF dead interval for the interface. The *seconds* variable indicates the number of seconds a router waits to see a neighbor router's Hello packets before declaring that the router is down (Range: 1–65535). This parameter must be the same for all routers attached to a network. This value should be some multiple of the Hello Interval.
ip ospf transmit-delay *seconds*	Set the OSPF Transit Delay for the interface. The *seconds* variable sets the estimated number of seconds it takes to transmit a link state update packet over this interface. (Range: 1–3600 seconds)
ip ospf mtu-ignore	Disable OSPF MTU mismatch detection on the received database description.
ip ospf network {**broadcast** | **point-to-point**}	Set the OSPF network type on the interface to broadcast or point-to-point. OSPF selects a designated router and originates network LSAs only for broadcast networks. No more than two OSPF routers may be present on a point-to-point link.
ip ospf authentication {**none** | {**simple** *key*} | {**encrypt** *key key-id*}}	Set the OSPF Authentication Type and Key for the specified interface. • **encrypt** — MD5 encrypted authentication key. • *key* — Authentication key for the specified interface. (Range: 8 bytes or less if the authentication type is **simple** and 16 bytes or less if the type is **encrypt**.) • *key-id* — Authentication key identifier for the authentication type **encrypt**. (Range: 0–25)
ip ospf cost *interface-cost*	Set the metric cost of the interface. The *interface-cost* variable specifies the cost (link-state metric) of the OSPF interface. (Range: 1–65535)
bandwidth *bw*	Set the interface bandwidth used in the formula to compute link cost for an interface: link cost = ref_bw÷interface bandwidth The *bw* variable is the interface bandwidth (Range: 1–10000000 Kbps).

Command	Purpose
exit	Exit to Global Configuration Mode
router ospf	Enter OSPF configuration mode.
passive-interface vlan *vlan-id*	Make an interface passive to prevent OSPF from forming an adjacency on an interface. OSPF advertises networks attached to passive interfaces as stub networks.
network *ip-address wildcard-mask* **area** *area-id*	Enable OSPFv2 on interfaces whose primary IP address matches this command, and make the interface a member of the specified area. • *ip-address* — Base IPv4 address of the network area. • *wildcard-mask* — The network mask indicating the subnet. • *area-id* — The ID of the area (Range: IP address or decimal from 0–4294967295).
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ip ospf interface [**vlan** *vlan-id*]	View summary information for all OSPF interfaces configured on the switch or for the specified routing interface.
show ip ospf interface stats vlan *vlan-id*	View per-interface OSPF statistics.

Configuring Stub Areas and NSSAs

Beginning in Privileged EXEC mode, use the following commands to configure OSPF stub areas and NSSAs.

Command	Purpose
configure	Enter global configuration mode.
router ospf	Enter OSPF configuration mode.
area *area-id* **stub**	Create a stub area for the specified area ID.
area *area-id* **stub no-summary**	Prevent Summary LSAs from being advertised into the stub area.

Command	Purpose
area *area-id* **default-cost** *integer*	Configure the metric value (default cost) for the type 3 summary LSA sent into the stub area. Range: 1–16777215)
area *area-id* **nssa**	Create an NSSA for the specified area ID.
area *area-id* **nssa no-summary**	Configure the NSSA so that summary LSAs are not advertised into the NSSA.
area *area-id* **nssa translator-role** {**always** | **candidate**}	Configure the translator role of the NSSA. • **always** — The router assumes the role of the translator when it becomes a border router. • **candidate** — The router can participate in the translator election process when it attains border router status.
area *area-id* **nssa translator-stab-intv** *integer*	Configure the translator stability interval of the NSSA. The *integer* variable is the period of time that an elected translator continues to perform its duties after it determines that its translator status has been deposed by another router. (Range: 0–3600)
area *area-id* **nssa default-information-originate** [**metric** *metric-value*] [**metric-type** *metric-type-value*]	Configure the metric value and type for the default route advertised into the NSSA. The metric type can be comparable (nssa-external 1) or non-comparable (nssa-external 2).
area *area-id* **nssa no-redistribution**	Prevent learned external routes from being redistributed to the NSSA.
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ip ospf area *area-id*	View the configuration and status of an OSPF area.

Configuring Virtual Links

Beginning in Privileged EXEC mode, use the following commands to configure OSPF Virtual Links.

Command	Purpose
configure	Enter global configuration mode.
router ospf	Enter OSPF configuration mode.
area *area-id* **virtual-link** *neighbor-id*	Create the OSPF virtual interface for the specified area-id and neighbor router. The *neighbor-id* variable is the IP address of the neighboring router.
area *area-id* **virtual-link** *router-id* [**authentication** [**message-digest** | **null**]] [[**authentication-key** *key*] | [**message-digest-key** *key-id* **md5** *key*]]	Create the OSPF virtual interface for the specified area-id and neighbor router. Use the optional parameters to configure authentication for the virtual link. If the area has not been previously created, it is created by this command. If the area already exists, the virtual-link information is added or modified. • **authentication**—Specifies authentication type. • **message-digest**—Specifies that message-digest authentication is used. • **null**—No authentication is used. Overrides password or message-digest authentication if configured for the area. • **md5**—Use MD5 Encryption for an OSPF Virtual Link • *key*—Authentication key for the specified interface. (Range: 8 bytes or less if the authentication type is simple and 16 bytes or less if the type is encrypt.) • *key-id*—Authentication key identifier for the authentication type encrypt. (Range: 0-255)
area *area-id* **virtual-link** *neighbor-id* **retransmit-interval** *seconds*	Set the OSPF retransmit interval for the virtual link interface. The *seconds* variable is the number of seconds to wait between retransmitting LSAs if no acknowledgement is received. (Range: 0–3600)

Command	Purpose
area *area-id* **virtual-link** *neighbor-id* **hello-interval** *seconds*	Set the OSPF hello interval for the virtual link. The *seconds* variable indicates the number of seconds to wait before sending Hello packets from the virtual interface. (Range: 1–65535).
area *area-id* **virtual-link** *neighbor-id* **dead-interval** *seconds*	Set the OSPF dead interval for the virtual link. The *seconds* variable indicates the number of seconds to wait before the virtual interface is assumed to be dead. (Range: 1–65535)
area *area-id* **virtual-link** *neighbor-id* **transmit-delay** *seconds*	Set the OSPF Transit Delay for the interface. The *seconds* variable is the number of seconds to increment the age of the LSA before sending, based on the estimated time it takes to transmit from the interface. (Range: 0–3600)
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ip ospf virtual-link brief	View summary information about all virtual links configured on the switch.

Configuring OSPF Area Range Settings

Beginning in Privileged EXEC mode, use the following commands to configure an OSPF area range.

Command	Purpose
configure	Enter global configuration mode.
router ospf	Enter OSPF configuration mode.
area *area-id* **range** *ip-address mask* **{summarylink | nssaexternallink} [advertise |not-advertise]**	Configure a summary prefix for routes learned in a given area. • *area-id* — Identifies the OSPF NSSA to configure. (Range: IP address or decimal from 0–4294967295) • *ip-address* — IP address. • *subnet-mask* — Subnet mask associated with IP address. • **summarylink** — Specifies a summary link LSDB type. • **nssaexternallink** — Specifies an NSSA external link LSDB type. • **advertise** — Advertisement of the area range. • **not-advertise** — Suppresses advertisement of the area range.
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ip ospf range *area-id*	View information about the area ranges for the specified *area-id*.

Configuring OSPF Route Redistribution Settings

Beginning in Privileged EXEC mode, use the following commands to configure OSPF route redistribution settings.

Command	Purpose
configure	Enter global configuration mode.
router ospf	Enter OSPF configuration mode.

Command	Purpose
distribute-list *accesslistname* **out** {**rip** | **static** | **connected**}	Specify the access list to filter routes received from the source protocol. The ACL must already exist on the switch. For information about the commands used for configuring ACLs, see "Configuring ACLs (CLI) " on page 664. • *accesslistname* — The name used to identify an existing ACL. • **rip** — Apply the specified access list when RIP is the source protocol. • **static** — Apply the specified access list when packets come through the static route. • **connected** — Apply the specified access list when packets come from a directly connected route.
redistribute {**rip** | **static** | **connected**} [**metric** *integer*] [**metric-type** {**1** | **2**}] [**tag** *integer*] [**subnets**]	Configure OSPF to allow redistribution of routes from the specified source protocol/routers. • **rip** — Specifies RIP as the source protocol. • **static** — Specifies that the source is a static route. • **connected** — Specifies that the source is a directly connected route. • *metric* — Specifies the metric to use when redistributing the route. (Range: 0–16777214) • **metric-type 1** — Type 1 external route. • **metric-type 2** — Type 2 external route. • *tag* — Value attached to each external route. (Range: 0–4294967295) • **subnets**—Unless this keyword is configured, OSPF distributes only class A, class B, and class C prefixes.
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ip ospf	View OSPF configuration and status information, including route distribution information.

Configuring NSF Settings for OSPF

Beginning in Privileged EXEC mode, use the following commands to configure the non-stop forwarding settings for OSPF.

Command	Purpose
configure	Enter global configuration mode.
router ospf	Enter OSPF configuration mode.
nsf [**ietf**] **helper strict-lsa-checking**	Require that an OSPF helpful neighbor exit helper mode whenever a topology change occurs. Use the **ietf** keyword to distinguish the IETF standard implementation of graceful restart from other implementations.
nsf [**ietf**] **restart-interval** *seconds*	Configure the length of the grace period on the restarting router. The *seconds* keyword is the number of seconds that the restarting router asks its neighbors to wait before exiting helper mode. The restarting router includes the restart interval in its grace LSAs (range 1–1800 seconds)
nsf helper [**planned-only**]	Allow OSPF to act as a helpful neighbor for a restarting router. Include the **planned-only** keyword to indicate that OSPF should only help a restarting router performing a planned restart.
nsf [**ietf**] [**planned-only**]	Enable a graceful restart of OSPF. • **ietf** — This keyword is used to distinguish the IETF standard implementation of graceful restart from other implementations. Since the IETF implementation is the only one supported, this keyword is optional. • **planned-only** — This keyword indicates that OSPF should only perform a graceful restart when the restart is planned (i.e., when the restart is a result of the initiate failover command).

Configuring OSPFv3 Features (CLI)

This section provides information about the commands used for configuring OSPFv3 settings on the switch. For more information about the commands and about additional **show** commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Global OSPFv3 Settings

Beginning in Privileged EXEC mode, use the following commands to configure various global OSPFv3 settings for the switch.

Command	Purpose
configure	Enter global configuration mode.
ipv6 router ospf	Enter OSPFv3 configuration mode.
router-id *ip-address*	Set the 4-digit dotted-decimal number that uniquely identifies the router.
auto-cost reference-bandwidth *ref_bw*	Set the reference bandwidth used in the formula to compute link cost for an interface: link cost = ref_bw÷interface bandwidth The *ref_bw* variable is the reference bandwidth in Mbps (Range: 1–4294967).
default-information originate [**always**] [**metric** *metric-value*] [**metric-type** *type-value*]	Control the advertisement of default routes. • **always** — Normally, OSPFv3 originates a default route only if a default route is redistributed into OSPFv3 (and default-information originate is configured). When the always option is configured, OSPFv3 originates a default route, even if no default route is redistributed. • *metric-value* — The metric (or preference) value of the default route. (Range: 1–16777214) • *type-value* — The value is either 1 or 2: External type-1 route or External type-2 route.
default-metric *metric-value*	Set a default for the metric of distributed routes. (Range: 1–16777214).

Command	Purpose
distance ospf {**external** | **inter-area** | **intra-area** } *distance*	Set the preference values of OSPFv3 route types in the router. The range for the *distance* variable is 1–255. Lower route preference values are preferred when determining the best route.
enable	Enable OSPFv3.
exit-overflow-interval *seconds*	Specify the exit overflow interval for OSPFv3 as defined in RFC 1765. The interval is the number of seconds after entering overflow state that a router will wait before attempting to leave the overflow state. (Range: 0–2147483647)
external-lsdb-limit *limit*	Configure the external LSDB limit for OSPFv3 as defined in RFC 1765. If the value is -1, then there is no limit. The *limit* variable is the maximum number of non-default AS external-LSAs allowed in the router's link-state database. (Range: –1 to 2147483647)
maximum-paths *maxpaths*	Set the number of paths that OSPFv3 can report for a given destination. (Range: 1–4.) **Note**: The upper limit of this command depends on the selected SDM template. Use the **show sdm prefer** command to display the maximum ECMP next hops limit.
passive-interface default	Configure OSPFv3 interfaces as passive by default. This command overrides any interface-level passive mode settings. OSPFv3 does not form adjacencies on passive interfaces but does advertise attached networks as stub networks.
exit	Exit to Global Configuration mode.
exit	Exit to Privileged EXEC mode.
show ipv6 ospf	View OSPFv3 global configuration and status.
clear ipv6 ospf [{**configuration** | **redistribution** | **counters** | **neighbor** [**interface vlan** *vlan-id* [*neighbor-id*]]}]	Reset specific OSPFv3 states. If no parameters are specified, OSPFv3 is disabled and then re-enabled.

Configuring OSPFv3 Interface Settings

Beginning in Privileged EXEC mode, use the following commands to configure per-interface OSPFv3 settings.

Command	Purpose
configure	Enter global configuration mode.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ipv6 ospf areaid *area-id*	Enables OSPFv3 on the interface and sets the area ID of an interface. This command supersedes the effects of network area command. The *area-id* variable is the ID of the area (Range: IP address or decimal from 0 –4294967295)
ipv6 ospf priority *number-value*	Set the OSPFv3 priority for the interface. The *number-value* variable specifies the priority of an interface (Range: 0 to 255). The default priority is 1, which is the highest router priority. A value of 0 indicates that the router is not eligible to become the designated router on this network.
ipv6 ospf retransmit-interval *seconds*	Set the OSPFv3 retransmit interval for the interface. The *seconds* variable is the number of seconds between link-state advertisements for adjacencies belonging to this router interface. This value is also used when retransmitting database descriptions and link-state request packets. Valid values range from 0 to 3600 seconds (1 hour).
ipv6 ospf hello-interval *seconds*	Set the OSPFv3 hello interval for the interface. This parameter must be the same for all routers attached to a network. The *seconds* variable indicates the number of seconds to wait before sending Hello packets from the interface. (Range: 1–65535).

Command	Purpose
ipv6 ospf dead-interval *seconds*	Set the OSPFv3 dead interval for the interface. The *seconds* variable indicates the number of seconds a router waits to see a neighbor router's Hello packets before declaring that the router is down (Range: 1–65535). This parameter must be the same for all routers attached to a network. This value should be some multiple of the Hello Interval.
ipv6 ospf transmit-delay *seconds*	Set the OSPFv3 Transit Delay for the interface. The *seconds* variable sets the estimated number of seconds it takes to transmit a link state update packet over this interface. (Range: 1–3600 seconds)
ip ospf mtu-ignore	Disable OSPFv3 MTU mismatch detection on received database description packets.
ipv6 ospf network {**broadcast** | **point-to-point** }	Set the OSPFv3 network type on the interface to broadcast or point-to-point. OSPFv3 selects a designated router and originates network LSAs only for broadcast networks. No more than two OSPFv3 routers may be present on a point-to-point link.
ipv6 ospf cost *interface-cost*	Set the metric cost of the interface. The *interface-cost* variable specifies the cost (link-state metric) of the OSPFv3 interface. (Range: 1–65535)
bandwidth *bw*	Set the interface bandwidth used in the formula to compute link cost for an interface: link cost = ref_bw÷interface bandwidth The *bw* variable is the interface bandwidth (Range: 1–10000000 Kbps).
exit	Exit to Global Configuration Mode
ipv6 router ospf	Enter OSPFv3 configuration mode.
passive-interface {vlan *vlan-id* | tunnel *tunnel-id*}	Make an interface passive to prevent OSPFv3 from forming an adjacency on an interface. OSPFv3 advertises networks attached to passive interfaces as stub networks.
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.

Command	Purpose
show ipv6 ospf interface [*interface-type interface-number*]	View summary information for all OSPFv3 interfaces configured on the switch or for the specified routing interface.
show ipv6 ospf interface stats *interface-type interface-number*	View per-interface OSPFv3 statistics.

Configuring Stub Areas and NSSAs

Beginning in Privileged EXEC mode, use the following commands to configure OSPFv3 stub areas and NSSAs.

Command	Purpose
configure	Enter global configuration mode.
ipv6 router ospf	Enter OSPFv3 configuration mode.
area *area-id* **stub**	Create a stub area for the specified area ID.
area *area-id* **stub no-summary**	Prevent Summary LSAs from being advertised into the stub area.
area *area-id* **default-cost** *cost*	Configure the metric value (default cost) for the type 3 summary LSA sent into the stub area. Range: 1–16777215)

Command	Purpose

area *area-id* **nssa** [**no-redistribution**] [**default-information-originate** [**metric** *metric-value*] [**metric-type** *metric-type-value*]] [**no-summary**] [**translator-role** *role*] [**translator-stab-intv** *interval*]	Create and configure an NSSA for the specified area ID.
	• *metric-value*—Specifies the metric of the default route advertised to the NSSA. (Range: 1–16777214)
	• *metric-type-value*—The metric type can be one of the following :
	  • A metric type of nssa-external 1 (comparable)
	  • A metric type of nssa-external 2 (non-comparable)
	• **no-summary—**Summary LSAs are not advertised into the NSSA
	• **role**—The translator role where role is one of the following :
	  • **always**—The router assumes the role of the translator when it becomes a border router.
	  • **candidate**—The router to participate in the translator election process when it attains border router status.
	• **interval**—The period of time that an elected translator continues to perform its duties after it determines that its translator status has been deposed by another router. (Range: 0–3600)

area *area-id* **nssa no-redistribution**	Prevent learned external routes from being redistributed to the NSSA.
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ipv6 ospf area *area-id*	Show configuration and status of an OSPF area.

Configuring Virtual Links

Beginning in Privileged EXEC mode, use the following commands to configure OSPFv3 Virtual Links.

Command	Purpose
configure	Enter global configuration mode.
ipv6 router ospf	Enter OSPFv3 configuration mode.
area *area-id* **virtual-link** *neighbor-id*	Create the OSPFv3 virtual interface for the specified *area-id* and neighbor router. The *neighbor-id* variable is the IP address of the neighboring router.
area *area-id* **virtual-link** *neighbor-id* **retransmit-interval** *seconds*	Set the OSPFv3 retransmit interval for the virtual link interface. The *seconds* variable is the number of seconds to wait between retransmitting LSAs if no acknowledgement is received. (Range: 0–3600)
area *area-id* **virtual-link** *neighbor-id* **hello-interval** *seconds*	Set the OSPFv3 hello interval for the virtual link. The *seconds* variable indicates the number of seconds to wait before sending Hello packets from the virtual interface. (Range: 1–65535).
area *area-id* **virtual-link** *neighbor-id* **dead-interval** *seconds*	Set the OSPFv3 dead interval for the virtual link. The *seconds* variable indicates the number of seconds to wait before the virtual interface is assumed to be dead. (Range: 1–65535)
area *area-id* **virtual-link** *neighbor-id* **transmit-delay** *seconds*	Set the OSPFv3 Transit Delay for the interface. The *seconds* variable is the number of seconds to increment the age of the LSA before sending, based on the estimated time it takes to transmit from the interface. (Range: 0–3600)
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ipv6 ospf virtual-link brief	View summary information about all virtual links configured on the switch.

Configuring an OSPFv3 Area Range

Beginning in Privileged EXEC mode, use the following commands to configure an OSPFv3 area range.

Command	Purpose
configure	Enter global configuration mode.
ipv6 router ospf	Enter OSPFv3 configuration mode.
area *area-id* **range** *ipv6-prefix/prefix-length* **{summarylink | nssaexternallink} [advertise |not-advertise]**	Configure a summary prefix for routes learned in a given area. • *area-id* — Identifies the OSPFv3 NSSA to configure. (Range: IP address or decimal from 0–4294967295) • *ipv6-prefix/prefix-length* — IPv6 address and prefix length. • **summarylink** — Specifies a summary link LSDB type. • **nssaexternallink** — Specifies an NSSA external link LSDB type. • **advertise** — Advertisement of the area range. • **not-advertise** — Suppresses advertisement of the area range.
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ipv6 ospf range *area-id*	View information about the area ranges for the specified area-id.

Configuring OSPFv3 Route Redistribution Settings

Beginning in Privileged EXEC mode, use the following commands to configure OSPFv3 route redistribution settings.

Command	Purpose
configure	Enter global configuration mode.
ipv6 router ospf	Enter OSPFv3 configuration mode.
redistribute {**static** | **connected**} [**metric** *metric*] [**metric-type** {1 | 2}] [**tag** *tag*]	Configure OSPFv3 to allow redistribution of routes from the specified source protocol/routers. • **static** — Specifies that the source is a static route. • **connected** — Specifies that the source is a directly connected route. • *metric* — Specifies the metric to use when redistributing the route. (Range: 0–16777214) • **metric-type 1** — Type 1 external route. • **metric-type 2** — Type 2 external route. • *tag* — Value attached to each external route, which might be used to communicate information between ASBRs. (Range: 0–4294967295)
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ipv6 ospf	View OSPFv3 configuration and status information, including information about redistributed routes.

Configuring NSF Settings for OSPFv3

Beginning in Privileged EXEC mode, use the following commands to configure the non-stop forwarding settings for OSPFv3.

Command	Purpose
configure	Enter global configuration mode.
ipv6 router ospf	Enter OSPFv3 configuration mode.
nsf [ietf] helper strict-lsa-checking	Require that an OSPFv3 helpful neighbor exit helper mode whenever a topology change occurs. Use the **ietf** keyword to distinguish the IETF standard implementation of graceful restart from other implementations.
nsf [ietf] restart-interval *seconds*	Configure the length of the grace period on the restarting router. The *seconds* keyword is the number of seconds that the restarting router asks its neighbors to wait before exiting helper mode. The restarting router includes the restart interval in its grace LSAs (range 1–1800 seconds)
nsf helper [planned-only]	Allow OSPFv3 to act as a helpful neighbor for a restarting router. Include the **planned-only** keyword to indicate that OSPFv3 should only help a restarting router performing a planned restart.
nsf [ietf] [planned-only]	Enable a graceful restart of OSPFv3. • **ietf** — This keyword is used to distinguish the IETF standard implementation of graceful restart from other implementations. Since the IETF implementation is the only one supported, this keyword is optional. • **planned-only** — This keyword indicates that OSPF should only perform a graceful restart when the restart is planned (i.e., when the restart is a result of the initiate failover command).

OSPF Configuration Examples

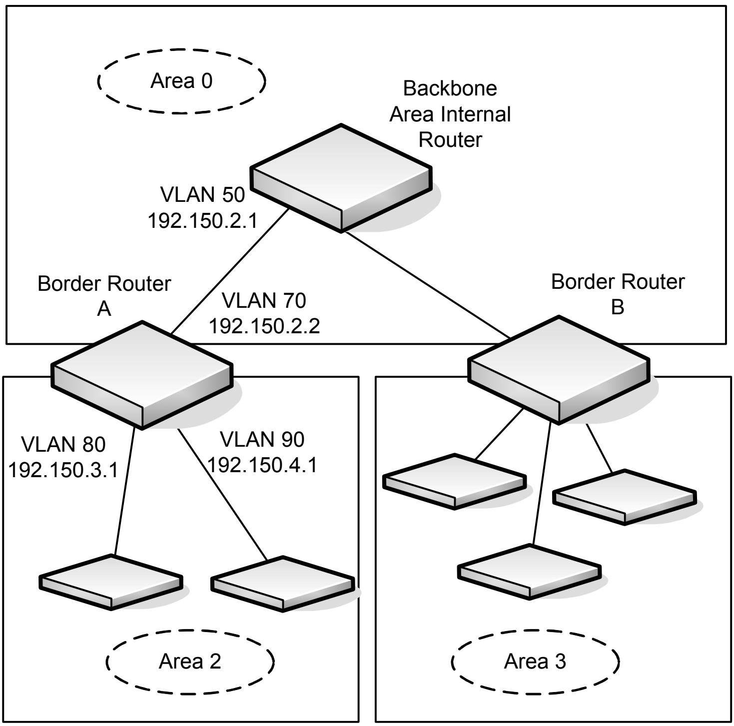
This section contains the following examples:

- Configuring an OSPF Border Router and Setting Interface Costs
- Configuring Stub and NSSA Areas for OSPF and OSPFv3
- Configuring a Virtual Link for OSPF and OSPFv3

Configuring an OSPF Border Router and Setting Interface Costs

This example shows how to configure the Dell Networking N-Series switch as an OSPF border router. The commands in this example configure the areas and interfaces on Border Router A shown in Figure 35-35.

Figure 35-35. OSPF Area Border Router



To Configure Border Router A:

1 Enable routing on the switch.

```
console#configure
console(config)#ip routing
```

2 Create VLANS 70, 80, and 90 and assign them to interfaces.

```
console(config)#vlan 70,80,90
console(config-vlan70,80,90)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 70
console(config-if-Gi1/0/1)#interface gi1/0/2
console(config-if-Gi1/0/2)#switchport access vlan 80
console(config-if-Gi1/0/1)#interface gi1/0/3
console(config-if-Gi1/0/2)#switchport access vlan 90
```

3 Assign IP addresses for VLANs 70, 80 and 90.

```
console(config)#interface vlan 70
console(config-if-vlan70)#ip address 192.150.2.2 255.255.255.0
console(config-if-vlan70)#exit

console(config)#interface vlan 80
console(config-if-vlan80)#ip address 192.150.3.1 255.255.255.0
console(config-if-vlan80)#exit

console(config)#interface vlan 90
console(config-if-vlan90)#ip address 192.150.4.1 255.255.255.0
console(config-if-vlan90)#exit
```

4 Enable OSPF on the switch and specify a router ID.

```
console(config)#router ospf
console(config-router)#router-id 192.150.9.9
console(config-router)#exit
```

5 Configure the OSPF area ID, priority, and cost for each interface.

NOTE: OSPF is globally enabled by default. To make it operational on the router, you configure OSPF for particular interfaces and identify which area the interface is associated with.

```
console(config)#interface vlan 70
console(config-if-vlan70)#ip ospf area 0.0.0.0
console(config-if-vlan70)#ip ospf priority 128
console(config-if-vlan70)#ip ospf cost 32
console(config-if-vlan70)#exit

console(config)#interface vlan 80
console(config-if-vlan80)#ip ospf area 0.0.0.2
console(config-if-vlan80)#ip ospf priority 255
console(config-if-vlan80)#ip ospf cost 64
console(config-if-vlan80)#exit

console(config)#interface vlan 90
console(config-if-vlan90)#ip ospf area 0.0.0.2
console(config-if-vlan90)#ip ospf priority 255
console(config-if-vlan90)#ip ospf cost 64
console(config-if-vlan90)#exit
```

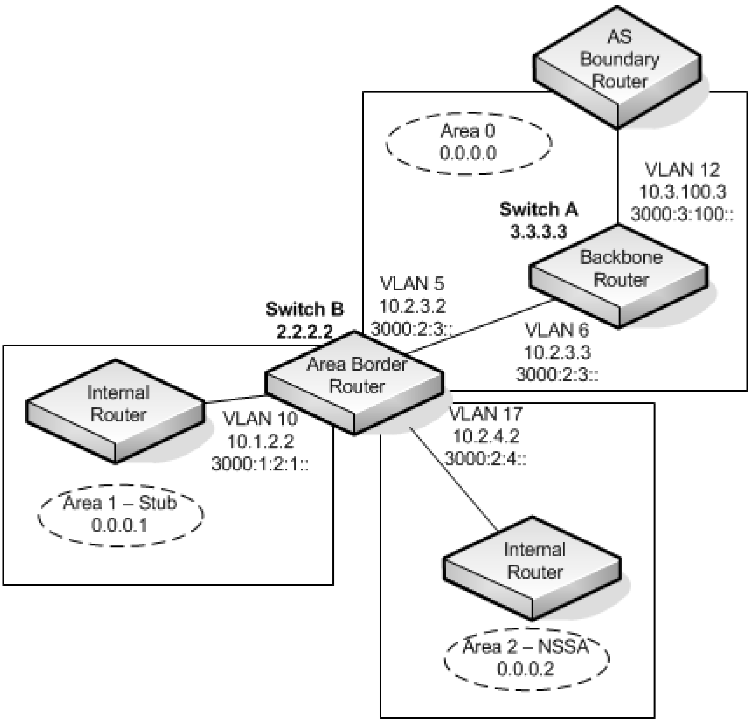
Configuring Stub and NSSA Areas for OSPF and OSPFv3

In this example, Area 0 connects directly to two other areas: Area 1 is defined as a stub area and Area 2 is defined as an NSSA area.

NOTE: OSPFv2 and OSPFv3 can operate concurrently on a network and on the same interfaces (although they do not interact). This example configures both protocols simultaneously.

Figure 35-36 illustrates this example OSPF configuration.

Figure 35-36. OSPF Configuration—Stub Area and NSSA Area



Switch A is a backbone router. It links to an ASBR (not defined here) that routes traffic outside the AS.

To configure Switch A:

1 Globally enable IPv6 and IPv4 routing:

```
console#configure
console(config)#ipv6 unicast-routing
console(config)#ip routing
```

2 Create VLANs 6 and 12 and assign them to interfaces.

```
console(config)#vlan 6,12
console(config-vlan6,12)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 6
console(config-if-Gi1/0/1)#interface gi1/0/2
console(config-if-Gi1/0/2)#switchport access vlan 12
```

3 Configure IP and IPv6 addresses on VLAN routing interface 6.

```
console(config-if)#interface vlan 6
console(config-if-vlan6)#ip address 10.2.3.3 255.255.255.0
console(config-if-vlan6)#ipv6 address 3000:2:3::/64 eui64
```

4 Associate the interface with area 0.0.0.0 and enable OSPFv3.

```
console(config-if-vlan6)#ip ospf area 0.0.0.0
console(config-if-vlan6)#ipv6 ospf
console(config-if-vlan6)#exit
```

5 Configure IP and IPv6 addresses on VLAN routing interface 12.

```
console(config)#interface vlan 12
console(config-if-vlan12)#ip address 10.3.100.3 255.255.255.0
console(config-if-vlan12)#ipv6 address 3000:3:100::/64 eui64
```

6 Associate the interface with area 0.0.0.0 and enable OSPFv3.

```
console(config-if-vlan12)#ip ospf area 0.0.0.0
console(config-if-vlan12)#ipv6 ospf
console(config-if-vlan12)#exit
```

7 Define the OSPF and OSPFv3 router IDs for the switch:

```
console(config)#ipv6 router ospf
console(config-rtr)#router-id 3.3.3.3
console(config-rtr)#exit

console(config)#router ospf
console(config-router)#router-id 3.3.3.3
console(config-router)#exit
```

Switch B is a ABR that connects Area 0 to Areas 1 and 2.

To configure Switch B:

1 Configure IPv6 and IPv4 routing. The static routes are included for illustration only: Redistributed static routes, like routes distributed from other protocols, are not injected into stub areas such as Area 1:

```
console#configure
console(config)#ipv6 unicast-routing

console(config)#ipv6 route 3000:44:44::/64
3000:2:3::210:18ff:fe82:c14
console(config)#ip route 10.23.67.0 255.255.255.0 10.2.3.3
```

2 Create VLANs 5, 10, and 17.

```
console(config)#vlan 5,10,17
console(config-vlan5,10,17)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 5
console(config-if-Gi1/0/1)#interface gi1/0/2
console(config-if-Gi1/0/2)#switchport access vlan 10
console(config-if-Gi1/0/1)#interface gi1/0/3
console(config-if-Gi1/0/2)#switchport access vlan 17
```

3 On VLANs 5, 10, and 17, configure IPv4 and IPv6 addresses and enable OSPFv3. For IPv6, associate VLAN 5 with Area 0, VLAN 10 with Area 1, and VLAN 17 with Area 2.

```
console(config)#interface vlan 5
console(config-if-vlan5)#ip address 10.2.3.2 255.255.255.0
console(config-if-vlan5)#ipv6 address 3000:2:3::/64 eui64
console(config-if-vlan5)#ipv6 ospf
console(config-if-vlan5)#ipv6 ospf areaid 0
console(config-if-vlan5)#exit
console(config)#interface vlan 10
console(config-if-vlan10)#ip address 10.1.2.2 255.255.255.0
console(config-if-vlan10)#ipv6 address 3000:1:2::/64 eui64
console(config-if-vlan10)#ipv6 ospf
console(config-if-vlan10)#ipv6 ospf areaid 1
console(config-if-vlan10)#exit
console(config)#interface vlan 17
console(config-if-vlan17)#ip address 10.2.4.2 255.255.255.0
console(config-if-vlan17)#ipv6 address 3000:2:4::/64 eui64
console(config-if-vlan17)#ipv6 ospf
console(config-if-vlan17)#ipv6 ospf areaid 2
console(config-if-vlan17)#exit
```

4 For IPv4: Configure the router ID, define an OSPF router, and define Area 1 as a stub., and define Area 2 as an NSSA.

```
console(config)#router ospf
console(config-router)#router-id 2.2.2.2
console(config-router)#area 0.0.0.1 stub
console(config-router)#area 0.0.0.2 nssa
```

5 For IPv4: Enable OSPF for IPv4 on VLANs 10, 5, and 17 by globally defining the range of IP addresses associated with each interface, and then associating those ranges with Areas 1, 0, and 2, respectively.

```
console(config-router)#network 10.1.2.0 0.0.0.255 area 0.0.0.1
console(config-router)#network 10.2.3.0 0.0.0.255 area 0.0.0.0
console(config-router)#network 10.2.4.0 0.0.0.255 area 0.0.0.2
```

6 For IPv4: Configure a metric cost to associate with static routes when they are redistributed via OSPF:

```
console(config-router)#redistribute static metric 1 subnets
console(config-router)#exit
```

7 For IPv6: Define an OSPF router. Define Area 1 as a stub and area 2 as a Not-So-Stubby-Area (NSSA). Configure a metric cost to associate with static routes when they are redistributed via OSPF:

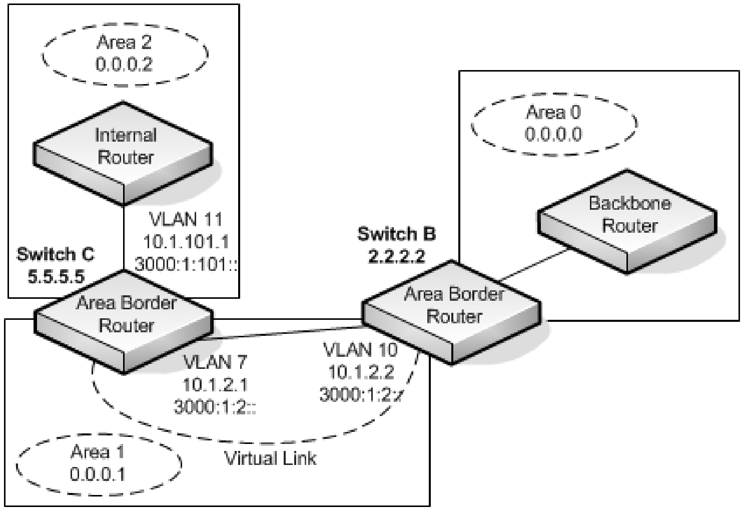
```
console(config)#ipv6 router ospf
console(config-rtr)#router-id 2.2.2.2
console(config-rtr)#area 0.0.0.1 stub
console(config-rtr)#area 0.0.0.2 nssa
console(config-rtr)#redistribute static metric 105 metric-type
1
console(config-rtr)#exit
```

Configuring a Virtual Link for OSPF and OSPFv3

In this example, Area 0 connects directly to Area 1. A virtual link is defined that traverses Area 1 and connects to Area 2. This example assumes other OSPF settings, such as area and interface configuration, have already been configured.

Figure 35-37 illustrates the relevant components in this example OSPF configuration.

Figure 35-37. OSPF Configuration—Virtual Link



Switch B is an ABR that directly connects Area 0 to Area 1. Note that in the previous example, Switch B connected to a stub area and an NSSA. Virtual links cannot be created across stub areas or NSSAs.

The following commands define a virtual link that traverses Area 1 to Switch C (5.5.5.5).

To configure Switch B:

1 Configure the virtual link to Switch C for IPv4.

```
console#configure
console(config)#router ospf
console(config-router)#area 0.0.0.1 virtual-link 5.5.5.5
console(config-router)#exit
```

2 Configure the virtual link to Switch C for IPv6.

```
console#configure
console(config)#ipv6 router ospf
console(config-rtr)#area 0.0.0.1 virtual-link 5.5.5.5
console(config-rtr)#exit
```

Switch C is a ABR that enables a virtual link from the remote Area 2 in the AS to Area 0. The following commands define a virtual link that traverses Area 1 to Switch B (2.2.2.2).

To configure Switch C:

1 For IPv4, assign the router ID, create the virtual link to Switch B, and associate the VLAN routing interfaces with the appropriate areas.

```
console(config)#router ospf
console(config-router)#area 0.0.0.1 virtual-link 2.2.2.2
console(config-router)#exit
```

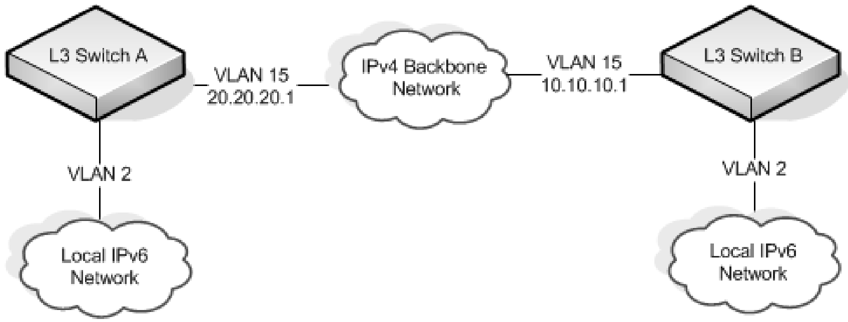
2 For IPv6, assign the router ID and create the virtual link to Switch B.

```
console(config)#ipv6 router ospf
console(config-rtr)#area 0.0.0.1 virtual-link 2.2.2.2
console(config-rtr)#exit
```

Interconnecting an IPv4 Backbone and Local IPv6 Network

In Figure 35-38, two Dell Networking L3 switches are connected as shown in the diagram. The VLAN 15 routing interface on both switches connects to an IPv4 backbone network where OSPF is used as the dynamic routing protocol to exchange IPv4 routes. OSPF allows device 1 and device 2 to learn routes to each other (from the 20.20.20.x network to the 10.10.10.x network and vice versa). The VLAN 2 routing interface on both devices connects to the local IPv6 network. OSPFv3 is used to exchange IPv6 routes between the two devices. The tunnel interface allows data to be transported between the two remote IPv6 networks over the IPv4 network.

Figure 35-38. IPv4 and IPv6 Interconnection Example



To configure Switch A:

1 Create the VLANs.

```
console(config)#vlan 2,15
console(config-vlan70,80,90)#interface te1/0/1
console(config-if-Te1/0/1)#switchport mode trunk
console(config-if-Te1/0/1)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 2
```

2 Enable IPv4 and IPv6 routing on the switch.

```
console(config)#ip routing
console(config)#ipv6 unicast-routing
```

3 Set the OSPF router ID.

```
console(config)#router ospf
console(config-router)#router-id 1.1.1.1
console(config-router)#exit
```

4 Set the OSPFv3 router ID.

```
console(config)#ipv6 router ospf
console(config-rtr)#router-id 1.1.1.1
console(config-rtr)#exit
```

5 Configure the IPv4 address and OSPF area for VLAN 15.

```
console(config)#interface vlan 15
console(config-if-vlan15)#ip address 20.20.20.1 255.255.255.0
console(config-if-vlan15)#ip ospf area 0.0.0.0
console(config-if-vlan15)#exit
```

6 Configure the IPv6 address and OSPFv3 information for VLAN 2.

```
console(config)#interface vlan 2
console(config-if-vlan2)#ipv6 address 2020:1::1/64
console(config-if-vlan2)#ipv6 ospf
console(config-if-vlan2)#ipv6 ospf network point-to-point
console(config-if-vlan2)#exit
```

7 Configure the tunnel.

```
console(config)#interface tunnel 0
console(config-if-tunnel0)#ipv6 address 2001::1/64
console(config-if-tunnel0)#tunnel mode ipv6ip
console(config-if-tunnel0)#tunnel source 20.20.20.1
console(config-if-tunnel0)#tunnel destination 10.10.10.1
console(config-if-tunnel0)#ipv6 ospf
console(config-if-tunnel0)#ipv6 ospf network point-to-point
console(config-if-tunnel0)#exit
```

8 Configure the loopback interface. The switch uses the loopback IP address as the OSPF and OSPFv3 router ID.

```
console(config)#interface loopback 0
console(config-if-loopback0)#ip address 1.1.1.1 255.255.255.0
console(config-if-loopback0)#exit
console(config)#exit
```

To configure Switch B:

1 Create the VLANs.

```
console(config)#vlan 2,15
console(config-vlan70,80,90)#interface te1/0/1
console(config-if-Te1/0/1)#switchport mode trunk
console(config-if-Te1/0/1)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 2
```

2 Enable IPv4 and IPv6 routing on the switch.

```
console(config)#ip routing
console(config)#ipv6 unicast-routing
```

3 Set the OSPF router ID.

```
console(config)#router ospf
console(config-router)#router-id 2.2.2.2
console(config-router)#exit
```

4 Set the OSPFv3 router ID.

```
console(config)#ipv6 router ospf
console(config-rtr)#router-id 2.2.2.2
console(config-rtr)#exit
```

5 Configure the IPv4 address and OSPF area for VLAN 15.

```
console(config)#interface vlan 15
console(config-if-vlan15)#ip address 10.10.10.1 255.255.255.0
console(config-if-vlan15)#ip ospf area 0.0.0.0
console(config-if-vlan15)#exit
```

6 Configure the IPv6 address and OSPFv3 information for VLAN 2.

```
console(config)#interface vlan 2
console(config-if-vlan2)#ipv6 address 2020:2::2/64
console(config-if-vlan2)#ipv6 ospf
console(config-if-vlan2)#ipv6 ospf network point-to-point
console(config-if-vlan2)#exit
```

7 Configure the tunnel.

```
console(config)#interface tunnel 0
console(config-if-tunnel0)#ipv6 address 2001::2/64
console(config-if-tunnel0)#tunnel mode ipv6ip
console(config-if-tunnel0)#tunnel source 10.10.10.1
console(config-if-tunnel0)#tunnel destination 20.20.20.1
console(config-if-tunnel0)#ipv6 ospf
console(config-if-tunnel0)#ipv6 ospf network point-to-point
console(config-if-tunnel0)#exit
```

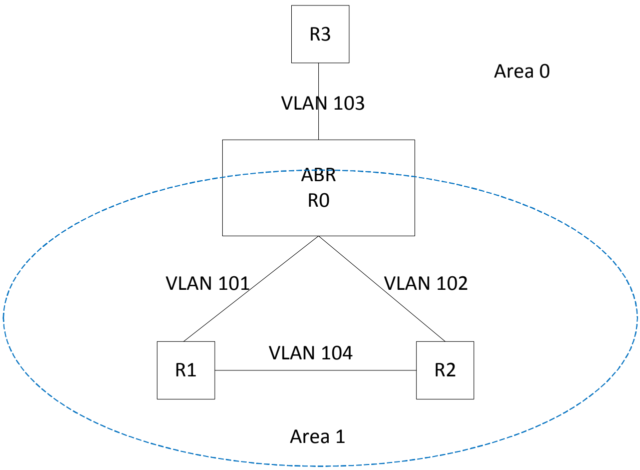
8 Configure the loopback interface. The switch uses the loopback IP address as the OSPF and OSPFv3 router ID.

```
console(config)#interface loopback 0
console(config-if-loopback0)#ip address 2.2.2.2 255.255.255.0
console(config-if-loopback0)#exit
console(config)#exit
```

Configuring the Static Area Range Cost

Figure 35-39 shows a topology for the configuration that follows.

Figure 35-39. Static Area Range Cost Example Topology



1 Configure R0.

```
terminal length 0
config
hostname ABR-R0
line console
exec-timeout 0
exit
vlan 101-103
exit
ip routing
router ospf
router-id 10.10.10.10
```

```
network 172.20.0.0 0.0.255.255 area 0
network 172.21.0.0 0.0.255.255 area 1
area 1 range 172.21.0.0 255.255.0.0 summarylink
timers spf 3 5
exit
interface vlan 101
ip address 172.21.1.10 255.255.255.0
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/21
switchport mode trunk
description “R1”
exit
interface vlan 102
ip address 172.21.2.10 255.255.255.0
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/22
description “R2”
switchport mode trunk
exit
interface vlan 103
ip address 172.20.1.10 255.255.255.0
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/23
switchport mode trunk
description “R3”
exit
exit
```

2 Configure R1.

```
terminal length 0
config
hostname R1
line console
exec-timeout 0
exit
vlan 101,104
exit
```

```
ip routing
router ospf
router-id 1.1.1.1
network 172.21.0.0 0.0.255.255 area 1
timers spf 3 5
exit
interface vlan 101
ip address 172.21.1.1 255.255.255.0
routing
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/21
switchport mode trunk
exit
interface vlan 104
ip address 172.21.3.1 255.255.255.0
routing
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/22
switchport mode trunk
exit
interface loopback 0
ip address 172.21.254.1 255.255.255.255
exit
exit
```

3 Configure R2.

```
terminal length 0
config
line console
serial timeout 0
exit
ip routing
router ospf
router-id 2.2.2.2
network 172.21.0.0 0.0.255.255 area 1
timers spf 3 5
exit
vlan 102,104
exit
interface vlan 102
```

```
ip address 172.21.2.2 255.255.255.0
routing
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/21
switchport mode trunk
exit
interface vlan 104
ip address 172.21.3.2 255.255.255.0
routing
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/22
switchport mode trunk
exit
interface loopback 0
ip address 172.21.254.2 255.255.255.255
exit
exit
```

4 R3 config:

```
terminal length 0
config
line console
serial timeout 0
exit
ip routing
router ospf
router-id 3.3.3.3
network 172.21.0.0 0.0.255.255 area 0
timers spf 3 5
exit
vlan 103
exit
interface vlan 103
ip address 172.21.1.1 255.255.255.0
routing
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/21
```

```
switchport mode trunk
exit
interface loopback 0
ip address 172.21.254.2 255.255.255.255
exit
exit
```

Discussion

With no area range cost specified, the range uses auto cost:

```
(ABR-R0) #show ip ospf range 1

         Prefix      Subnet Mask   Type     Action       Cost   Active
     172.21.0.0      255.255.0.0      S  Advertise       Auto        Y

(ABR-R0) #show ip ospf database summary

                Network Summary States (Area 0.0.0.0)
LS Age: 644
LS options: (E-Bit)
LS Type: Network Summary LSA
LS Id: 172.21.0.0 (network prefix)
Advertising Router: 10.10.10.10
LS Seq Number: 0x80000002
Checksum: 0x8ee1
Length: 28
 Network Mask: 255.255.0.0
 Metric: 2
```

Min—The cost can be set to 0, the minimum value. OSPF re-advertises the summary LSA with a metric of 0:

```
(ABR-R0) (config-router)#area 1 range 172.21.0.0 255.255.0.0 summarylink
advertise cost ?

<0-16777215>             Set area range cost

(ABR-R0) (config-router)#area 1 range 172.21.0.0 255.255.0.0
summarylink advertise cost 0

(ABR-R0) #show ip ospf range 1

         Prefix      Subnet Mask   Type     Action       Cost   Active
     172.21.0.0      255.255.0.0      S  Advertise          0        Y

(ABR-R0) #show ip ospf 0 database summary
```

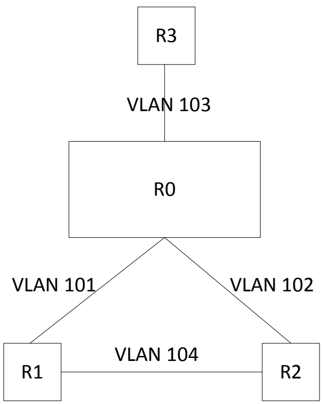
```
                Network Summary States (Area 0.0.0.0)
LS Age: 49
LS options: (E-Bit)
LS Type: Network Summary LSA
LS Id: 172.21.0.0 (network prefix)
Advertising Router: 10.10.10.10
LS Seq Number: 0x80000003
Checksum: 0x78f8
Length: 28
Network Mask: 255.255.0.0
Metric: 0
```

The cost can be set to the maximum value, 16,777,215, which is LSInfinity. Since OSPF cannot send a type 3 summary LSA with this metric (according to RFC 2328), the summary LSA is flushed. The individual routes are not re-advertised.

Configuring Flood Blocking

Figure 35-40 shows an example topology for flood blocking. The configuration follows.

Figure 35-40. Flood Blocking Topology



1 Configure R0:

```
terminal length 0
config
```

```
hostname R0
line console
exec-timeout 0
exit
vlan 101-103
exit
ip routing
router ospf
router-id 10.10.10.10
network 172.20.0.0 0.0.255.255 area 0
network 172.21.0.0 0.0.255.255 area 0
timers spf 3 5
exit
interface vlan 101
ip address 172.21.1.10 255.255.255.0
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/21
switchport mode trunk
description “R1”
exit
interface vlan 102
ip address 172.21.2.10 255.255.255.0
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/22
description “R2”
switchport mode trunk
exit
interface vlan 103
ip address 172.20.1.10 255.255.255.0
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/23
switchport mode trunk
description “R3”
exit
exit
```

2 Configure R1:

```
terminal length 0
config
hostname R1
line console
exec-timeout 0
exit
vlan 101,104
exit
ip routing
router ospf
router-id 1.1.1.1
network 172.21.0.0 0.0.255.255 area 0
timers spf 3 5
exit
interface vlan 101
ip address 172.21.1.1 255.255.255.0
routing
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/21
switchport mode trunk
exit
interface vlan 104
ip address 172.21.3.1 255.255.255.0
routing
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/22
switchport mode trunk
exit
interface loopback 0
ip address 172.21.254.1 255.255.255.255
exit
exit
```

3 Configure R2:

```
terminal length 0
config
line console
serial timeout 0
exit
```

```
ip routing
router ospf
router-id 2.2.2.2
network 172.21.0.0 0.0.255.255 area 0
timers spf 3 5
exit
vlan 102,104
exit
interface vlan 102
ip address 172.21.2.2 255.255.255.0
routing
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/21
switchport mode trunk
exit
interface vlan 104
ip address 172.21.3.2 255.255.255.0
routing
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/22
switchport mode trunk
exit
interface loopback 0
ip address 172.21.254.2 255.255.255.255
exit
exit
```

4 Configure R3:

```
terminal length 0
config
line console
serial timeout 0
exit
ip routing
router ospf
router-id 3.3.3.3
network 172.21.0.0 0.0.255.255 area 0
timers spf 3 5
exit
```

```
vlan 103
exit
interface vlan 103
ip address 172.21.1.1 255.255.255.0
routing
ip ospf hello-interval 1
ip ospf dead-interval 4
ip ospf network point-to-point
exit
interface te1/0/21
switchport mode trunk
exit
interface loopback 0
ip address 172.21.254.2 255.255.255.255
exit
exit
```

Discussion

With flood blocking disabled on all interfaces, sending a T3 summary LSA from R3 to R0 will cause R0 to forward the LSA on its interface to R1. Enabling flood blocking on R0's interface to R1 will inhibit this behavior.

```
(R0)(config-if-vlan101)ip ospf database-filter all out
```

A trace on the R3-R0 link shows that the LSA is actually flooded from R1 to R0, since R1 received the LSA via R2. Even though R1 does not receive this LSA directly from R0, it still correctly computes the route through the R0:

```
(R1) #show ip route
console#show ip route

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2

O IA   100.0.0.0/24 [110/2] via 172.21.1.10,   00h:01m:35s,  0/25
```

OSPF also blocks external LSAs on the blocked interface. Stopping and restarting R3's OSPF protocol causes R3 to re-originate its router LSA. R0 does not send R3's router LSA on the blocked interface.

With flood blocking enabled on the R0 interface, if the link from R0 to R1 bounces, R0 Database Description packets do not include any LSAs. However, database synchronization still occurs (through R2) and R1 computes the correct routes after the link is restored.

Configuring OSPF VRFs

Dell Networking VRF is an implementation of Virtual Routing and Forwarding (VRF) for OSPF for IPv4 networks. Virtual Routing and Forwarding allows multiple independent instances for the forwarding plane to exist simultaneously. Refer to "VRF " on page 1273 for more information.

VRF configuration follows the same steps as configuration for the default routing instance with two additional steps: creating the VRF instance and associating VLANs to the instance. Existing commands which have been enabled for VRF accept an additional VRF instance identifier (name). VRF names can be up to 32 characters in length. If a VRF instance identifier is not used in the command, it applies to the global routing instance by default.

Follow the steps below to create a VRF and enable OSPF routing in the VRF:

First, create the VLAN instances associated to the VRF. It is recommended that a VLAN numbering scheme be developed to allow for future growth and to assist in the easy recognition of which VLANs are associated to which VRFs.

In global config mode, create the pool of VLANs:

```
console#configure terminal
console(config)#vlan 100-109
console(config-vlan100-109)#exit
```

Assign the VLAN to an interface:

```
console(config)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 100
console(config-if-Gi1/0/1)#exit
```

Create the VRF and enable routing:

```
console(config)#ip vrf red
console(config-vrf-red)#ip routing
console(config-vrf-red)#exit
```

Assign IP addresses to the interfaces:

```
console(config)#interface vlan 100
```

```
console(config-if-vlan100)#ip address 192.168.0.1 /24
```

Put the VLAN interface into the VRF:

```
console(config-if-vlan100)#ip vrf forwarding red
console(config-if-vlan100)#exit

Routing interface moved from Default router instance to red router
instance.
```

Enable OSPF on the VRF, assign a network and enable OSPF for the VRF:

```
console(config)#router ospf vrf red
console(Config-router-vrf-red)#network 192.168.0.0 255.255.255.0
area 0
console(Config-router-vrf-red)#router-id 192.168.0.253
console(Config-router-vrf-red)#redistribute connected
console(Config-router-vrf-red)#enable
console(Config-router-vrf-red)#exit


console(config)#show ip ospf vrf red

Router ID...................................... 192.168.0.253
OSPF Admin Mode................................ Enable
RFC 1583 Compatibility......................... Enable
External LSDB Limit............................ No Limit
Exit Overflow Interval......................... 0
Spf Delay Time................................. 5 seconds
Spf Hold Time.................................. 10 seconds
Flood Pacing Interval.......................... 33 ms
LSA Refresh Group Pacing Time.................. 60 seconds
Opaque capability.............................. Enable
AutoCost Ref BW................................ 100 Mbps
Default Passive Setting........................ Disabled
Maximum Paths.................................. 16
Default Metric................................. Not configured
Stub Router Configuration...................... None
Summary LSA Metric Override.................... Disabled

Default Route Advertise........................ Disabled
Always......................................... False
Metric......................................... Not configured
Metric Type.................................... External Type 2
```

```
Number of Active Areas......................... 0 (0 normal, 0
stub, 0 nssa)
ABR Status..................................... Disable
ASBR Status.................................... Disable
Stub Router Status............................. Inactive
External LSDB Overflow......................... False
External LSA Count............................. 0
External LSA Checksum.......................... 0
AS_OPAQUE LSA Count............................ 0
AS_OPAQUE LSA Checksum......................... 0
New LSAs Originated............................ 0
LSAs Received.................................. 0
LSA Count...................................... 0
Maximum Number of LSAs......................... 66408
LSA High Water Mark............................ 0
AS Scope LSA Flood List Length................. 0
Retransmit List Entries........................ 0
Maximum Number of Retransmit Entries........... 265632
Retransmit Entries High Water Mark............. 0

NSF Support.................................... Disabled
NSF Restart Interval........................... 120 seconds
NSF Restart Status............................. Not Restarting
NSF Restart Age................................ 0 seconds
NSF Restart Exit Reason........................ Not attempted
NSF Helper Support............................. Always
NSF Helper Strict LSA Checking................. Enabled
```

36

VRF

Dell Networking N3000 and N4000 Series Switches

NOTE: This feature is not available on Dell Networking N1500/N2000 Series switches.

Virtual Routing and Forwarding (VRF) allows multiple independent instances of the forwarding plane to exist simultaneously. (The terms VRF, VRF instance, and virtual forwarding instance all refer to the same thing.) VRF allows the administrator to segment the network without incurring the costs of multiple routers. Each VRF instance operates as an independent VPN. The IP addresses assigned to each VPN may overlap. Static route leaking to and from the global instance is supported. Configuration of static route leaking among non-default VRFs results in undefined behavior.

VRF-associated VLANs may not overlap with other VRF instances.

VRF is supported on Dell Networking N3000 and N4000 Series switches. In addition to the default global instances, the following number of VRFs are supported.

Table 36-1. VRF Instances Supported in Dell Networking N-Series Switches

N3000 Series	N4000 Series
12	52

The following capabilities are supported for VRFs:

- Static routing (including route leaking)
- OSPF
- ARP
- Ping
- VRRP
- Trace route
- DHCP relay (IP helper)

- ICMP echo reply configuration
- ICMP error interval configuration

VRF Resource Sharing

Hardware resources such as routes and ARP entries are shared between VRFs. If a VRF allocates the maximum routes supported by the system, no VRF will be able to add a new route.

VRF ARP Entries

There is no support to reserve ARP entries per VRF instance as the system purges the least recently used ARP entry automatically. The maximum number of static ARP entries is enforced on a per VR instance basis.

VRF Route Entries

Routes are shared among the VR instances. The number of routes supported can never exceed the platform supported number. Initially, the number of “free” routes is the platform supported maximum. “Free” routes are available for any VR to use.

Two schemes are imposed on sharing of routes between the VR instances: Reservation and Restriction. The administrator can use the **maximum routes** command to reserve a number of routes for a VRF or to restrict the maximum number of routes available to a VR instance.

Reserved routes are deducted from the “free” routes available in the system. In-use routes are also deducted from the “free” routes available in the system.

The dynamic number of routes available to be allocated to a VRF instance is the lower of the number of “free” routes available in the system and the administrator-configured maximum routes.

The system-wide limit on static route entries is enforced on a per-VR-instance basis. That is, each VRF may allocate the system limit of static routes.

VRF configuration follows the same steps as configuration for the default routing instance with two additional steps: creating the VRF instance and associating VLANs to the instance. Existing commands which have been enabled for VRF accept an additional VRF instance identifier (name). VRF names can be up to 32 characters in length. If a VRF instance identifier is not used in the command, it applies to the global routing instance by default.

Follow the steps below to create a VRF and enable OSPF routing in the VRF:

First, create the VLAN instances associated to the VRF. It is recommended that a VLAN numbering scheme be developed to allow for future growth and to assist in the easy recognition of which VLANs are associated to which VRFs.

1. In global config mode, create the pool of VLANs:

```
console#config
console(config)#vlan 100-109
console(config-vlan100-109)#exit
```

2. Assign the VLAN to an interface:

```
console(config)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 100
console(config-if-Gi1/0/1)#exit
```

3. Create the VRF and enable routing:

```
console(config)#ip vrf red
console(config-vrf-red)#ip routing
console(config-vrf-red)#exit
```

4. Assign IP addresses to the interfaces:

```
console(config)#interface vlan 100
console(config-if-vlan100)#ip address 192.168.0.1 /24
```

5. Put the VLAN interface into the VRF:

```
console(config-if-vlan100)#ip vrf forwarding red

Routing interface moved from Default router instance to red
router instance.

console(config-if-vlan100)#exit
```

6. Enable OSPF on the VRF, assign a network and enable OSPF for the VRF:

```
console(config)#router ospf vrf red
console(Config-router-vrf-red)#network 192.168.0.0
255.255.255.0 area 0
console(Config-router-vrf-red)#router-id 192.168.0.253
console(Config-router-vrf-red)#redistribute connected
console(Config-router-vrf-red)#enable
console(Config-router-vrf-red)#exit
```

Use the **show ip ospf vrf** command to view the configuration of the VRF:

```
console(config)#show ip ospf vrf red

Router ID...................................... 192.168.0.253
```

```
OSPF Admin Mode................................. Enable
RFC 1583 Compatibility.......................... Enable
External LSDB Limit............................. No Limit
Exit Overflow Interval.......................... 0
Spf Delay Time.................................. 5 seconds
Spf Hold Time................................... 10 seconds
Flood Pacing Interval........................... 33 ms
LSA Refresh Group Pacing Time................... 60 seconds
Opaque capability............................... Enable
AutoCost Ref BW................................. 100 Mbps
Default Passive Setting......................... Disabled
Maximum Paths................................... 16
Default Metric.................................. Not configured
Stub Router Configuration....................... None
Summary LSA Metric Override..................... Disabled

Default Route Advertise......................... Disabled
Always.......................................... False
Metric.......................................... Not configured
Metric Type..................................... External Type 2


Number of Active Areas.......................... 0 (0 normal, 0
stub, 0 nssa)
ABR Status...................................... Disable
ASBR Status..................................... Disable
Stub Router Status.............................. Inactive
External LSDB Overflow.......................... False
External LSA Count.............................. 0
External LSA Checksum........................... 0
AS_OPAQUE LSA Count............................. 0
AS_OPAQUE LSA Checksum.......................... 0
New LSAs Originated............................. 0
LSAs Received................................... 0
LSA Count....................................... 0
Maximum Number of LSAs.......................... 66408
LSA High Water Mark............................. 0
AS Scope LSA Flood List Length.................. 0
Retransmit List Entries......................... 0
Maximum Number of Retransmit Entries............ 265632
Retransmit Entries High Water Mark.............. 0

NSF Support..................................... Disabled
NSF Restart Interval............................ 120 seconds
NSF Restart Status.............................. Not Restarting
NSF Restart Age................................. 0 seconds
```

```
NSF Restart Exit Reason........................ Not attempted
NSF Helper Support............................. Always
NSF Helper Strict LSA Checking................. Enabled
```

37

RIP

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure Routing Information Protocol (RIP) on the switch. RIP is a dynamic routing protocol for IPv4 networks.

The topics covered in this chapter include:

- RIP Overview
- Default RIP Values
- Configuring RIP Features (Web)
- Configuring RIP Features (CLI)
- RIP Configuration Example

RIP Overview

RIP is an Interior Gateway Protocol (IGP) that performs dynamic routing within a network. Dell Networking N-Series switches support two dynamic routing protocols: OSPF and Routing Information Protocol (RIP).

Unlike OSPF, RIP is a distance-vector protocol and uses UDP broadcasts to maintain topology information and hop counts to determine the best route to transmit IP traffic. RIP is best suited for small, homogenous networks.

How Does RIP Determine Route Information?

The routing information is propagated in RIP update packets that are sent out both periodically and in the event of a network topology change. On receipt of a RIP update, depending on whether the specified route exists or does not exist in the route table, the router may modify, delete or add the route to its route table.

RIP uses hop count, which is the number of routers an IP packet must pass through, to calculate the best route for a packet. A route with a low hop count is preferred over a route with a higher hop count. A directly-connected route has a hop-count of 0. With RIP, the maximum number of hops from source to destination is 15. Packets with a hop count greater than 15 are dropped because the destination network is considered unreachable.

What Is Split Horizon?

RIP uses a technique called split horizon to avoid problems caused by including routes in updates sent to the router from which the route was originally learned. With simple split horizon, a route is not included in updates sent on the interface on which it was learned. In split horizon with poison reverse, a route is included in updates sent on the interface where it was learned, but the metric is set to infinity.

What RIP Versions Are Supported?

There are two versions of RIP:

- RIP-1 defined in RFC 1058
  - Routes are specified by IP destination network and hop count
  - The routing table is broadcast to all stations on the attached network
- RIP-2 defined in RFC 1723
  - Route specification is extended to include subnet mask and gateway
  - The routing table is sent to a multicast address, reducing network traffic
  - An authentication method is used for security

The Dell Networking N-Series switches support both versions of RIP. You may configure a given port:

- To receive packets in either or both formats
- To transmit packets formatted for RIP-1 or RIP-2 or to send RIP-2 packets to the RIP-1 broadcast address
- To prevent any RIP packets from being received
- To prevent any RIP packets from being transmitted

Default RIP Values

RIP is globally enabled by default. To make it operational on the router, you configure and enable RIP for particular VLAN routing interfaces.

Table 37-1 shows the global default values for RIP.

Table 37-1. RIP Global Defaults

Parameter	Default Value
Admin Mode	Enabled
Split Horizon Mode	Simple
Auto Summary Mode	Disabled
Host Routes Accept Mode	Enabled
Default Information Originate	Disabled
Default Metric	None configured
Route Redistribution	Disabled for all sources.

Table 37-2 shows the per-interface default values for RIP.

Table 37-2. RIP Per-Interface Defaults

Parameter	Default Value
Admin Mode	Disabled
Send Version	RIPv2
Receive Version	Both
Authentication Type	None

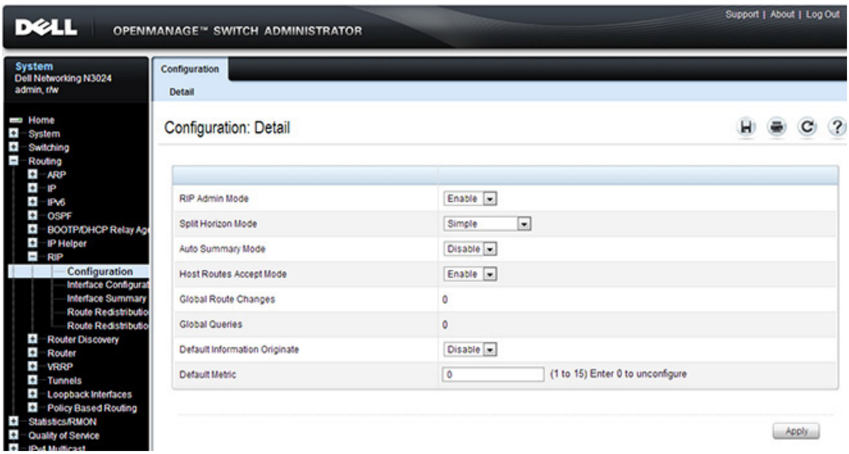
Configuring RIP Features (Web)

This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring RIP features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click at the top of the page.

RIP Configuration

Use the **Configuration** page to enable and configure or disable RIP in Global mode. To display the page, click **Routing** → **RIP** → **Configuration** in the navigation panel.

Figure 37-1. RIP Configuration

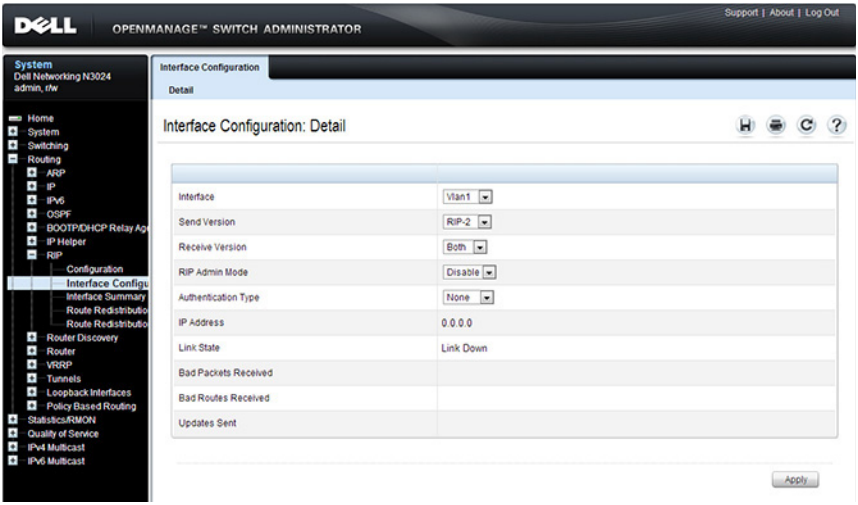


RIP Interface Configuration

Use the **Interface Configuration** page to enable and configure or to disable RIP on a specific interface.

To display the page, click **Routing** → **RIP** → **Interface Configuration** in the navigation panel.

Figure 37-2. RIP Interface Configuration

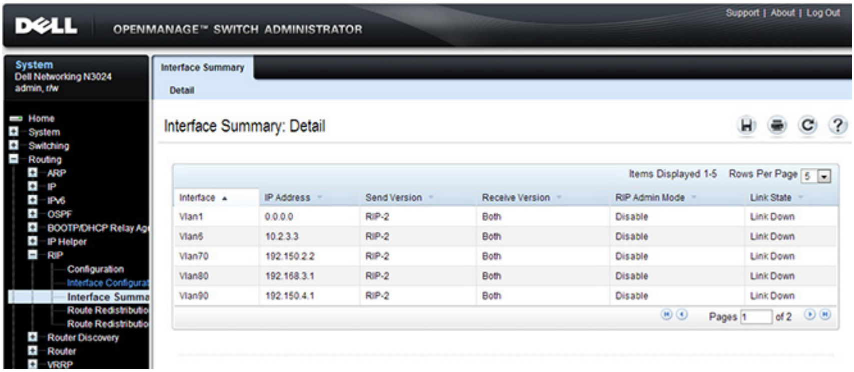


RIP Interface Summary

Use the **Interface Summary** page to display RIP configuration status on an interface.

To display the page, click **Routing → RIP → Interface Summary** in the navigation panel.

Figure 37-3. RIP Interface Summary

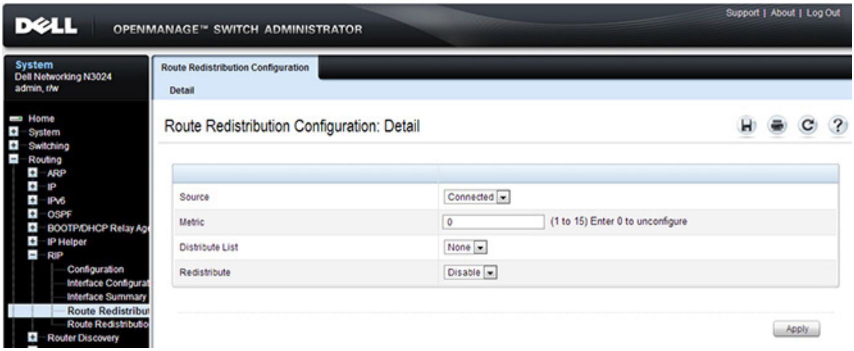


RIP Route Redistribution Configuration

Use the **Route Redistribution Configuration** page to configure the RIP Route Redistribution parameters. The allowable values for each fields are displayed next to the field. If any invalid values are entered, an alert message is displayed with the list of all the valid values.

To display the page, click **Routing** → **RIP** → **Route Redistribution Configuration** in the navigation panel.

Figure 37-4. RIP Route Redistribution Configuration



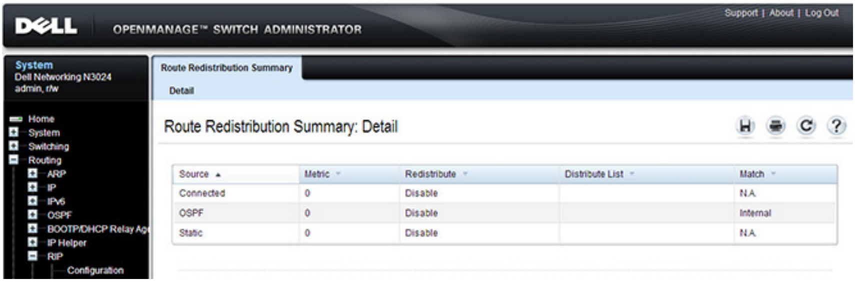
NOTE: Static reject routes are not redistributed by RIP. For a static reject route, the next hop interface value is Null0. Packets to the network address specified in static reject routes are intentionally dropped.

RIP Route Redistribution Summary

Use the **Route Redistribution Summary** page to display Route Redistribution configurations.

To display the page, click **Routing** → **RIP** → **Route Redistribution Summary** in the navigation panel.

Figure 37-5. RIP Route Redistribution Summary



Configuring RIP Features (CLI)

This section provides information about the commands used for configuring RIP settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Global RIP Settings

Beginning in Privileged EXEC mode, use the following commands to configure various global RIP settings for the switch.

NOTE: RIP is enabled by default. The Global RIP Settings are optional.

Command	Purpose
configure	Enter global configuration mode.
router rip	Enter RIP configuration mode.
split-horizon {none | simple | poison}	Set the RIP split horizon mode. • **none** — RIP does not use split horizon to avoid routing loops. • **simple** — RIP uses split horizon to avoid routing loops. • **poison** — RIP uses split horizon with poison reverse (increases routing packet update size).
auto-summary	Enable the RIP auto-summarization mode.
no hostroutesaccept	Prevent the switch from accepting host routes.
default-information originate	Control the advertisement of default routes.
default-metric *metric-value*	Set a default for the metric of distributed routes. The *metric-value* variable is the metric (or preference) value of the default route. (Range: 1–15)
enable	Reset the default administrative mode of RIP in the router (active)
CTRL + Z	Exit to Privileged EXEC mode.
show ip rip	View various RIP settings for the switch.

Configuring RIP Interface Settings

Beginning in Privileged EXEC mode, use the following commands to configure per-interface RIP settings.

Command	Purpose
configure	Enter global configuration mode.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ip rip	Enable RIP on the interface.
ip rip send version {rip1 rip1c | rip2 |none}	Configure the interface to allow RIP control packets of the specified version(s) to be sent.
ip rip receive version {rip1 | rip2 | both | none}	Configure the interface to allow RIP control packets of the specified version(s) to be received.
ip rip authentication {none | {simple *key***} | {encrypt** *key key-id***}**	set the RIP Version 2 Authentication Type and Key for the interface. • *key* — Authentication key for the specified interface. (Range: 16 bytes or less) • **encrypt** — Specifies the Ethernet unit/port of the interface to view information. • *key-id* — Authentication key identifier for authentication type encrypt. (Range: 0-255)
exit	Exit to Global Configuration Mode
exit	Exit to Privileged Exec mode.
show ip rip interface vlan *vlan-id*	View RIP configuration information for the specified routing interface.
show ip rip interface brief	View summary information about the RIP configuration on all interfaces.

Configuring Route Redistribution Settings

Beginning in Privileged EXEC mode, use the following commands to configure an OSPF area range and to configure route redistribution settings.

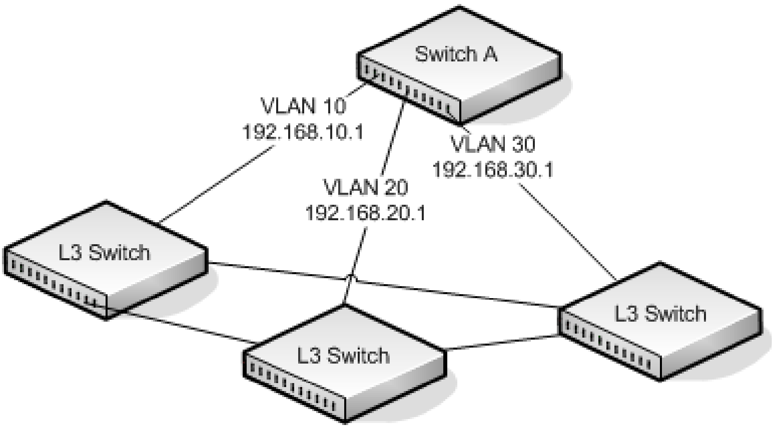
Command	Purpose
configure	Enter global configuration mode.
router rip	Enter RIP configuration mode.
distribute-list *accesslistname* **out** {**ospf** | **static** | **connected**}	Specify the access list to filter routes received from the source protocol. The ACL must already exist on the switch. For information about the commands used for configuring ACLs, see "Configuring ACLs (CLI) " on page 664. • *accesslistname* — The name used to identify an existing ACL. • **ospf** — Apply the specified access list when OSPF is the source protocol. • **static** — Apply the specified access list when packets come through the static route. • **connected** — Apply the specified access list when packets come from a directly connected route.
redistribute {**static** | **connected**} [**metric** *integer*]	Configure RIP to allow redistribution of routes from the specified source protocol/routers. • **static** — Specifies that the source is a static route. • **connected** — Specifies that the source is a directly connected route. • *metric* — Specifies the metric to use when redistributing the route. Range: 1-15.

Command	Purpose
redistribute ospf [metric *metric*] **[match [internal] [external 1] [external 2] [nssa-external 1] [nssa-external 2]]**	Configure RIP to allow redistribution of routes from the OSPF. • **ospf**— Specifies OSPF as the source protocol. • *metric* — Specifies the metric to use when redistributing the route. Range: 1-15. • **internal** — Adds internal matches to any match types presently being redistributed. • **external 1** — Adds routes imported into OSPF as Type-1 external routes into any match types presently being redistributed. • **external 2** — Adds routes imported into OSPF as Type-2 external routes into any match types presently being redistributed. • **nssa-external 1** — Adds routes imported into OSPF as NSSA Type-1 external routes into any match types presently being redistributed. • **nssa-external 2** — Adds routes imported into OSPF as NSSA Type-2 external routes into any match types presently being redistributed.
distance rip *integer*	Set the route preference value of RIP in the router. Lower route preference values are preferred when determining the best route.
exit	Exit to Global Config mode.
exit	Exit to Privileged Exec mode.
show ip rip	View information about the RIP route distribution configuration.

RIP Configuration Example

This example includes four Dell Networking N-Series switches that use RIP to determine network topology and route information. The commands in this example configure Switch A shown in Figure 37-6.

Figure 37-6. RIP Network Diagram



To configure the switch:

1 Enable routing on the switch

```
console#config
console(config)#ip routing
```

2 Create VLANs 10, 20, and 30.

```
console(config)#vlan 10,20,30
console(config-vlan10,20,30)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 10
console(config-if-Gi1/0/1)#interface gi1/0/2
console(config-if-Gi1/0/2)#switchport access vlan 20
console(config-if-Gi1/0/2)#interface gi1/0/3
console(config-if-Gi1/0/3)#switchport access vlan 30
```

3 Assign an IP address and enable RIP on each interface. Additionally, the commands specify that each interface can receive both RIP-1 and RIP-2 frames but send only RIP-2 formatted frames.

```
console(config)#interface vlan 10
```

```
console(config-if-vlan10)#ip address 192.168.10.1 255.255.255.0
console(config-if-vlan10)#ip rip
console(config-if-vlan10)#ip rip receive version both
console(config-if-vlan10)#ip rip send version rip2
console(config-if-vlan10)#exit

console(config)#interface vlan 20
console(config-if-vlan20)#ip address 192.168.20.1 255.255.255.0
console(config-if-vlan20)#ip rip
console(config-if-vlan20)#ip rip receive version both
console(config-if-vlan20)#ip rip send version rip2
console(config-if-vlan20)#exit

console(config)#interface vlan 30
console(config-if-vlan30)#ip address 192.168.30.1 255.255.255.0
console(config-if-vlan30)#ip rip
console(config-if-vlan30)#ip rip receive version both
console(config-if-vlan30)#ip rip send version rip2
console(config-if-vlan30)#exit
```

4 Enable auto summarization of subprefixes when crossing classful boundaries.

```
console(config)#router rip
console(config-router)#auto-summary
console(config-router)#exit
console(config)#exit
```

5 Verify the configuration

```
console#show ip rip

RIP Admin Mode.......................... Enable
Split Horizon Mode...................... Simple
Auto Summary Mode....................... Enable
Host Routes Accept Mode................. Enable
Global route changes.................... 0
Global queries.......................... 0

Default Metric.................... Not configured
Default Route Advertise........... 0

console#show ip rip interface brief

Interface  IP Address     Send      Receive  RIP       Link
                          Version   Version  Mode      State
---------- ------------   --------- -------  --------  ------
Vl1        0.0.0.0        RIP-2     RIP-2    Disable   Down
```

```
Vl10        192.168.10.1   RIP-2      Both      Enable     Down
Vl20        192.168.10.1   RIP-2      Both      Enable     Down
Vl30        192.168.10.1   RIP-2      Both      Disable    Down
```

38

VRRP

Dell Networking N1500, N3000, and N4000 Series Switches

NOTE: This feature is not available on Dell Networking N2000 Series switches.

This chapter describes how to configure Virtual Routing Redundancy Protocol (VRRP) on the switch. VRRP can help create redundancy on networks in which end-stations are statically configured with the default gateway IP address.

The topics covered in this chapter include:

- VRRP Overview
- Default VRRP Values
- Configuring VRRP Features (Web)
- Configuring VRRP Features (CLI)
- VRRP Configuration Example

VRRP Overview

The Virtual Router Redundancy (VRRP) protocol is designed to handle default router (L3 switch) failures by providing a scheme to dynamically elect a backup router. VRRP can help minimize black hole periods due to the failure of the default gateway router during which all traffic directed towards it is lost until the failure is detected.

How Does VRRP Work?

VRRP eliminates the single point of failure associated with static default routes by enabling a backup router to take over from a master router without affecting the end stations using the route. The end stations will use a virtual IP address that will be recognized by the backup router if the master router fails. Participating routers use an election protocol to determine which router is the master router at any given time. A maximum of 50 virtual routers may be configured. A given port may appear as more than one virtual router to the network, also, more than one port on a switch may be configured as a virtual router.

With VRRP, a virtual router is associated with one or more IP addresses that serve as default gateways. In the event that the VRRP router controlling these IP addresses (formally known as the master) fails, the group of IP addresses and the default forwarding role is taken over by a Backup VRRP router.

What Is the VRRP Router Priority?

The VRRP router priority is a value from 1–255 that determines which router is the master. The greater the number, the higher the priority. If the virtual IP address is the IP address of a VLAN routing interface on one of the routers in the VRRP group, the router with IP address that is the same as the virtual IP address is the interface owner and automatically has a priority of 255. By default, this router is the VRRP master in the group.

If no router in the group owns the VRRP virtual IP address, the router with the highest configured priority is the VRRP master. If multiple routers have the same priority, the router with the highest IP address becomes the VRRP master.

If the VRRP master fails, other members of the VRRP group will elect a master based on the configured router priority values. For example, router A is the interface owner and master, and it has a priority of 255. Router B is configured with a priority of 200, and Router C is configured with a priority of 190. If Router A fails, Router B assumes the role of VRRP master because it has a higher priority.

What Is VRRP Preemption?

If preempt mode is enabled and a router with a higher priority joins the VRRP group, it takes over the VRRP master role if the current VRRP master is not the owner of the virtual IP address. The preemption delay controls how long to wait to determine whether a higher priority Backup router preempts a lower priority master. In certain cases, for example, during periods of network congestion, a backup router might fail to receive advertisements from the master. This could cause members in the VRRP group to change their states frequently, i.e. flap. The problem can be resolved by setting the VRRP preemption delay timer to a non-zero value.

What Is VRRP Accept Mode?

The accept mode allows the switch to respond to pings (ICMP Echo Requests) sent to the VRRP virtual IP address. The VRRP specification (RFC 3768) indicates that a router may accept IP packets sent to the virtual router IP address only if the router is the address owner. In practice, this restriction makes it more difficult to troubleshoot network connectivity problems. When a host cannot communicate, it is common to ping the host's default gateway to determine whether the problem is in the first hop of the path to the destination. When the default gateway is a virtual router that does not respond to pings, this troubleshooting technique is unavailable. In the Dell Networking N-Series switch VRRP feature, Accept Mode can be enabled to allow the system to respond to pings that are sent to the virtual IP address.

This capability adds support for responding to pings, but does not allow the VRRP master to accept other types of packets. The VRRP master responds to both fragmented and un-fragmented ICMP Echo Request packets. The VRRP master responds to Echo Requests sent to the virtual router's primary address or any of its secondary addresses.

Members of the virtual router who are in backup state discard ping packets destined to VRRP addresses, just as they discard any Ethernet frame sent to a VRRP MAC address.

When the VRRP master responds with an Echo Reply, the source IPv4 address is the VRRP address and source MAC address is the virtual router's MAC address.

What Are VRRP Route and Interface Tracking?

The VRRP Route/Interface Tracking feature extends VRRP capability to allow tracking of specific routes and interface IP states within the router that can alter the priority level of a virtual router for a VRRP group.

VRRP interface tracking monitors a specific interface IP state within the router. Depending on the state of the tracked interface, the feature can alter the VRRP priority level of a virtual router for a VRRP group.

NOTE: An exception to the priority level change is that if the VRRP group is the IP address owner, its priority is fixed at 255 and cannot be reduced through the tracking process.

With standard VRRP, the backup router takes over only if the router goes down. With VRRP interface tracking, if a tracked interface goes down on the VRRP master, the priority decrement value is subtracted from the router priority. If the master router priority becomes less than the priority on the backup router, the backup router takes over. If the tracked interface becomes up, the value of the priority decrement is added to the current router priority. If the resulting priority is more than the backup router priority, the original VRRP master resumes control.

VRRP route tracking monitors the reachability of an IP route. A tracked route is considered up when a routing table entry exists for the route and the route is accessible. When the tracked route is removed from the routing table, the priority of the VRRP router will be reduced by the priority decrement value. When the tracked route is added to the routing table, the priority will be incremented by the same.

Default VRRP Values

Table 38-1 shows the global default values for VRRP.

Table 38-1. VRRP Defaults

Parameter	Default Value
Admin Mode	Disabled
Virtual Router ID (VRID)	None (Range 1-255)
Preempt Mode	Enabled
Preempt Delay	0 Seconds
Learn Advertisement Timer Interval	Enabled
Accept Mode	Disabled
Priority	100
Advertisement Interval	1
Authentication	None
Route Tracking	No routes tracked
Interface Tracking	No interfaces tracked

Configuring VRRP Features (Web)

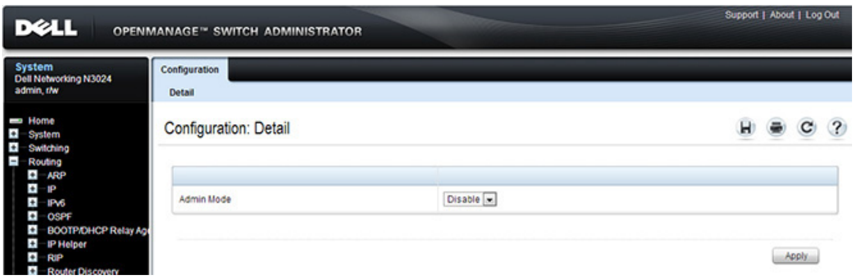
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring VRRP features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

VRRP Configuration

Use the **Configuration** page to enable or disable the administrative status of a virtual router.

To display the page, click **Routing** → **VRRP** → **Configuration** in the navigation panel.

Figure 38-1. VRRP Configuration

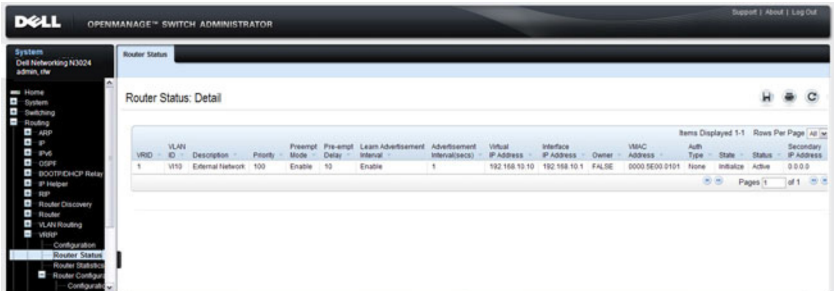


VRRP Virtual Router Status

Use the **Router Status** page to display virtual router status.

To display the page, click **Routing** → **VRRP** → **Router Status** in the navigation panel.

Figure 38-2. Virtual Router Status

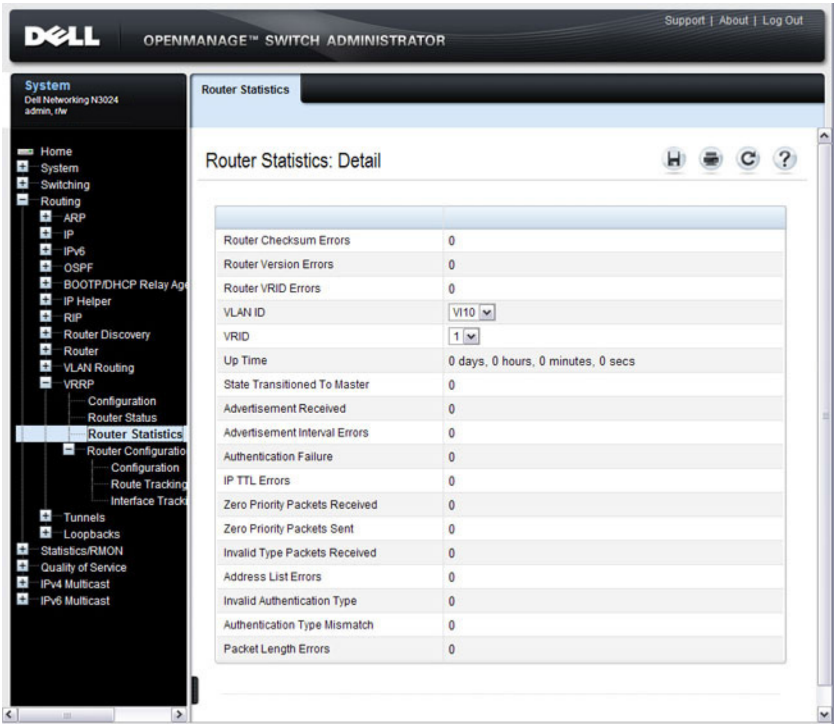


VRRP Virtual Router Statistics

Use the **Router Statistics** page to display statistics for a specified virtual router.

To display the page, click **Routing** → **VRRP** → **Router Statistics** in the navigation panel.

Figure 38-3. Virtual Router Statistics

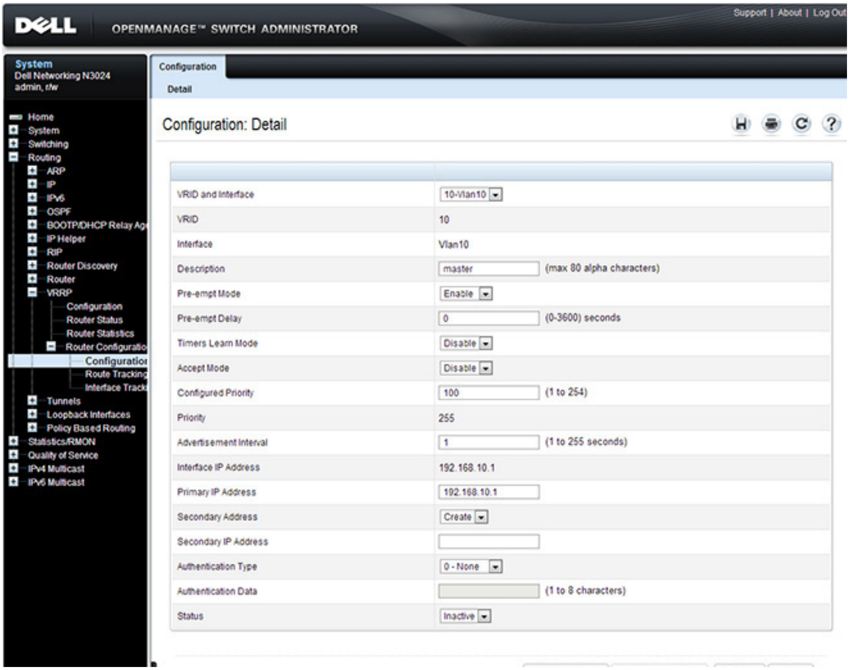


VRRP Router Configuration

Use the **Configuration** page to configure a virtual router.

To display the page, click **Routing** → **VRRP** → **Router Configuration** → **Configuration** in the navigation panel.

Figure 38-4. VRRP Router Configuration

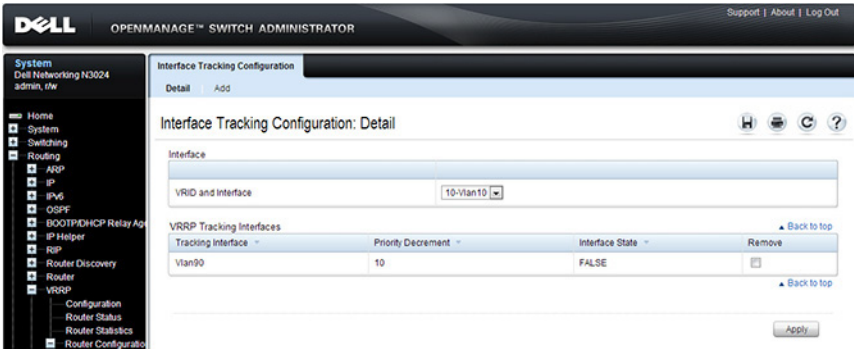


VRRP Route Tracking Configuration

Use the **Route Tracking Configuration** page to view routes that are tracked by VRRP and to add new tracked routes.

To display the page, click **Routing** → **VRRP** → **Router Configuration** → **Route Tracking Configuration** in the navigation panel.

Figure 38-5. VRRP Route Tracking Configuration



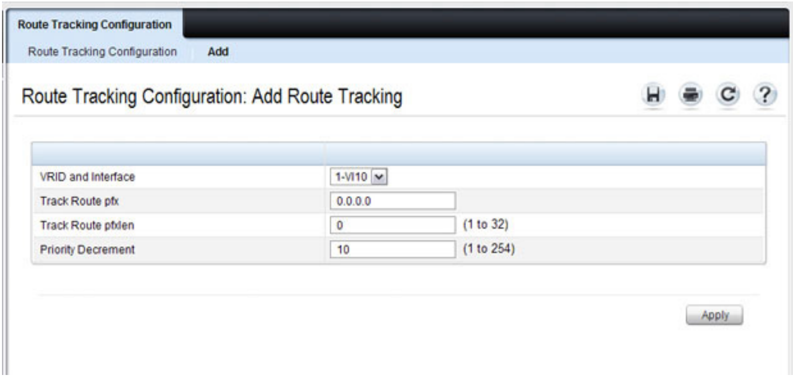
Configuring VRRP Route Tracking

To configure VRRP route tracking:

1 From the **Route Tracking Configuration** page, click **Add**.

The **Add Route Tracking** page displays.

Figure 38-6. Add Route Tracking



2 Select the virtual router ID and VLAN routing interface that will track the route.

3 Specify the destination network address (track route prefix) for the route to track. Use dotted decimal format, for example 192.168.10.0.

4 Specify the prefix length for the tracked route.

5 Specify a value for the **Priority Decrement** to define the amount that the router priority will be decreased when a tracked route becomes unreachable.

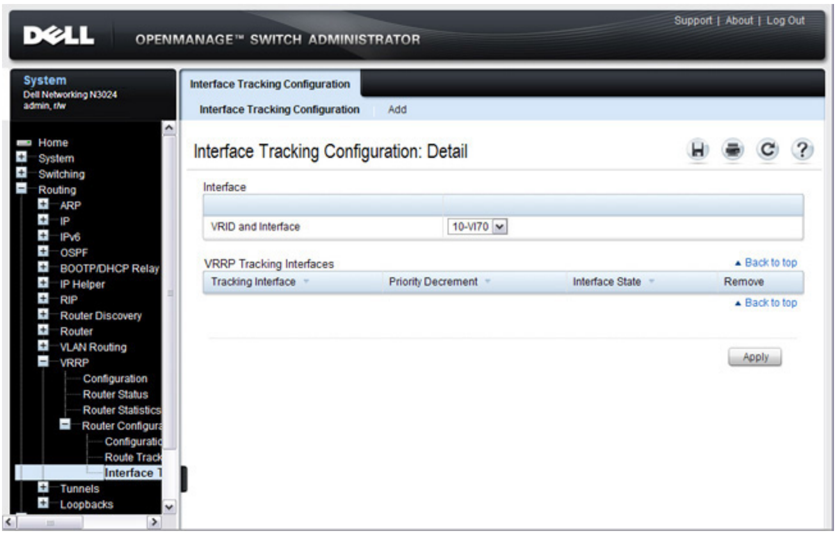
6. Click **Apply** to update the switch.

VRRP Interface Tracking Configuration

Use the **Interface Tracking Configuration** page to view interfaces that are tracked by VRRP and to add new tracked interfaces.

To display the page, click **Routing** → **VRRP** → **Router Configuration** → **Interface Tracking Configuration** in the navigation panel.

Figure 38-7. VRRP Interface Tracking Configuration



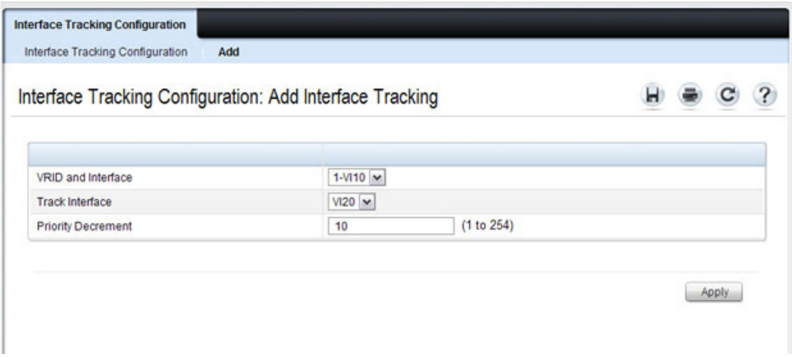
Configuring VRRP Interface Tracking

To configure VRRP interface tracking:

1 From the **Interface Tracking Configuration** page, click **Add**.

The **Add Interface Tracking** page displays.

Figure 38-8. VRRP Interface Tracking Configuration



2 Select the virtual router ID and VLAN routing interface that will track the interface.

3 Specify the interface to track.

4 Specify a value for the **Priority Decrement** to define the amount that the router priority will be decreased when a tracked interface goes down.

5. Click **Apply** to update the switch.

Configuring VRRP Features (CLI)

This section provides information about the commands used for configuring VRRP settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring VRRP Settings

Beginning in Privileged EXEC mode, use the following commands to configure switch and interface VRRP settings. This set of commands also describes how to configure VRRP interface and route tracking.

Command	Purpose
configure	Enter global configuration mode.
ip vrrp	Enable the administrative mode of VRRP for the router (L3 switch).
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
vrrp *vr-id*	Allow the interface to create in the VRRP group specified by the *vr-id* parameter, which is a number from 1–255.
vrrp *vr-id* **description**	(Optional) Create a text description that identifies the VRRP group.
vrrp *vr-id* **preempt** [**delay** *seconds*]	Enable the preemption mode value for the virtual router configured on a specified interface. A preempt delay can optionally be configured. A preempt delay is the number of seconds the VRRP router waits before the VRRP router sends an advertisement to claim master ownership.
vrrp *vr-id* **accept-mode**	Allow the VRRP master to accept ping packets sent to one of the virtual router's IP addresses.
vrrp *vr-id* **priority** *priority*	Set the priority value for the virtual router configured on the interface.
vrrp *vr-id* **ip** *ip-address* [**secondary**]	Set the virtual router IP address value for an interface.

Command	Purpose
vrrp *vr-id* **timers** {**learn** | **advertise** *seconds*}	Configure the VRRP timer settings. Use the keyword **learn** to enable VRRP to learn the advertisement timer interval of the master router. Use the keyword **advertise** to set the frequency, in seconds, that an interface on the specified virtual router sends a virtual router advertisement.
vrrp *vr-id* **authentication** {**none** | **simple** *key*}	Set the authorization details value for the virtual router configured on a specified interface. • *vr-id* — The virtual router identifier. (Range: 1-255) • **none** — Indicates authentication type is none. • **simple** — Authentication type is a simple text password. • *key* — The key for simple authentication. (Range: String values)
vrrp *vr-id* **mode**	Enable the virtual router configured on an interface, which starts the virtual router.
vrrp *vr-id* **track interface vlan** *vlan-id* [**decrement** *priority*]	Specify an interface the virtual router (*vr-id*) on the interface will track. If the interface goes down, the virtual router priority is decreased by the amount specified by the *priority* value.
vrrp *vr-id* **track ip route** *ip-address/prefix-length* [**decrement** *priority*]	Specify a route that the virtual router (*vr-id*) on the interface will track. If the route to the destination network specified by the *ip-address/prefix-length* variable is removed from the routing table, the virtual router priority is decreased by the amount specified by the *priority* value.
CTRL + Z	Exit to Privileged EXEC mode.
show vrrp [*vr-id*]	View settings for all VRRP groups or for the specified VRRP group for the switch.
show vrrp brief	View a summary of interfaces configured to participate in VRRP groups.
show vrrp interface {**brief** | **vlan** *vlan-id* [**stats**]}	View information about VRRP settings configured on all interfaces or on the specified interface. If you specify an interface, use the keyword **stats** to view VRRP statistics for the interface.

VRRP Configuration Example

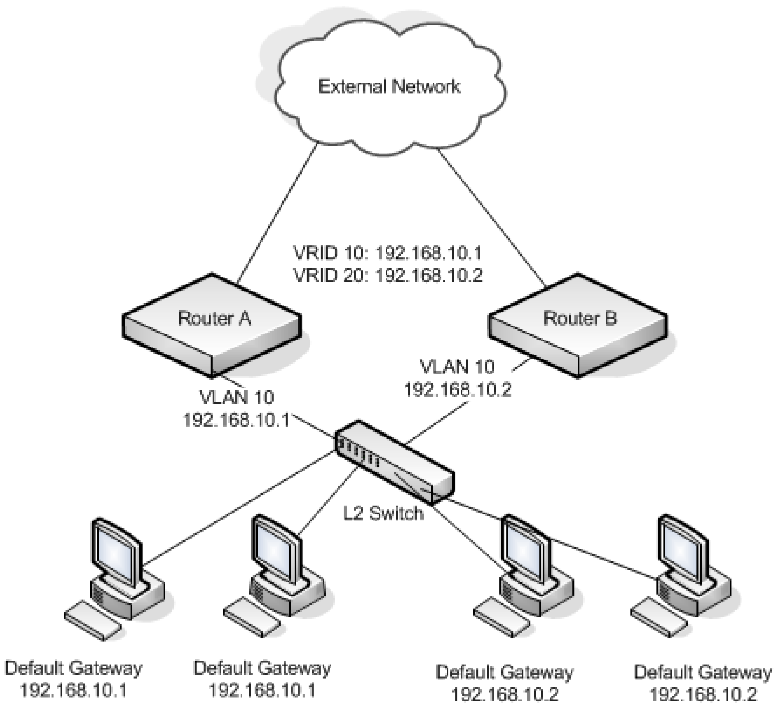
This section contains the following VRRP examples:

- VRRP with Load Sharing
- Troubleshooting VRRP
- VRRP with Route and Interface Tracking
- Configuring VRRP in a VRF

VRRP with Load Sharing

In Figure 38-9, two L3 Dell Networking N-Series switches are performing the routing for network clients. Router A is the default gateway for some clients, and Router B is the default gateway for other clients.

Figure 38-9. VRRP with Load Sharing Network Diagram



This example configures two VRRP groups on each router. Router A is the VRRP master for the VRRP group with VRID 10 and the backup for VRID 20. Router B is the VRRP master for VRID 20 and the backup for VRID 10. If Router A fails, Router B will become the master of VRID 10 and will use the virtual IP address 192.168.10.1. Traffic from the clients configured to use Router A as the default gateway will be handled by Router B.

To configure Router A:

1 Enable routing for the switch.

```
console#config
console(config)#ip routing
```

2 Create and configure the VLAN routing interface to use as the default gateway for network clients. This example assumes all other routing interfaces, such as the interface to the external network, have been configured.

```
console(config)#interface vlan 10
console(config-if-vlan10)#ip address 192.168.10.1 255.255.255.0
console(config-if-vlan10)#exit
```

3 Enable VRRP for the switch.

```
console(config)#ip vrrp
```

4 Assign a virtual router ID to the VLAN routing interface for the first VRRP group.

```
console(config)#interface vlan 10
console(config-if-vlan10)#vrrp 10
```

5 Specify the IP address that the virtual router function will use. This router is the virtual IP address owner (because the routing interface has the same IP address as the virtual IP address for the VRRP group), so the priority value is 255.

```
console(config-if-vlan10)#vrrp 10 ip 192.168.10.1
```

6 Configure an optional description to help identify the VRRP group.

```
console(config-if-vlan10)#vrrp 10 description master
```

7 Assign a virtual router ID to the VLAN routing interface for the second VRRP group.

```
console(config-if-vlan10)#vrrp 20
```

8 Specify the IP address that the virtual router function will use.

```
console(config-if-vlan10)#vrrp 20 ip 192.168.10.2
```

9 Configure an optional description to help identify the VRRP group.

```
console(config-if-vlan10)#vrrp 20 description backup
```

10 Enable the VRRP groups on the interface.

```
console(config-if-vlan10)#vrrp 10 mode
console(config-if-vlan10)#vrrp 20 mode
console(config-if-vlan10)#exit
console(config)#exit
```

The only difference between the Router A and Router B configurations is the IP address assigned to VLAN 10. On Router B, the IP address of VLAN 10 is 192.168.10.2. Because this is also the actual IP address of VRID 20, Router B is the interface owner and VRRP master of VRRP group 20.

To configure Router B:

1 Enable routing for the switch.

```
console#config
console(config)#ip routing
```

2 Create and configure the VLAN routing interface to use as the default gateway for network clients. This example assumes all other routing interfaces, such as the interface to the external network, have been configured.

```
console(config)#interface vlan 10
console(config-if-vlan10)#ip address 192.168.10.2 255.255.255.0
console(config-if-vlan10)#exit
```

3 Enable VRRP for the switch.

```
console(config)#ip vrrp
```

4 Assign a virtual router ID to the VLAN routing interface for the first VRRP group.

```
console(config)#interface vlan 10
console(config-if-vlan10)#vrrp 10
```

5 Specify the IP address that the virtual router function will use.

```
console(config-if-vlan10)#vrrp 10 ip 192.168.10.2
```

6 Configure an optional description to help identify the VRRP group.

```
console(config-if-vlan10)#vrrp 10 description master
```

7 Assign a virtual router ID to the VLAN routing interface for the second VRRP group.

```
console(config-if-vlan10)#vrrp 20
```

8 Specify the IP address that the virtual router function will use.

The router is the virtual IP address owner of this address, so the priority value is 255 by default.

```
console(config-if-vlan10)#vrrp 20 ip 192.168.10.1
```

9 Configure an optional description to help identify the VRRP group.

```
console(config-if-vlan10)#vrrp 20 description backup
```

10 Enable the VRRP groups on the interface.

```
console(config-if-vlan10)#vrrp 10 mode
console(config-if-vlan10)#vrrp 20 mode
console(config-if-vlan10)#exit
console(config)#exit
```

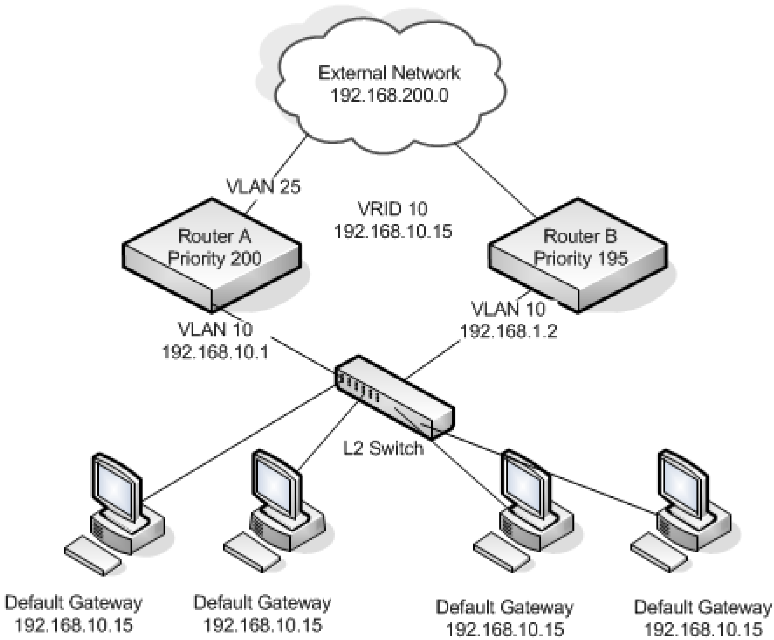
Troubleshooting VRRP

When configuring VRRP, ensure that the layer-2 network facing the VRRP router is up. The VRRP peers will show as being in the “Master” or “Initializing” state until the layer 2 network is operational. Check the spanning-tree state on any routed links. Routed links must show as forwarding. Disable spanning tree on the routed links if necessary to prevent spanning tree from blocking routed links.

VRRP with Route and Interface Tracking

In Figure 38-10, the VRRP priorities are configured so that Router A is the VRRP master, and Router B is the VRRP backup. Router A forwards IP traffic from clients to the external network through the VLAN 25 routing interface. The clients are configured to use the virtual IP address 192.168.10.15 as the default gateway.

Figure 38-10. VRRP with Tracking Network Diagram



Without VRRP interface or route tracking, if something happened to VLAN 25 or the route to the external network, as long as Router A remains up, it will continue to be the VRRP master even though traffic from the clients does not have a path to the external network. However, if the interface and/or route tracking features are configured, Router A can decrease its priority value when the problems occur so that Router B becomes the master.

To configure Router A:

1 Enable routing for the switch.

```
console#config
console(config)#ip routing
```

2 Create and configure the VLAN routing interface to use as the default gateway for network clients. This example assumes all other routing interfaces, such as the interface to the external network, have been configured.

```
console(config)#interface vlan 10
console(config-if-vlan10)#ip address 192.168.10.1 255.255.255.0
console(config-if-vlan10)#exit
```

3 Enable VRRP for the switch.

```
console(config)#ip vrrp
```

4 Assign a virtual router ID to the VLAN routing interface for the VRRP group.

```
console(config)#interface vlan 10
console(config-if-vlan10)#vrrp 10
```

5 Specify the IP address that the virtual router function will use.

```
console(config-if-vlan10)#vrrp 10 ip 192.168.10.15
```

6 Configure the router priority.

```
console(config-if-vlan10)#vrrp 10 priority 200
```

7 Enable preempt mode so that the router can regain its position as VRRP master if its priority is greater than the priority of the backup router.

```
console(config-if-vlan10)#vrrp 10 preempt
```

8 Enable the VRRP groups on the interface.

```
console(config-if-vlan10)#vrrp 10 mode
console(config-if-vlan10)#exit
```

9 Track the routing interface VLAN 25 on VRID 10 so that if it goes down, the priority of VRID 10 on Router A is decreased by 10, which is the default decrement priority value.

```
console(config-if-vlan10)#vrrp 10 track interface vlan 25
```

10 Track the route to the 192.168.200.0 network. If it becomes unavailable, the priority of VRID 10 on Router A is decreased by 10, which is the default decrement priority value.

```
console(config-if-vlan10)#vrrp 10 track ip route 192.168.200.0/24
console(config-if-vlan10)#exit
```

Router B is the backup router for VRID 10. The configured priority is 195. If the VLAN 25 routing interface or route to the external network on Router A go down, the priority of Router A will become 190 (or 180, if both the interface and router are down). Because the configured priority of Router B is greater than the actual priority of Router A, Router B will become the master for VRID 10. When VLAN 25 and the route to the external network are back up, the priority of Router A returns to 200, and it resumes its role as VRRP master.

To configure Router B:

1 Enable routing for the switch.

```
console#config
console(config)#ip routing
```

2 Create and configure the VLAN routing interface to use as the default gateway for network clients. This example assumes all other routing interfaces, such as the interface to the external network, have been configured.

```
console(config)#interface vlan 10
console(config-if-vlan10)#ip address 192.168.10.2 255.255.255.0
console(config-if-vlan10)#exit
```

3 Enable VRRP for the switch.

```
console(config)#ip vrrp
```

4 Assign a virtual router ID to the VLAN routing interface for the VRRP group.

```
console(config)#interface vlan 10
console(config-if-vlan10)#vrrp 10
```

5 Specify the IP address that the virtual router function will use.

```
console(config-if-vlan10)#vrrp 10 ip 192.168.10.15
```

6 Configure the router priority.

```
console(config-if-vlan10)#vrrp 10 priority 195
```

7 Enable preempt mode so that the router can regain its position as VRRP master if its priority is greater than the priority of the backup router.

```
console(config-if-vlan10)#vrrp 10 preempt
```

8 Enable the VRRP groups on the interface.

```
console(config-if-vlan10)#vrrp 10 mode
console(config-if-vlan10)#exit
console(config)#exit
```

Configuring VRRP in a VRF

In this example, a VRRP master is configured in VRF red-1. Interface gi1/0/1 on each of the VRRP peers is connected to the other switch. The configuration steps are as follows:

1 Create the VRRP VLAN:

```
console#config
console(config)#vlan 10
console(config-vlan10)#exit
```

2 Create a VRF and enable routing:

```
console(config)#ip vrf red-1
console(config-vrf-red-1)#ip routing
console(config-vrf-red-1)#exit
```

3 Enable ip routing globally:

```
console(config)#ip routing
```

4 Enable VRRP globally:

```
console(config)#ip vrrp
```

5 Configure a VLAN interface:

```
console(config)#interface vlan 10
```

6 Make the VRF a member of the VLAN:

```
console(config-if-vlan10)#ip vrf forwarding red-1
```

7 Add an IP address to the VLAN to make it a routing VLAN:

```
console(config-if-vlan10)#ip address 129.168.0.1 255.255.255.0
```

8 Create a VRRP instance:

```
console(config-if-vlan10)#vrrp 1
```

9 Set the VRRP virtual address:

```
console(config-if-vlan10)#vrrp 1 ip 129.168.0.100
```

10 Set the VRRP priority and accept pings:

```
console(config-if-vlan10)#vrrp 1 priority 1
console(config-if-vlan10)#vrrp 1 accept-mode
console(config-if-vlan10)#exit
```

11 Configure the physical interface as a VLAN 10 member:

```
console(config)#interface Gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 10
console(config-if-Gi1/0/1)#exit
```

The following steps provide configure the companion VRRP peer:

1 Create a VLAN:

```
console#configure
console(config)#vlan 10
console(config-vlan)#exit
```

2 Create a VRF and enable routing:

```
console(config)#ip vrf red-1
console(config-ip-vrf-red-1)#ip routing
console(config-ip-vrf-red-1)#exit
```

3 Enable ip routing globally:

```
console(config)#ip routing
```

4 Enable VRRP globally:

```
console(config)#ip vrrp
```

5 Configure a VLAN interface:

```
console(config)#interface vlan 10
```

6 Make the VRF a member of the VLAN:

```
console(config-if-vlan10)#ip vrf forwarding red-1
```

7 Add an IP address to the VLAN:

```
console(config-if-vlan10)#ip address 129.168.0.2 255.255.255.0
```

8 Create a VRRP instance:

```
console(config-if-vlan10)#vrrp 1
```

9 Set the VRRP virtual address:

```
console(config-if-vlan10)#vrrp 1 ip 129.168.0.100
```

10 Set the VRRP priority to indicate the other router is the VRRP master and to accept pings:

```
console(config-if-vlan10)#vrrp 1 priority 2
console(config-if-vlan10)#vrrp 1 accept-mode
console(config-if-vlan10)#exit
```

11 Configure the physical interface as a VLAN 10 member:

```
console(config)#interface Gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 10
console(config-if-Gi1/0/1)#exit
```

For VRRP to become active, other interfaces need to be enabled for VLAN 10 such that the VRRP peers are able to establish connectivity to each other over those interfaces as well as over Gi1/0/1.

39

BGP

Dell Networking N3000 and N4000 Series Switches

NOTE: This feature is not available on Dell Networking N1500 and N2000 Series switches. BGP is enabled on Dell Networking N3000 Series switches through use of the AGGREGATION ROUTER firmware.

Border Gateway Protocol (BGP) is a standardized exterior gateway path-vector or distance-vector protocol. BGP makes routing decisions based upon paths and network policies configured by the administrator.

This chapter includes the following topics:

- Overview
- BGP Operations
- BGP Limitations
- BGP Configuration Examples

The following terms and acronyms are used in this chapter.

Table 39-1. BGP-Related Terms

Term	Definition
Accept-RIB-In	The collection of routing information that has passed inbound policy and been accepted as candidate BGP routes.
Adj-RIB-In	The collection of routing information received from peers
Adj-RIB-Out	The collection of routing information sent to peers
AS	Autonomous System
ASN	Autonomous System Number
BGP	Border Gateway Protocol
eBGP	Exterior Border Gateway Protocol
iBGP	Interior Border Gateway Protocol
MED	Multi Exit Discriminator
RIB	Routing Information Base

Table 39-1. BGP-Related Terms

Term	Definition
RTO	Routing Table Object. The common routing table, or "RIB," which collects routes from all sources (local, static, dynamic) and determines the most preferred route to each destination.
TCP	Transmission Control Protocol

Overview

BGP operates by establishing adjacencies (connections) with other BGP peers (routers). BGP peers are configured manually. A BGP speaker (peer) sends a keep-alive message every 30 seconds to the BGP peer to maintain the connections. BGP uses TCP as its transport protocol.

BGP speakers distribute routing information via Network Layer Reachability Information (NLRI) Updates. Normally, Dell Networking BGP distributes routes learned from interior sources only to exterior peers and distributes routes learned from exterior sources to all peers.

Dell Networking BGP supports filtering of learned routes using route-maps for both the in and out directions.

Dell Networking BGP supports IPv4 and IPv6 unicast routes only. Both IPv4 and IPv6 peering are supported. IPv4 routes may be carried over IPv4 peering sessions. IPv6 routes may be carried over IPv4 or IPv6 peering sessions.

The only optional parameters recognized in an OPEN message are the Capabilities option (RFC 5492) and the multiprotocol capabilities option (RFC 4760). The RFC 4271 deprecated Authentication Information option is not supported. If a neighbor includes the deprecated authentication parameter in its OPEN message, Dell Networking BGP rejects the OPEN and will not form an adjacency.

Dell Networking BGP allows the network operator to configure a maximum number of prefixes accepted from a peer. The limit defaults to the maximum number of routes that can be installed in the routing table. When the limit is reached, by default, BGP shuts down the peer. BGP may be configured to instead discard new address prefixes but not terminate the peer (RFC 4271 section 6.7).

Dell Networking BGP supports the following RFCs in whole or in part as indicated:

- RFC 1997 – BGP Communities Attribute
- RFC 2385 – Protection of BGP Sessions via the TCP MD5 Signature Option
- RFC 2545 – Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing
- RFC 2918 – Route Refresh Capability for BGP-4
- RFC 4271 – A Border Gateway Protocol 4 (BGP-4)
- RFC 4273 – Definitions of Managed Objects for BGP-4
- RFC 4456 – BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)
- RFC 4486 – Sub-codes for BGP Cease Notification Message
- RFC 4760 – Multiprotocol Extensions for BGP-4
- RFC 5492 – Capabilities Advertisement with BGP-4

Dell Networking BGP supports both IPv4 and IPv6 peering sessions and supports both IPv4 and IPv6 routes (RFC 4760 and RFC 2545). IPv6 peering sessions support IPv6 routes only, but IPv4 addresses may be embedded within IPv6 routes.

TCP MD5 authentication is supported, however, RFC 5295 is not supported. Dell Networking BGP also supports a private MIB with information regarding:

- Internal BGP message queue status
- Transmit and receive message counters
- Decision process statistics
- Per-peer message and prefix counters

Dell Networking BGP supports configuration via the CLI only.

Routing must be enabled to enable Dell Networking BGP. Both the AS number and the router ID are required to be configured. Enabling of BGP is automatic when the AS number and router ID are configured. The **no enable** command may be used to temporarily disable BGP without removing the BGP configuration.

Autonomous Systems

Dell Networking BGP supports both exterior routing (eBGP) between autonomous systems (inter-AS) and interior routing within an AS (iBGP). Dell Networking BGP is suitable for use in enterprise and data center deployments. Dell Networking switches do not have sufficient capacity to hold a full Internet routing table.

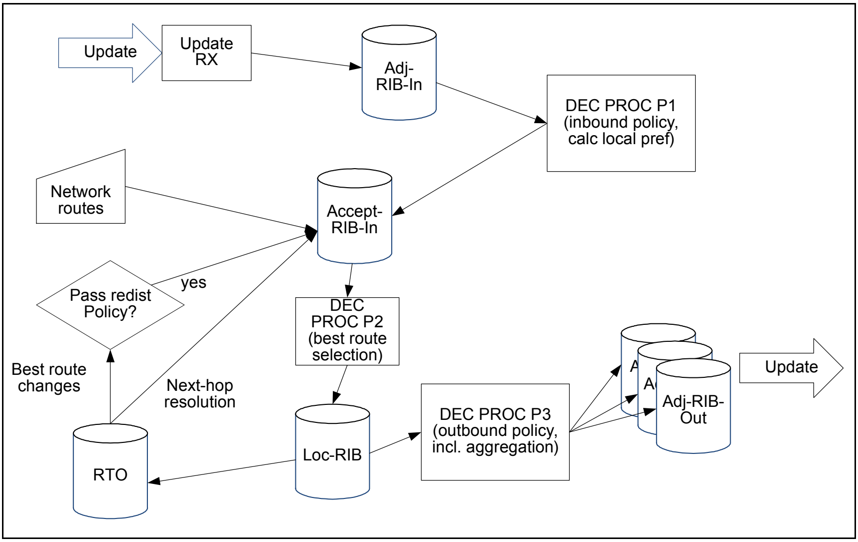
Dell Networking supports BGP version 4 with 2-byte Autonomous System Numbers (ASN). An autonomous system number is a globally unique identifier for a group of IP networks that has a single, clearly defined external routing policy.

BGP Operations

Decision Process Overview

The BGP decision process is logic that applies inbound policy to routing information from peers, computes routes, and advertises routes to peers. Figure 39-1 shows an overview of the decision process. BGP parses incoming UPDATE messages, storing routing information in Adj RIB-In. Phase 1 of the decision process applies inbound policy to routes in Adj RIB In. Routes that pass inbound policy are copied to Accept-RIB-In and LOCAL_PREF is set. BGP uses the routing table to resolve a BGP next hop to a local next hop. Locally originated routes (those configured with the **network** command or redistributed from another protocol) go directly to Accept-RIB-In. Phase 2 of the decision process selects the best route to each destination in Accept-RIB-In. Each best route is stored in the local RIB and given to RTO. Phase 3 of the decision process applies outbound policy to routes in the local RIB and determines the status of aggregate routes. Active aggregates and individual routes that pass outbound policy are placed in an Adj-RIB-Out specific to each update group, and UPDATE messages are sent to communicate the routes to neighbors.

Figure 39-1. BGP Decision Process



BGP Route Selection

Dell Networking BGP uses the following route selection rules in order from 1 to N:

1. Prefer the route with the higher local preference
2. Prefer a locally-originated route over a non-locally originated route
3. Prefer the route with the shorter AS Path
4. Prefer the route with the lower ORIGIN. IGP is better than EGP is better than INCOMPLETE.
5. Prefer the route with the lower MED. By default, MEDs are compared for routes from any AS, but a configuration option limits comparison of MEDs to the same AS. A route with no MED is considered to have a MED of 0.
6. Prefer an eBGP route to an iBGP route
7. Prefer the route with the lower IGP cost to the BGP NEXT HOP
8. Prefer the route learned from the peer with the lower router ID
9. Prefer the route learned from the peer with the lower peer IP address

Limiting Phase 2 CPU Usage

In a network with a large number of prefixes, phase 2 of the decision process can consume a significant amount of time. If the BGP hold timers are configured to be shorter than the duration of the decision process, the timers can expire causing a loss of adjacency. If the decision process runs frequently, it may consume significant CPU resources, starving other processes. Two mechanisms mitigate these potential issues. First, a hold timer prevents phase 2 from running too often. The hold time is explained in more detail in the next paragraph. Second, phase 2 is limited to running no more than approximately one second without yielding to other tasks.

When an event triggers phase 2, a short delay (usually 100 ms) is imposed before the decision process runs. This delay allows other RIB changes to complete before computing new routes. When the trigger occurs, if the decision process was previously run within the hold time, the next decision process is scheduled to run after the hold time expires. The initial hold time is one second. Each time one or more new triggers occur while the hold timer is running, the hold timer is doubled, up to a maximum of 4 seconds. When the hold timer expires without BGP receiving a new phase 2 trigger, the hold time is reset to the minimum hold time.

Path Attributes

Dell Networking supports all path attributes described in RFC 4271.

Dell Networking BGP sets the ORIGIN path attribute to IGP for routes originated through the **network** command and to INCOMPLETE for routes originated through route redistribution. Dell Networking BGP never sets the ORIGIN path attribute to EGP.

Dell Networking BGP sets the AS_PATH path attribute in compliance with RFC 4271. Dell Networking BGP does require that paths from external peers include the configured AS number of the peer as the first AS in the path. Dell Networking BGP enforces a configurable limit to the length of the AS_PATH attribute in received paths. Paths that exceed the limit are discarded.

Dell Networking BGP offers a configuration option (**neighbor next-hop-self**) to set the NEXT_HOP attribute to a local IP address when sending an UPDATE message to an internal peer. Otherwise, Dell Networking BGP follows the guidance in RFC 4271 when sending to internal peers. When sending an UPDATE message to an external peer, Dell Networking BGP retains the NEXT_HOP address if it is an address on the subnet used to connect the peers but is not the peer's IP address. Otherwise, Dell Networking BGP sets the NEXT_HOP path attribute to the local IP address on the interface to the peer. Dell Networking BGP does not support “first party” next hop. Dell Networking does not allow the network operator to disable third party next hop. Dell Networking does not support multihop EBGP. (RFC 4271 section 5.1.3)

The Multi Exit Discriminator (MED) attribute is sent to external peers when advertising routes that originate within the local AS. The MED value may be configured for redistributed routes, either using the **metric** option on the redistribution command or by configuring a **default-metric**. If the MED is not configured for a redistributed route, the route is advertised without a MED attribute. Routes originated through the **network** command set the MED to the metric of the IGP route to the same network. The MED may also be set on locally-originated routes using a route map. The MED for non-locally-originated routes is propagated to internal peers. By default, MEDs are only compared when two routes are received from external peers in the same AS. There is a configuration option to force BGP to compare MEDs for paths received from different autonomous systems.

When BGP receives an UPDATE message from an external peer, it assigns a local preference value during phase 1 of the decision process. Local preference is set to a fixed, configured value which is the same for all paths received from all neighbors. This value is attached to the path in the LOCAL_PREF path attribute when the path is advertised to internal peers. The configured default local preference is assigned to all locally-originated routes and to the paths for all active aggregate addresses. LOCAL_PREF can be configured to different values on different routers to influence the exit point from the AS that other routers select for each destination. An inbound route map can override the default local preference. LOCAL_PREF is never included in paths sent to external peers. If the user changes the default local preference while BGP is running, BGP automatically initiates an immediate soft inbound reset for all external peers, updates the local preference for all locally-originated routes, and re-computes routes.

For each aggregate address configured, the network administrator may specify whether to advertise an AS_SET of the AS numbers in the paths from which the aggregate was formed. When the aggregate is advertised with an empty AS Path, the ATOMIC_AGGREGATE path attribute is attached to the path. In either case, the AGGREGATOR path attribute is attached.

BGP Finite State Machine (FSM)

Dell Networking BGP supports all mandatory FSM session attributes and the following optional session attributes (RFC 4271 section 8):

- AllowAutomaticStart—Connections are automatically restarted after an error closes a connection. An adjacency to an external peer in the IDLE state is automatically started if the routing interface to that peer comes up. An adjacency to an internal peer in the IDLE state is automatically started when the peer's IP address becomes reachable.
- AllowAutomaticStop—When a neighbor sends more prefixes than the configured limit, the connection may be automatically shut down, depending on configuration. AutoStop is also used for fast fallover. When the routing interface to an external peer goes down, the peering session is automatically stopped. Similarly, if an internal peer becomes unreachable, the peering session is automatically stopped.
- CollisionDetectEstablishedState—When an OPEN message is received on a TCP connection and the adjacency using that connection has already reached ESTABLISHED state, the adjacency is cleared. If an OPEN message is received on a different TCP connection than the one used to reach ESTABLISHED state, the new TCP connection is cleared and the adjacency remains up.
- DampPeerOscillations—An idle hold time is enforced between automatic restarts. The length of the idle time depends on the reason the adjacency entered the idle state.
- IdleHoldTime/IdleHoldTimer—After an error clears a connection or a TCP connection fails, Dell Networking BGP waits before attempting to reestablish the adjacency. The waiting time varies depending on the event. When a TCP connection fails, BGP waits 30 to 60 seconds. When a NOTIFICATION is received, BGP waits 1 to 2 seconds. Other events trigger a wait of 10 to 20 seconds. The delay time is not configurable.
- SendNOTIFICATIONWithoutOPEN—Dell Networking will accept a NOTIFICATION packet from a peer that has not first sent an OPEN packet. Dell Networking will not send a NOTIFICATION without first sending an OPEN.

None of the optional session attributes are configurable.

Dell Networking BGP supports manual start and stop events. A manual start event occurs when the user first configures a peer (**neighbor remote-as**) or administratively enables a peer (**no neighbor shutdown**). A manual stop event occurs when the user administratively disables a neighbor (**neighbor shutdown**).

Of the optional events in RFC 4271 section 8.1.2 - 8.1.5, the following events are supported:

- AutomaticStart_with_DampPeerOscillations (Event 6)
- AutomaticStop (Event 8)
- IdleHoldTimer_Expires (Event 13)

When an attempt to establish an adjacency fails, Dell Networking puts the adjacency in the IDLE state and starts the idle hold timer. When the idle hold timer expires, Dell Networking moves the adjacency to the CONNECT state and initiates a new TCP connection. If the neighbor does not respond to the connection request, Dell Networking retries three times. The first retry is done after the configured retry interval. Each subsequent retry doubles the previous retry interval. So by default, the TCP connection is retried after 2, 4, and 8 seconds. Configuring the initial retry interval to a large value can prevent retries. The TCP stack times out a connection attempt in 20 seconds (after retransmitting the SYN segment according to normal retransmit procedures), sends a TCP reset, and notifies the application of the connection failure. The connection failure resets the BGP connect retry timer, puts the adjacency in IDLE state, and starts the idle hold timer.

Dell Networking BGP allows multiple BGP sessions between the same two routers. However, each session must be established between different pairs of IP addresses.

Dell Networking BGP includes two capabilities in every OPEN message it sends. The first is the Route Refresh capability described in RFC 2918. The second is the multiprotocol capability described in RFC 4760. Dell Networking always advertises the IPv4/unicast AFI/SAFI pair. If the user has activated IPv6 for the peer, the OPEN also includes the IPv6/unicast pair. Even though Dell Networking BGP does not support any AFI/SAFI pairs other than IPv4/unicast when IPv6 is not enabled, Dell Networking advertises the multiprotocol capability with IPv4/unicast because some other implementations appear to require this in order to establish an adjacency.

Detecting Loss of Adjacency

Dell Networking optionally drops an adjacency with an external peer when the routing interface to that peer goes down. This behavior can be enabled globally or on specific interfaces using the **bgp fast-external-fallover** and **ip bgp fast-external-fallover** commands. BGP accomplishes this behavior by listening to router events. When BGP gets a routing interface down event, BGP drops the adjacency with all external peers whose IPv4 address is in one of the subnets on the failed interface. Both **fast-external-fallover** and **fast-internal-fallover** are enabled by default.

Dell Networking also offers an option to quickly detect loss of reachability to internal peers, and drop the BGP adjacency when such a loss occurs. Because internal peers are often not on a local subnet (and an internal peer can be reached through multiple local interfaces), BGP cannot determine internal peer reachability based on local link state. Instead, when this feature is enabled, BGP registers for address resolution changes for each internal peer's IPv4 address. When a peer's address becomes unreachable (i.e., the route table manager deletes the route to the peer and no non-default route to the peer remains), BGP drops the adjacency to the peer. BGP considers an internal peer to be unreachable if the only route to the peer is a default route. This feature can enabled or disabled globally for all internal peers using **bgp fast-internal-fallover**. Because internal peers are not associated with a single interface, there is no interface configuration option.

When fast fallover is disabled for a peer, the adjacency remains in the ESTABLISHED state until the hold timer expires. When connectivity to the peer is lost, the BGP Next Hop for routes learned from affected peers becomes unreachable. This change makes the routes unusable, and BGP immediately removes them from the routing table. So even without the fast fallover behavior enabled, the routing table reacts quickly to changes in local interface state. However, when the adjacency remains in ESTABLISHED state even though the neighbor is unreachable, BGP cannot send UPDATE messages to the neighbor. If the link is restored before the dead interval expires, there is no event to cause BGP to resend the failed UPDATEs. Because BGP does not periodically refresh routing state, and the loss of UPDATEs is permanent. To avoid this situation, when an UPDATE message fails to be sent to any member of an outbound update group, BGP reschedules the update send process to resend the data. Thus, having a neighbor in an ESTABLISHED but unreachable state causes duplicate data to be sent to other members of the update group. With fast fallover enabled, the adjacency to the unreachable neighbor is no longer ESTABLISHED, and if an UPDATE is sent to the neighbor's update group, BGP does not try to send to the failed neighbor. When the failed adjacency is reestablished, BGP resends all routing information to the neighbor.

Both internal and external fallover should happen within a second of the loss of reachability. Enabling fast fallover should relax the need to set a short hold time and send KEEPALIVE messages rapidly.

Authentication

RFC 4271 requires support for TCP MD5 authentication as specified in RFC 2385. Dell Networking supports TCP MD5 authentication. The network administrator may optionally enable TCP MD5 for a specific peering session by configuring a password on each end of the connection.

Because of concerns about the increasing vulnerability of MD5, the IETF has recently obsoleted RFC 2385, replacing it with more robust mechanisms specified in RFC 5925, *The TCP Authentication Option*. In spite of this, support for TCP MD5 has some near-term value: it allows interoperability with other implementations that do not yet support RFC 5925. Dell Networking BGP does not support for RFC 5925.

Outbound Update Groups

To reduce the memory required for the Adj-RIB-Out and to reduce the processing required by the phase 3 decision process, BGP sorts peers into update groups. Every peer in an update group has the same configured (or default) value for minRouteAdvertisementInterval and the same set of outbound policies. Each update group contains only internal or external peers. Thus, the same information is advertised to every peer in the update group and may be advertised at the same time. A single advertised path list (Adj-RIB-Out) is retained for each update group. A single UPDATE message is constructed and a copy sent to each peer in the update group. When a peer in the ESTABLISHED state moves from one update group to another because of a configuration change, BGP withdraws all prefixes previously advertised to the peer and advertises to the peer the Adj-RIB-Out of the new update group.

BGP maintains separate update groups for IPv4 and IPv6. If IPv6 is active for a peer with an IPv4 address, the peer is in both an IPv4 update group and in an IPv6 update group. A neighbor may be in an IPv6 update group for an IPv4 peer session (if the network administrator activates IPv6 on the peer session) and in an IPv6 update group for an IPv6 peer session. Such a configuration is probably a misconfiguration. BGP will send IPv6 NLRI to the neighbor twice.

BGP assigns peers to update groups automatically. The Dell Networking UI has no configuration associated with update groups and the UI does report update group membership.

Removing Private AS Numbers

An organization may use private AS numbers internally. Private ASNs must be removed from routes to destinations within an AS before the routes are advertised in order to avoid conflicts with other networks also using overlapping private ASNs. Dell Networking BGP may be configured to remove private AS numbers from the AS_PATH attribute of paths advertised to external peers as an outbound policy.

Two-byte ASNs in the range from 64,512 to 65,535 are removed when this option is enabled. The administrator may optionally configure BGP to replace private ASNs with the local AS number. The replace option maintains the original length of an AS path, which can be important when the AS path length is used in route selection. The option to remove or replace private ASNs can be configured independently for each address family (IPv4 or IPv6).

Templates

Dell Networking BGP supports configuration of neighbor parameters in named peer templates. A template defines a set of peer parameters. Multiple peers can inherit parameters from a template, eliminating the need to repeat common configuration for every peer. A neighbor can inherit from a single template. BGP accepts configuration of up to 32 templates.

Neighbor configuration parameters can be divided into two groups, session parameters and policy parameters. Session parameters apply to the peering session. Session parameters include configuration options such as keep-alive and hold timers. Policy parameters are specific to the routes for an address family (e.g., IPv4 and IPv6), such as the maximum number of routes accepted from a peer or prefix lists used to filter routes received from or sent to a peer. Peer templates allow both session parameters and policies to be configured within the same template. With a template, policy parameters are configured for a specific address family.

Session parameters that may be configured in a template are as follows:

Table 39-2. Configurable Session Parameters in BGP Peer Templates

Parameter	Description
allowas-in	Configure to accept routes with my ASN in the as-path.
connect-retry-interval	Configure the connection retry interval for the peer.
description	Configure a description for the peer.
ebgp-multihop	Configure to allow non-directly-connected eBGP neighbors.
fall-over	Configure fast fall-over.
local-as	Configure local-as.
password	Configure a TCP password.
remote-as	Configure remote-as.
rfc5549-support	Configure support of RFC 5549.
shutdown	Configure the administrative status.
timers	Configure keepalive and hold time.
update-source	Configure a source address.

Policy parameters that may be configured per address family within a template are as follows:

Table 39-3. Session Parameters in BGP Peer Templates—Configurable Per-Address Family

Parameter	Description
advertisement-interval	Configure the BGP advertisement interval for the peer.
default-originate	Configure this peer to generate a default route.
filter-list	Configure filter lists for the peer.
maximum-prefix	Configure the maximum number of prefixes learned from the peer.
next-hop-self	Configure the router as next hop.
prefix-list	Configure prefix lists for the peer.

Table 39-3. Session Parameters in BGP Peer Templates—Configurable Per-Address Family

Parameter	Description
remove-private-as	Remove private ASNs from AS_PATH when sending to inheriting peers.
route-map	Configure a route map for the peer.
route-reflector-client	Configure a peer as a route reflector client.
send-community	Configure this peer to send BGP communities.

Resolving Interface Routes

In Dell Networking, the next hop of a route is always a set of next-hop IP addresses. Dell Networking does not support routes whose next hop is simply an interface. Thus, the second route resolvability condition in RFC 4271 section 9.1.2.1 does not apply.

Originating BGP Routes

A router running Dell Networking BGP can originate a BGP route through route redistribution and through configuration (the **network** command). Attributes of locally-originated routes may be set through a route map. Locally-originated BGP routes are sent to both internal and external peers unless filtered by outbound policy.

Locally-originated routes are added to Accept-RIB-In. Phase 2 of the decision process considers locally-originated routes along with routes received from peers when selecting the best BGP route to each destination.

BGP can be configured to originate the same prefix through a network command and through redistribution. Dell Networking BGP creates a different path for each if the path attributes differ. BGP only advertises the prefix with the preferred path.

RFC 4271 section 9.2.1.2 specifies “a minimum amount of time that must elapse between successive advertisements of UPDATE messages that report changes within the advertising BGP speaker's own autonomous systems” and refers to this as minASOriginationInterval. RFC 4271 section 10 suggests a default of 15 seconds. Dell Networking BGP does not enforce minASOriginationInterval, but relies on minRouteAdvertisementInterval, which is applied to all advertisements, to dampen flaps of locally-originated routes. Delay and hold timers limit how often phase 2 of the decision process runs. This phase 2 dampening limits route origination, as does IP event dampening when interface flaps would otherwise cause rapid origination.

BGP originates a default route to all neighbors if the **default-information originate** command is given and the default route is among the routes BGP redistributes. Because this default origination depends on redistribution, BGP normally only originates a default if a default is in the routing table. The **always** option can be configured to relax this requirement. If the routing table does not contain a default route, but the network administrator wants BGP to originate a default route, the administrator can configure a static default route. To prevent the static default route from affecting the local router's forwarding, the default route can be given a preference of 255 (ip route 0.0.0.0 0.0.0.0 next-hop 255) or it can be configured as a reject route (**ip route 0.0.0.0 0.0.0.0 Null0**).

BGP can also originate a default to a specific neighbor using **neighbor default-originate**. This form of default origination does not install a default route in the BGP routing table (it will not appear in **show ip bgp**), nor does it install a default route in the Adj-RIB-Out for the update group of peers so configured (it will not appear in **show ip bgp neighbor advertised-routes**). A neighbor specific default has no MED and the Origin is IGP. A neighbor-specific default is only advertised if the Adj-RIB-Out does not include a default learned by other means, either from **default-information originate** or a default learned from a peer. This type of default origination is not conditioned on the presence of a default route in the routing table.

Equal Cost Multipath (ECMP)

By default, Dell Networking BGP selects a single next hop for each BGP route. Dell Networking BGP can be configured to install BGP routes with up to 16 next hops in the common routing table (RTO). The network administrator can independently configure the maximum number of next hops for routes through internal and external peers.

Paths can be used to form an ECMP route when they are both internal or both external, the resolved next hop is different, and the following attributes are the same:

- local preference
- AS path length

- origin
- MED
- IGP distance to the BGP next hop

Dell Networking BGP does not require ECMP next hops to be in a common AS. This behavior is enabled by default. To disable this behavior, use the **no bgp always-compare-med** command.

When advertising to neighbors, BGP always advertises the single best path to each destination prefix, even if BGP has an ECMP route to a destination.

NOTE: The maximum ECMP width is limited by the chosen SDM template. Both the Dell Networking N3000 Series switches and the Dell Networking N4000 Series switches can support 16-wide ECMP when using a non-default SDM template.

BGP Next-Hop Resolution

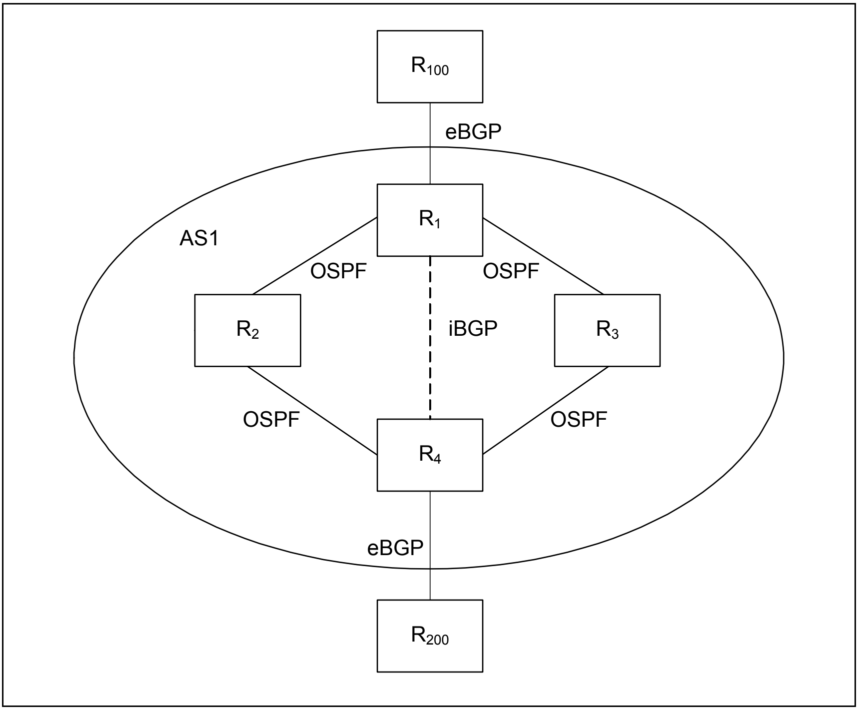
BGP UPDATE messages specify a NEXT_HOP attribute for each prefix. The NEXT_HOP attribute may be on an attached subnet for the receiver when the UPDATE is received from an external peer. But the NEXT_HOP on routes from internal peers or a multihop external peer is not always on a local subnet. Thus, BGP has to resolve the BGP NEXT_HOP to one or more local next hops (similar to how a router resolves a tunnel endpoint to a local next hop). BGP resolves a remote NEXT_HOP by asking RTO for the longest prefix match. As the routing table changes, the resolution for a NEXT_HOP may change. BGP registers each remote BGP NEXT_HOP with RTO for next hop resolution changes.

When RTO notifies BGP of a next hop resolution change, BGP finds all the paths whose BGP NEXT_HOP is the IP address whose resolution changed and updates the immediate next hops for each path. A next hop resolution change triggers phase 2 of the decision process for the affected prefixes.

Dell Networking allows up to 512 addresses to be registered for next-hop resolution changes. This should be sufficient for BGP. The number of addresses BGP needs to track is limited to the number of external peers to the router's autonomous system (not just the external peers for the router itself) or, if routers are configured to advertise themselves as the next hop (next-hop-self), to the number of internal peers.

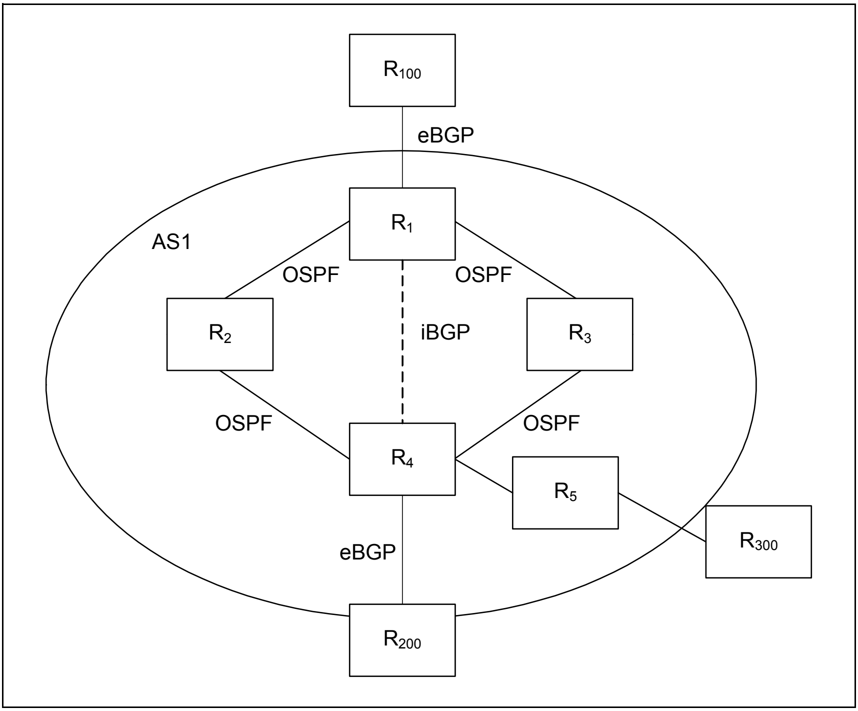
A BGP NEXT_HOP can resolve to an ECMP IGP route. When BGP is configured to allow ECMP iBGP routes, the BGP NEXT_HOP resolves to multiple next hops. BGP retains up to the number of resolved next hops allowed for an iBGP route. For example, in Figure 39-2, R4 receives an iBGP route from internal peer R1. The BGP NEXT_HOP of this path resolves to an ECMP OSPF route through R2 and R3. If BGP is configured on R4 to allow ECMP iBGP routes, then R4 will resolve the path's BGP NEXT_HOP to a pair of next hops through R2 and R3.

Figure 39-2. ECMP NEXT_HOP Resolution



When iBGP paths are combined into an ECMP route, their next-hop sets are merged to form the set of next hops for the route. For example, in Figure 39-3, if R4 learns another route via R5 and R300 with the same destination as the route in the previous example, and the path from R300 is equivalent to the path through R100, then R4 will install a route using R2, R3, and R5 as next hops.

Figure 39-3. Combining iBGP Routes



Address Aggregation

Dell Networking BGP supports address aggregation. The network administrator can configure up to 128 aggregate addresses. BGP compares active prefixes in the local RIB to the set of aggregate addresses. To be considered a match for an aggregate address, a prefix must be more specific (i.e., have a longer prefix length) than the aggregate address. A prefix whose prefix length equals the length of the aggregate address is not considered for aggregation. If one or more prefixes fall within an aggregate, the aggregate is considered active. A prefix must be used for forwarding to be considered for inclusion in an aggregate address (unless it is a locally-originated prefix). Aggregate addresses may overlap (for example, 10.1.0.0/16 and 10.0.0.0/8). A prefix that matches overlapping aggregates is considered to match only the aggregate with the longest mask. When an aggregate address becomes active (that is, when the first contained route is matched to the aggregate), BGP adds a discard route to RTO with prefix and network mask equal to those defined for the aggregate address. Aggregate addresses apply to both locally-originated routes and routes learned from peers.

Address aggregation is done prior to application of outbound policy. Thus, an active aggregate may be advertised to a neighbor, even if the outbound policy to the neighbor filters all of the aggregate's more specific routes (but permits the aggregate itself).

An aggregate address is advertised with a set of path attributes derived from the best paths for each NLRI included in the aggregate. Path attributes of the aggregate are formed as follows:

- If one or more aggregated routes have ORIGIN set to INCOMPLETE, the aggregate path sets ORIGIN to INCOMPLETE. Otherwise, if one or more routes has an ORIGIN set to EGP, the aggregate path sets the ORIGIN to EGP. Otherwise, the ORIGIN is set to IGP in the aggregate path.
- The local preference is set to the default local preference configured on the router that creates the aggregate. (Of course, if the aggregate is advertised to an external peer, the local preference is not included.)
- The NEXT_HOP is not imported from the aggregated routes. It is always set to the local IPv4 address of the TCP connection to the peer.
- If the **as-set** option is configured for an aggregate, then the aggregate is advertised with a non-empty AS_PATH. If the AS_PATH of all contained routes is the same, then the AS_PATH of the aggregate is the AS_PATH of the contained routes. Otherwise, if the contained routes have different AS_PATHs, the AS_PATH attribute includes an AS_SET with each of the AS numbers listed in the AS PATHs of the aggregated routes. If the **as-set** option is not configured, the aggregate is advertised with an empty AS_PATH.
- If BGP is configured to aggregate routes with different MEDs, no MED is included in the path for the aggregate. Otherwise, if the **as-set** option is not configured the aggregate MED is set to the MED for the aggregated routes. If the **as-set** option is configured and the first segment in the AS Path is an AS SET, then no MED is advertised.
- If the **as-set** option is configured, the aggregate's path does not include the ATOMIC_AGGREGATE attribute. Otherwise, it does.
- The AGGREGATOR attribute is always included.

- If the individual routes have communities and the aggregate does not have the ATOMIC_AGGREGATE attribute set, the aggregate is advertised with the union of the communities from the individual routes. If the aggregate carries the ATOMIC_AGGREGATE attribute, the aggregate is advertised with no communities.

Dell Networking BGP never aggregates paths with unknown attributes. By default, Dell Networking BGP does not aggregate paths with different MEDs, but there is a configuration option to allow this.

Routing Policy

Route maps are used to implement redistribution policy, i.e., they are used to filter routes and change route attributes. Route maps may be applied in the inbound direction (before the P2 decision process) or in the outbound direction (after best route selection but before redistribution to peer routers). In either direction, the list of configured route maps is processed in order of increasing sequence number.

Match clauses in a route map are processed as a logical AND, i.e., all match clauses in a route-map must match for the route to be processed. If a route matches all relevant match clauses in a route map, the route is permitted or denied and route-map list processing ceases for that route. (Some match clauses are irrelevant to routing policy, e.g., match length). ACLs are not supported for matching prefixes when filtering routes. Use the **ip prefix-list** command instead. Matching on AS numbers, community, or extended communities is also supported.

In processing the list of route maps, each route map is processed individually. When a match occurs, if the route is permitted, the set clauses are processed and list processing ceases. If the route is denied, no set clauses are processed and list processing ceases. If no match occurs, list processing continues with the next highest numbered route map.

Unlike PBR, routing policy route maps implement an implicit deny at the end of the last route map that filters out all routes. Routes that do not match any match clause reach the end of the route map list and are removed.

An empty route map has a defined action. An deny route map with no match clause matches all routes and does not perform any set actions since the action is deny. Processing of the route-map list ceases. An empty permit route map permits all routes, performs any set clauses, and stops processing of the route-map list since it matched a route.

Inbound Policy

An inbound policy is a policy applied to UPDATE messages received from peers. Dell Networking BGP supports the following inbound policies which are matched against incoming route updates in the order below:

- A global prefix filter that applies to all neighbors (**distribute-list in**)
- A per-neighbor AS path filter (**neighbor filter-list in**)
- A prefix filter that applies to a specific neighbor (**neighbor prefix-list in**)
- A per-neighbor route map (**neighbor route-map in**)

These policy mechanisms determine whether to accept or reject routes received from neighbors. A route map may also change the attributes of received routes.

Within a route map, match terms are considered in the following order:

- AS path list
- Prefix list
- Community list

When processing list terms, a match for any term indicates a match and processing stops.

Outbound Policy

An outbound policy is a policy applied to BGP's best routes (those in the local RIB and active aggregates) and determines which routes are advertised to each peer. The route map option may also change the attributes advertised to a peer. Dell Networking BGP supports the following outbound policies which are matched against outgoing routes in the order below:

- A per-neighbor AS path filter (**neighbor filter-list out**)
- A prefix filter that applies to all neighbors (**distribute-list out**)
- A prefix filter that applies to a specific neighbor (**neighbor prefix-list out**)
- A per-neighbor route map (**neighbor route-map out**)

Within a route map, match terms are considered in the following order:

- AS path list
- Prefix list
- Community list

When processing list terms, a match for any term indicates a match and processing stops.

Routing Policy Changes

When the user makes a routing policy configuration change, Dell Networking BGP automatically applies the new policy. Like any other configuration change, routing policy changes are immediately saved in the running configuration, as soon as the user enters the command.

Even though policy configuration changes are committed to the running configuration immediately, they do not take operational effect until three minutes after the last configuration change. The delay allows the user time to make other configuration changes or correct any mistakes before the change takes effect. If another event, such as receipt of an UPDATE message or a neighbor established event triggers the decision process while waiting for the three minutes to expire, then the decision process runs at the time of the event using the old policy configuration. To immediately apply policy changes, the **clear ip bgp** command can be issued to trigger an immediate soft reset.

In response to a change to an outbound policy, BGP recomputes update group membership and advertises updates to the affected peer to reflect the change in policy.

In response to a change of an inbound policy, BGP schedules phase 1 of the decision process. If the policy change is neighbor-specific, phase 1 only re-evaluates routes received from that neighbor. If the change is global, phase 1 re-evaluates all routes. If an affected neighbor supports Route Refresh, BGP sends a ROUTE REFRESH message to the neighbor, and applies the new policy to the UPDATE messages received in response. If a neighbor does not support Route Refresh, BGP applies the new policy to path information previously received from the neighbor. As with outbound policy, inbound policy changes are immediately committed to the running configuration but do not take effect for three minutes. The soft reset is deferred for three minutes to allow configuration changes to be finalized before they are applied. The **clear ip bgp** command can be issued to trigger an immediate soft reset, if desired.

At startup, when the saved configuration is applied, there could potentially be a lot of churn to outbound update groups and filtering of routing information. This startup churn is avoided by keeping BGP globally disabled until after the entire configuration is applied and the status of all routing interfaces is known.

BGP Timers

Dell Networking BGP supports the five mandatory timers described in RFC 4271 section 10. Dell Networking BGP employs the optional IdleHoldTimer, but does not support a DelayOpenTimer.

When Dell Networking BGP initiates a TCP connection to a peer, it starts a retry timer (ConnectRetryTimer from RFC 4271). If the connection is not established before the retry timer expires, BGP initiates a new TCP connection attempt. Up to 3 retries are attempted with an exponential back-off of the retry time. The initial retry time is configurable per neighbor.

The IDLE hold timer runs when a peer has automatically transitioned to the IDLE state. When the IDLE hold timer expires, BGP attempts to form an adjacency. The IDLE hold time is a jittered to avoid synchronization of retries. The idle hold time varies depending on the event that triggered the transition to IDLE.

Dell Networking BGP starts hold and keep-alive timers for each peer. When BGP establishes an adjacency, the neighbors agree to use the minimum hold time configured on either neighbor. BGP sends KEEPALIVE messages at either 1/3 of the negotiated hold time or the configured keepalive interval, whichever is more frequent. Keepalive times are jittered.

RFC 4271 section 9.2.1.1 specifies a “minimum amount of time that must elapse between an advertisement and/or withdrawal of routes to a particular destination by a BGP speaker to a peer.” In Dell Networking BGP, this advertisement interval is configurable independently for each peer, defaulting to 30 seconds for external peers and 5 seconds for internal peers. The advertisement interval may be configured to 0. Dell Networking BGP enforces the advertisement interval by limiting how often phase 3 of the decision process can run for each outbound update group. The advertisement interval is applied to withdrawals as well as active advertisements.

Communities

Dell Networking BGP supports BGP standard communities as defined in RFC 1997. Dell Networking supports community lists for matching routes based on community, and supports matching and setting communities in route maps. Dell Networking BGP recognizes and honors the following well-known communities (RFC 1997):

- NO_EXPORT—A route carrying this community is not advertised to external peers.
- NO_ADVERTISE—A route carrying this community is not advertised to any peer.
- NO_EXPORT_SUBCONFED—A route carrying this community is not advertised to external peers.

If Dell Networking receives an UPDATE message with more than 512 communities, a NOTIFICATION message is returned to the sender with error UPDATE message/attribute length error.

Routing Table Overflow

BGP Routing Table

Device configuration errors and other network transients can cause temporary or sustained spikes in the BGP routing table size. To protect the router from allocating too much memory in these scenarios, Dell Networking BGP limits the BGP routing table size. The limit is set to the number of routes supported by the routing table (RTO). BGP imposes separate limits for each address family it supports. Once the BGP routing table is full, new routes computed in phase 2 of the decision process are not added to RTO and are not used for forwarding, but are advertised to neighbors. When the BGP routing table becomes full, a log message is written to the log warning the administrator. While BGP remains in this state, it periodically writes a log message that states the number of NLRI routes that could not be added to the routing table.

BGP automatically recovers from a temporary spike in BGP routes above this limit. When BGP cannot add a route to the BGP routing table, it sets the phase 2 pending flag on that NLRI in the Accept RIB. While there are NLRI in this state, BGP periodically checks if there is space available in the BGP routing table, and if so, runs phase 2. When space becomes available in the BGP routing table, these routes are added.

RTO Full Condition

If BGP computes a new route but the routing table does not accept the route because it is full, BGP flags the route as one not added to RTO. BGP periodically tries to add these routes to RTO. BGP will continue to advertise the best routes to neighbors, even if they are not added to RTO. The only necessary condition for forwarding a route to the neighbor is that the route is the best route in the BGP database. When used in conjunction with VRFs, this rule may make routing black holes likely unless the network capacity is planned correctly.

Route Reflection

Dell Networking BGP can be configured as a route reflector as described in RFC 4456. Like any BGP implementation, Dell Networking BGP can also act as a route reflector client. Route reflection eliminates the need to configure a full mesh of iBGP peering sessions. As its name implies, this feature allows a router to reflect a route received from an internal peer to another internal peer. Under conventional BGP rules, a router can only send routes learned from an external peer or routes locally originated to an internal peer. A route reflector will advertise routes learned from an iBGP speaker designated as a route reflector client to other iBGP speakers.

The administrator can configure an internal BGP peer to be a route reflector client. Alternatively, the administrator can configure a peer template to make any inheriting peers route reflector clients. The client status of a peer can be configured independently for IPv4 and IPv6.

A cluster may have multiple route reflectors. Route reflectors within the same cluster are configured with a cluster ID. When a route reflector reflects a route, it prepends its cluster ID to a list of cluster IDs in the CLUSTER_LIST attribute.

RFC 4456 notes that:

> "...when a RR reflects a route, it SHOULD NOT modify the following path attributes: NEXT_HOP, AS_PATH, LOCAL_PREF, and MED. Their modification could potentially result in routing loops."

For this reason, if a route reflector client has an outbound neighbor route-map configured, the set statements in the route map are ignored.

VRF Support

Dell Networking switches that support BGP and VRFs also support BGP in VRFs. When configured in a VRF, BGP runs independent sessions to neighbors in the VRF and forwards independently.

BGP Neighbor Configuration

Dell Networking BGP supports configuration of eBGP neighbors that are not directly connected using the **ebgp-multihop** parameter to the **neighbor** command. Multi-hop is supported only for eBGP configurations where the neighbor has a different AS number. Dell Networking iBGP requires neighbors to be directly connected.

Dell Networking BGP also supports auto-detection of a neighbor's IPv6 link local address. The BGP neighbor must be directly connected and the link must be configured to use IPv6 addressing.

Extended Communities

Dell Networking BGP supports standard extended communities as defined in RFC 4360. Dell Networking BGP supports extended community lists for matching routes based on the extended community and supports matching and setting of extended communities in route maps. It also supports selective export and import of routes using export and import maps.

The extended community attribute provides a mechanism for labeling routes carried in BGP-4. These labels are then used to control the distribution of routes among VRFs. The extended community attribute is sent by BGP when configured in the default VRF only, i.e., as part of an MP-BGP configuration.

A BGP NLRI can carry both standard and extended community attributes and it can also carry multiple community attributes through the use of the additive keyword in the case of standard communities, and through the use of route-maps when exporting the VRF routes in the case of extended communities.

Extended Community Attribute Structure

Each Extended Community attribute has a community type code of 16 and is encoded into an 8-octet value. The first 2 octets are the attribute type and the remaining 6 octets contain the value of attribute. The values from 0 through 0x7FFF are assigned by IANA and values from 0x8000 through 0xFFFF are vendor-specific.

The Extended Community attribute may be represented in multiple ways but Dell Networking supports only the following two formats:

Two-octet AS specific Extended Community

This is an extended type with Type Field composed of 2 octets and Value Field composed of 6 octets.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 0x00 or 0x40  |   Sub-Type    |    Global Administrator       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Local Administrator                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The value of the high-order octet of this extended type is either 0x00 or 0x40. The low-order octet of this extended type is used to indicate sub-types.

The Value Field consists of two sub-fields:

- Global Administrator sub-field: 2 octets

  This sub-field contains an Autonomous System number assigned by IANA.

- Local Administrator sub-field: 4 octets

  The organization identified by Autonomous System number in the Global Administrator sub-field can encode any information in this sub-field. The format and meaning of the value encoded in this sub-field should be defined by the sub-type of the community.

IPv4 address specific Extended Community

This is an extended type with Type Field composed of 2 octets and Value Field composed of 6 octets.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 0x01 or 0x41  |   Sub-Type    |    Global Administrator       |
```

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Global Administrator (cont.)  |    Local Administrator        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The value of the high-order octet of this extended type is either 0x01 or 0x41. The low-order octet of this extended type is used to indicate sub-types.

The Value field consists of two sub-fields:

- Global Administrator sub-field: 4 octets

  This sub-field contains an IPv4 unicast address assigned by one of the Internet registries.

- Local Administrator sub-field: 2 octets

  The organization that has been assigned the IPv4 address in the Global Administrator sub-field can encode any information in this sub-field. The format and meaning of this value encoded in this sub-field should be defined by the sub-type of the community.

Types of Extended Communities

Dell Networking BGP recognizes and honors the following well-known extended community attributes from RFC 4360:

- Route target community
- Route origin community

Route Target Community

The Route Target Community identifies one or more routers that may receive a set of routes (attached with this community) carried by BGP. This community is transitive across the Autonomous System boundary.

The value of the high-order octet of the Type field for the Route Target Community can be 0x00, 0x01 or 0x02. The value of the low-order octet (Sub-type) of the Type field for this community is 0x02 (if represented in Two-octet AS specific format) and 0x102 (if represented in IPv4 address specific format).

Possible uses of the Route Target Community attribute are described in the following sections.

Route Origin Community Attribute

The Route Origin Community attribute identifies one or more routers that advertise routes via BGP. The attribute is transitive across Autonomous System boundaries.

The Route Origin Community attribute is used to prevent routing loops when BGP speakers are multi-homed to another site and that site uses the AS-Override feature. This Route Origin Community attribute identifies the site from where the routes are originated so that they are not re-distributed back to the originating site.

The value of the high-order octet of the Type field for the Route Origin Community can be 0x00, 0x01, or 0x02. The value of the low-order octet (Sub-type) of the Type field for this community is 0x03 (if represented in Two-octet AS specific format) and 0x103 (if represented in IPv4 address specific format).

VPNv4/VRF Route Distribution via MP-BGP

Some administrators may choose to use BGP to redistribute VPN routes. Each VRF has its own independent address space; meaning that the same address/net mask can be used in any number of VRFs, where in each VRF the address, in fact, identifies a different system. But a BGP speaker can only install and distribute one route for a specific address prefix. If multiple overlapping routes are received by BGP, only the last received route is installed in any particular per-site VRF route table. Dell Networking allows BGP to install and distribute multiple overlapping routes to a single IP address prefix in different VRFs. This is achieved by the use of a new VPNv4 address family as discussed below. It is recommended that the administrator use a policy to determine which sites can advertise and install routes.

VPNv4 Address Family

MP-BGP allows BGP to carry routes from different address families. To allow BGP to install and distribute overlapping address routes, each address/route must be made unique. To achieve this, a new VPNv4 address family is introduced. A VPN-IPv4 address is a 12-byte quantity, beginning with an 8-byte Route Distinguisher (RD) followed by a 4-byte IPv4 address. The RD attribute follows the same structuring mechanism as described in the 'Extended Community structure' above.

If two VRFs use the same IPv4 address prefix, the router translates these into unique VPN-IPv4 address prefixes by prepending the RD (configured per VRF) to the address. The purpose of the RD is to allow the router to install unique routes with an identical IPv4 address prefix. The structuring of the RD provides no semantics. When BGP compares two such addresses, it ignores the RD structure completely and compares it as a 12-byte entity. It is recommended that each VPN within a site utilize a unique RD.

A router may be configured to associate routes that belong to a particular VRF with a particular RD. When BGP redistributes these routes, BGP prepends the configured RD value to the route and re-distributes them as VPNv4 routes. The router that receives these VPNv4 routes installs them into the global BGP table along with the RD. If two routes have the same address prefix but different RD values, only the last route is installed to the RTO table of the router that imports the route and the rest are overwritten.

Dell Networking BGP doesn't advertise routes in the traditional IPv4 NLRI format when a neighbor is activated in VPNv4 address family mode.

Controlling Route Distribution

This section describes the methods by which VPNv4 route redistribution may be controlled.

The Route Target Attribute (RT)

A Route Target attribute identifies a set of VRFs belonging to a VPN. Every VRF is associated with one or more "Route Target" attributes that define the VPNs to which it belongs. Route targets are advertised using the VPNV4 address family.

When a VPNv4 route is advertised by a router, the "Route Target" attributes are carried in the BGP advertisement as attributes of the route.

An MP-BGP router that receives a VPNV4 route compares it with the Import Route Target attributes configured for one or multiple VRFs and depending on the match installs the route into the matching VRF table. When a BGP router advertises a route to a BGP neighbor, it attaches one or more Export Route Target attributes to the route (as configured for that VRF).

The Export Route Target attributes and the Import Route Target attributes are distinct sets and may or may not be the same in a VRF.

A BGP route can only have one RD but can have multiple Route Targets.

A VRF may be configured to associate all the routes that belong to the VRF with a particular Route Target attribute. Dell Networking allows a finer selection of routes with the use of Export and Import maps. Export and Import maps provides greater flexibility to the administrator where she can associate some routes of a VRF with a particular Route Target attribute and some other routes with a different Route Target attribute.

The Route Target attribute assists in configuring selective route leaking among VRFs using Multi-protocol BGP. Essentially, route leaking between VRFs within a single router can be achieved with just the import and export Route Target statements.

Dell Networking allows configuring Route Target attributes in VRF mode, using IP Extended community lists in association with inbound/outbound Route maps.

Behavior of VPNv4 Route Leaking into a VRF with Identical Prefixes & RTs

In this scenario, multiple routes with identical prefix/RTs but different RDs are advertised to a VPNv4 peer from a single router. The standard decision process rules handle the cases when receiving identical prefix/RTs with the same RD or with different RDs from different routers. On the receiving VPNv4 router, the decision process picks the best route from the identical prefix/neighbor combination routes as follows:

1. If the next hops of the prefixes are the same, the first received MP-NLRI/route is imported into the VRFs with a matching import statement. In the case of an update, the first received MP-NLRI/route is maintained and is not replaced, even if the next hop is different.
2. If the next hops of the prefixes are different, the prefix with the highest RD value is imported into the VRF.

The decision to use the highest RD is arbitrary; however, the scenario where routes with identical prefix/RT pairs are configured in different VRFs on a single router is almost certainly a misconfiguration.

How the VPNv4 NLRI Is Carried in BGP

The BGP Multiprotocol Extensions are used to encode the NLRI. If the Address Family Identifier (AFI) field is set to 1, and the Subsequent Address Family Identifier (SAFI) field is set to 128, the NLRI is a VPNv4 address. AFI 1 is used since the network layer protocol associated with the NLRI is still IP.

In order for two BGP speakers to exchange labeled VPN-IPv4 NLRI, they must use the BGP Capabilities Advertisement (in the OPEN message) to ensure that they both are capable of properly processing VPN-IPv4 NLRI. This is done by using capability code 1 (multiprotocol BGP), with an AFI of 1 and an SAFI of 128.

The VPNv4 NLRI is encoded as specified in the above sections, where the prefix consists of an 8-byte RD followed by an IPv4 prefix.

The Site of Origin Attribute (SOO)

A VPNv4 route may optionally carry a Site of Origin attribute that uniquely identifies a site (a topologically associated set of routers with mutual IP connectivity). This attribute identifies the corresponding route as having come from one of the site members.

The SOO attribute is used in the identification and prevention of routing loops. The SOO attribute is an extended community attribute used to identify routes that have originated from a site so that the re-advertisement of that prefix back to the source site can be prevented, thereby helping to eliminate routing loops.

Site of Origin is one of the attributes a router assigns to a prefix prior to redistributing any VPNv4 prefixes. All prefixes learned from a particular site must be assigned the same site of origin attribute, even if the site is multiply connected to a single router external to the site or is connected to multiple routers external to the site.

Dell Networking allows configuration of the SOO attribute using IP Extended community lists in association with inbound/outbound route maps.

IPv6

Dell Networking supports both IPv4 and IPv6 peering sessions. IPv4 routes are advertised on IPv4 peer sessions. Dell Networking does not support advertisement of IPv4 routes over IPv6 peer sessions. IPv6 routes can be advertised over either type of peer session as described in RFC 4760 and RFC 2545. The user must explicitly activate IPv6 route advertisement on either type of peer session. When IPv6 is enabled, the OPEN message includes the IPv6/unicast AFI/SAFI pair in the multiprotocol capability option.

IPv6 prefixes can be originated through route redistribution or a **network** command. Both can be configured with a route map to set path attributes. BGP can also originate an IPv6 default route. Default-origination can be neighbor-specific. IPv6 routes can be filtered using prefix lists, route maps with community lists, and using AS path access lists. BGP can compute IPv6 routes with up to 16 ECMP next hops.

IPv6 Peering Using A Link Local Address

Link local addresses are one class of IPv6 address that can be used as a BGP peer address. Allowing link local addresses to be used as peer addresses introduces some complications. These are discussed here:

First, consider whether it even makes sense to use a link local address as a peer address. As its name implies, a link local address has link scope; it is only valid on a single link. This characteristic implies that a BGP peer identified by a link local address must be attached to a local link (i.e., shares a layer-2 broadcast domain with the router where the peer is configured). This restriction is typically met for external peers (with the exception of external peers configured with ebgp-multihop). Internal peers are typically not on a local link. Even when two BGP speakers share a common link, a loopback address is often assigned as the peer address to avoid tying the fate of the adjacency to a single link or interface. However, it is possible to configure internal peers using an IPv6 address configured on the link that connects them.

Another feature of link local addresses that makes them less than ideal for BGP peering is that they are auto-generated, typically from the local MAC address. Because each BGP peer is configured with a specific IP address, the peer address must be known at configuration time. This is in contrast to OSPFv3 adjacencies, where neighbor addresses are learned dynamically. If the peer's MAC address changes (for example, if a router fails and is replaced with new hardware), the link local address may change and the BGP configuration will need to be updated. Dell Networking allows discovery of BGP peers using link-local addresses through the configuration of listen ranges.

To use link local addresses for eBGP peers, an administrator must use the next-hop-self option. Normally, when a BGP speaker forwards an external route to internal peers, it retains the BGP NEXT_HOP. If the NEXT_HOP is a link local address, internal peers will be unable to resolve it and thus not be able to use these routes. The router must therefore be configured to change the NEXT_HOP to one of its own global addresses before forwarding routes from an external peer with a link local address (or the implementation must do this automatically).

A primary consideration in using link-local addresses is the user interface. With IPv4 addresses and global IPv6 addresses, the user interface simply identifies the neighbor by IP address:

```
router bgp 1
     neighbor 10.1.1.1 remote-as 100
     neighbor 10.1.1.1 advertisement-interval 10
     neighbor 2001:db8::1 remote-as 200
     neighbor 2001:db8::1 advertisement-interval 20
```

Because two neighbors on different links may have the same link local address, the address itself may not uniquely identify a neighbor, nor does the link local address identify the interface where control packets should be sent to the neighbor. The interface must also be specified.

Dell Networking uses the same MAC address for all routing interfaces on a router and, thus, has the same link local address on each interface. If a user were to try to configure two parallel adjacencies between two Dell Networking routers and wanted to use link local peering, the peers would have the same link local address.

Dell Networking BGP uses the **interface** parameter to the neighbor command to identify a unique auto-configured link local address as shown in the following examples:

```
router bgp 1
     neighbor fe80::1 interface vlan 10 remote-as 100
     neighbor fe80::1 interface vlan 10 advertisement-interval 10
     neighbor fe80::1 interface vlan 20 remote-as 200
```

BGP Dynamic Neighbor Peering

This capability allows peering with BGP neighbors to be dynamically established when BGP connection requests are received from a configured IP address range. Dynamic neighbor peering avoids the need for explicit configuration of the neighbor by the administrator when it is acceptable to establish peering with a neighbor irrespective of their specific IP address. This capability is especially useful when deploying BGP in data centers where configuration of the specific neighbors is time-consuming and error-prone, and where security concerns are lessened due to the closed nature of the network.

Configuration includes the address range on which to listen and, optionally, a peer template from which the neighbor's properties may be inherited.

Because Dell Networking routing is configured on routed VLANs, it is required that dynamic neighbor peering never be configured on a multi-access VLAN. Only a single interface may be a member of a VLAN on which dynamic neighbor peering is configured. Dynamic neighbor configuration on multi-access VLANs is not supported and the behavior is undefined.

The number of configurable listen address ranges in the system is limited to 10. The number of dynamic neighbors sessions established as a result of this feature is limited to the total number of neighbor sessions supported in the system. Listen ranges may be configured for both IPv4 and IPv6 addresses.

The following example configures an interface with a BGP listen range:

1 Create VLAN 200 and set the switch hostname:

```
console#configure
console(config)#vlan 200
console(config-vlan200)#exit
console(config)#hostname "R3"
R3(config)#no ip domain-lookup
```

2 Enable IPv4 routing.

```
R3(config)#ip routing
```

3 Configure a loopback for the local router.

```
R3(config)#interface loopback 0
R3(config-if-loopback0)#ip address 11.11.11.11 255.255.255.255
R3(config-if-loopback0)#exit
```

4 Configure a routed interface to the peer.

```
R3(config)#interface vlan 200
R3(config-if-vlan200)#ip address 192.168.100.11 255.255.255.0
R3(config-if-vlan200)#exit
```

5 Assign the routed interface to a physical interface.

```
R3(config)#interface Gi1/0/16
R3(config-if-Gi1/0/16)#switchport access vlan 200
R3(config-if-Gi1/0/16)#exit
```

6 Configure the local BGP speaker.

```
R3(config)#router bgp 5500
R3(config-router)#bgp log-neighbor-changes
```

7 The router ID is required.

```
R3(config-router)#bgp router-id 11.11.11.11
```

8 Set the listen range to the local routed interface subnet and use template T1.

```
R3(config-router)#bgp listen range 192.168.100.0/24 inherit
peer T1
```

9 Configure template T1 to indicate an IGP peer.

```
R3(config-router)#template peer T1
R3(config-router-tmp)#remote-as 5500
R3(config-router-tmp)#exit
R3(config-router)#exit
R3(config)#exit
```

10 Display the dynamic neighbors.

```
R3#show ip bgp listen range

Listen Range .................................. 192.168.100.0/24
Inherited Template ........................... T1

  Member              ASN       State
  ---------------- ----- -----------
  192.168.100.10    5500 ESTABLISHED
```

IPv6 Source Address Selection

When BGP initiates a TCP connection to a peer, it selects a source IPv6 address. When the user has configured a source interface (using neighbor update-source), the source address is taken from this interface. When the peer's IPv6 address is a link local address, the local interface used to reach the peer is configured (in neighbor remote-as) and the source address is taken rom this interface.

If the neighbor address is a link local address, BGP selects a link local address as the source address. Otherwise, BGP selects a local address in the same subnet as the neighbor's address. If no such address is found, BGP selects the first active global IPv6 address on the source interface.

Network Address of Next Hop

When advertising IPv6 routes, the *Network Address of Next Hop* field in MP_REACH_NLRI is set according to RFC 2545. Under conditions specified in this RFC, both a global and a link local next-hop address may be included. The primary purpose of the global address is an address that can be re-advertised to internal peers. The primary purpose of the link local address is for use as the next hop of routes.

We expect interfaces to external peers will normally have both a link local and a global IPv6 address. Both addresses are included in the *Network Address of Next Hop* field when sending MP_REACH_NLRI to the peer. Normally, internal peers are not on a common subnet (even when they are, peering is normally to addresses on loopback interfaces) and the *Network Address of Next Hop* field includes only a global address. Even when the peer address of an internal peer is on a local link, Dell Networking BGP only advertises a global next-hop IPv6 address.

Identifying Local IPv6 Addresses

In some situations, a router sets the next-hop addresses to addresses configured on one of its own interfaces. When local IPv6 addresses are needed, Dell Networking BGP uses IPv6 addresses on the local end of the TCP connection to the peer. If the peering session uses IPv4, BGP finds IPv6 addresses on the same routing interface as the IPv4 address that terminates the TCP connection. If there are multiple global addresses on the interface, BGP uses the first one (essentially, a random choice). If the set of IPv6 addresses on the interface changes, BGP may alter its choice of local IPv6 addresses.

Using Policy to Specify Next Hop

The network administrator can override the normal rules for selecting a next hop address by configuring IPv6 next hops with outbound policy (a neighbor-specific route map with a **set ipv6 next-hop** term). When configuring an IPv6 next hop, the network administrator should ensure the neighbor can reach the next-hop address. For example, a link local next-hop address should not be configured to an internal peer not on a local link. Using per-neighbor outbound policy to set the IPv6 next hop has the disadvantage of putting each neighbor in a different outbound update group, thus losing the efficiency advantages of sharing an Adj-RIB-Out and of building an UPDATE message once and sending it to many peers.

Alternatively, the network administrator can configure inbound policy on the receiver to set IPv6 next hops.

BGP Limitations

Dell Networking BGP does not support configuration via the Web interface. Dell Networking supports the following RFCs with the exceptions listed in Table 39-4:

Table 39-4. BGP Limitations

Description	Source	Compliance
A BGP speaker MUST be able to support the disabling advertisement of third party NEXT_HOP attributes in order to handle imperfectly bridged media.	RFC 4271 section 5.1.3	No configuration option is available
The parameter MinASOriginationIntervalTimer determines the minimum amount of time that must elapse between successive advertisements of UPDATE messages that report changes within the advertising BGP speaker's own autonomous systems.	RFC 4271 section 9.2.1.2	No. Dell Networking does not enforce a MinASOriginationInte rval
BGP can be configured to remove or replace private ASNs from the AS_PATH attribute of paths advertised to external peers	Dell Networking requirement	Yes
The option to remove or replace private ASNs can be configured independently for each address family	Dell Networking requirement	Yes
ASNs in the range from 64512 to 65535 are removed when this option is enabled	Dell Networking requirement	Yes

Table 39-4. BGP Limitations (Continued)

Description	Source	Compliance
Dell Networking BGP can only be configured through the CLI. SNMP support is limited to the standard MIB, which primarily provides status reporting, and a proprietary MIB which provides additional status variables. Configuration through SNMP is not supported.	Dell Networking requirement	–
BGP may learn the maximum number of routes supported by each Dell Networking N-Series switch: • N3000: 8160 (IPv4), 4096 (IPv6) • N4000: 8160 (IPv4), 4096 (IPv6)	Dell Networking requirement	–

BGP Configuration Examples

This section includes the following configuration examples:

- Enabling BGP
- BGP Example
- Network Example
- BGP Redistribution of OSPF Example
- Configuring the Multi-Exit Discriminator in BGP Advertised Routes
- Configuring Communities in BGP
- Configuring a Route Reflector
- Campus Network MP-BGP and OSPF Configuration
- Configuring MP-eBGP and Extended Communities

Enabling BGP

The following are rules to remember when enabling BGP:

- IP routing must be enabled in order to enable BGP:

  ```
  console(config)#router bgp 4545

  IP routing is not enabled. Enable IP routing or IPv6 routing
  before configuring BGP.

  console(config)#ip routing
  console(config)#router bgp 4545
  ```

- The AS number is required when configuring BGP and will place the user into configuration mode for that BGP ASN. The router-id is required as the IPv4 address for BGP to use. The router-id may be the same as a loopback address.

  ```
  console(config-router)#bgp router-id 1.1.1.1
  ```

- BGP ASN is enabled by default when both the AS number and router-id are configured
- The BGP configuration mode **no enable** command is not shown in the running config unless both the AS number and router-id are configured.

BGP is now enabled in this example.

BGP Example

This example configures iBGP between two routers using the same AS and each using their own loopback address as update-source.

Router A Configuration

On a router, a loopback interface is created and assigned an IP address. The router ID is assigned (the same IPv4 address as the loopback interface) and the IPv4 address of the neighbor (Router B IP address) is assigned. Finally, the neighbor's update source is assigned to the local loopback interface. This ensures that the adjacency will remain up, helping to avoid reconvergence events.

```
console(config)#ip routing
console(config)#interface lo1
console(config-if-loopback1)#ip address 1.1.1.1 /31
console(config-if-loopback1)#exit

console(config)#router bgp 65001
console(config-router)#bgp router-id 1.1.1.1
console(config-router)#neighbor 1.1.1.3 remote-as 65001
console(config-router)#neighbor 1.1.1.3 update-source loopback 1
```

Router B Configuration

Router B is the mirror image of the Router A configuration shown above. The steps are the same except the IP address of the local router and the neighbor are reversed from the configuration above.

```
console(config)#ip routing
console(config)#interface lo1
console(config-if-loopback1)#ip address 1.1.1.3 /31
console(config-if-loopback1)#exit

console(config)#router bgp 65001
console(config-router)#bgp router-id 1.1.1.3
console(config-router)#neighbor 1.1.1.1 remote-as 65001
console(config-router)#neighbor 1.1.1.1 update-source loopback 1
```

Network Example

The following configuration uses the network command to inject received iBGP routes into the BGP routing table. The network mask allows sub-netting and super-netting. An alternative to the **network** command is to use the **redistribute** command.

Interface Gi1/0/1 is configured as a member of VLAN 10, VLAN 10 is assigned an IP address, IP routing is enabled, and BGP router 65001 is created with a router ID of 129.168.1.254. A static subnet route 129.168.0.X is created for VLAN 10. An iBGP neighbor 129.168.0.254 is configured and the network 129.168.x.X is configured to super-net the 129.168.0.x advertisement. The neighbor state is configured to follow loopback 0.

```
console#configure
console(config)#vlan 10
console(config-vlan)#exit
console(config)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 10
console(config-if-Gi1/0/1)#exit

console(config)#interface vlan 10
console(config-if-vlan10)#ip address 129.168.10.2 /24
console(config-if-vlan10)#exit

console(config)#int loopback 0
console(config-if-loopback0)#ip address 129.168.1.254 /24
console(config-if-loopback0)#exit

console(config)#ip routing
console(config)#router bgp 65001
console(config-router)#bgp router-id 129.168.1.254
console(config-router)#neighbor 129.168.0.254 remote-as 65001
console(config-router)#neighbor 129.168.0.254 update-source
loopback 0
console(config-router)#network 129.168.0.0 mask 255.255.0.0
console(config-router)#exit
```

BGP Redistribution of OSPF Example

The following configuration uses the redistribute command to inject received eBGP routes into the BGP routing table.

Interface Te1/0/1 is configured in trunk mode with a native VLAN 10 and VLAN 10 is assigned an IP address with a /30 subnet. BGP fast fallover is enabled for VLAN 10.

IP routing is enabled and a default route is configured that points to the neighbor router. BGP router 3434 is created with a router ID of 172.16.64.1. An eBGP neighbor 216.31.219.19 is configured. Private AS numbers are stripped before distribution to this neighbor.

The router is configured to redistribute static and OSPF type 1 & type 2 external routes. Internal and connected routes are not redistributed. An alternative to the **redistribute** command is to use the **network** command.

```
console#configure
console(config)#vlan 10
console(config-vlan)#exit

console(config)#interface te1/0/1
console(config-if-te1/0/1)#switchport mode trunk
console(config-if-te1/0/1)#switchport trunk native vlan 100
console(config-if-te1/0/1)#exit

console(config)#ip routing
console(config)#interface vlan 10
console(config-if-vlan10)#ip address 172.16.65.1 /30
console(config-if-vlan10)#ip bgp fast-external-fallover permit
console(config-if-vlan10)#exit

console(config)#ip route 0.0.0.0 0.0.0.0 172.16.65.2 name 'Default-
Route'
console(config)#router bgp 3434
console(config-router)#bgp router-id 172.16.64.1
console(config-router)#neighbor 216.31.219.19 remote-private-as
1402
console(config-router)#redistribute static
console(config-router)#redistribute ospf match external 1
console(config-router)#redistribute ospf match external 2
console(config-router)#exit
```

Configuring the Multi-Exit Discriminator in BGP Advertised Routes

The following example configures an egress routing policy that sets the metric for matching routes. In the example, VLAN 10 is created, followed by an access list matching directly connected source address 5.5.5.x for which the metric will be injected into the advertised routes.

A route map "Inject-MED" is created. This route map sets the match criteria as ACL MED-Hosts and configures the metric for matching routes to be 100.

Interface Gi1/0/1 is configured as a member of VLAN 10, VLAN 10 is assigned an IP address, IP routing is enabled, and BGP router 65001 is created with a router ID of 129.168.1.254. A static subnet route 129.168.0.X is created for VLAN 10. An iBGP neighbor 129.168.0.254 is configured and the network 129.168.x.X is configured to inject the metric of 100 into routes matching the prefix-list MED-Hosts. All other routes are permitted using the default metric.

```
console#configure
console(config)#vlan 10
console(config-vlan)#exit

console(config)#ip prefix-list MED-Hosts seq 10 permit 5.5.5.0
255.255.255.0
console(config)#ip bgp-community new-format
console(config)#route-map "Inject-MED" permit 100
console(route-map)#match ip address prefix-list MED-Hosts
console(route-map)#set metric 100
console(route-map)#exit

console(config)#route-map "Inject-MED" permit 200
console(route-map)#exit

console(config)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 10
console(config-if-Gi1/0/1)#exit

console(config)#interface vlan 10
console(config-if-vlan10)#ip address 129.168.10.2 /24
console(config-if-vlan10)#exit

console(config)#ip routing
console(config)#ip route 129.168.0.0 255.255.255.0 129.168.10.1
vlan 10
console(config-vlan)#router bgp 65001
console(config-router)#bgp router-id 129.168.1.254
```

```
console(config-router)#neighbor 129.168.0.254 remote-as 65001
console(config-router)#network 129.168.0.0 mask 255.255.0.0 route-
map Inject-MED
console(config0router)#redistribute connected
console(config-router)#exit
```

Configuring Communities in BGP

The following example configures an egress routing policy that sets the community attribute for matching routes. In the example, VLAN 10 is created, followed by an access list Comm-Hosts matching directly connected source address 5.5.5.x for which the community attribute will be injected into the advertised routes.

A route map "Subnet-5-5-5" is created. This route map sets the match criteria as ACL Comm-Hosts and configures the community attribute 65001:300 for the matching routes.

Interface Gi1/0/1 is configured as a member of VLAN 10, VLAN 10 is assigned an IP address, IP routing is enabled, and BGP router 65001 is created with a router ID of 129.168.1.254. An iBGP neighbor 129.168.0.254 with a remote AS of 65001 is configured (this is an iBGP configuration). BGP is configured to send the community attribute and to use the Subnet-5-5-5 match criteria as an egress policy.

```
console#configure
console(config)#vlan 10
console(config-vlan)#exit

console(config)#ip prefix-list MED-Hosts seq 10 permit 5.5.5.0
255.255.255.0
console(config)#ip bgp-community new-format
console(config)#route-map "Subnet-5-5-5" permit 100
console(route-map)#match ip address prefix-list Comm-Hosts
console(route-map)#set community 65001:300 additive
console(route-map)#exit

console(config)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 10
console(config-if-Gi1/0/1)#exit

console(config)#interface vlan 10
console(config-if-vlan10)#ip address 129.168.10.2 /24
console(config-if-vlan10)#exit

console(config)#int loopback 0
```

```
console(config-if-loopback0)#ip address 129.168.1.254 /24
console(config-if-loopback0)#exit

console(config)#ip routing
console(config)#router bgp 65001
console(config-router)#bgp router-id 129.168.1.254
console(config-router)#neighbor 129.168.0.254 remote-as 65001
console(config-router)#neighbor 129.168.0.254 send-community
console(config-router)#neighbor 129.168.0.254 route-map Subnet-5-5-
5 out
console(config-router)#redistribute connected
console(config-router)#exit
```

Configuring a Route Reflector

The following example configures an iBGP speaker as a route reflector. Each iBGP neighbor will have it's routes reflected to other iBGP neighbors. In this example, only a single neighbor is configured.

1 Interface Gi1/0/1 is configured as a member of VLAN 10:

```
console#configure
console(config)#vlan 10
console(config-vlan10)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport access vlan 10
```

2 VLAN 10 is assigned an IP address:

```
console(config-if-Gi1/0/1)#interface vlan 10
console(config-if-vlan10)#ip address 129.168.10.2 /24
```

3 Loopback 0 is assigned an IP address. The iBGP peer will be configured to follow the loopback state in a later step:

```
console(config-if-vlan10)#interface loopback 0
console(config-if-loopback0)#ip address 129.168.1.254 /24
console(config-if-loopback0)#exit
```

4 IP routing is enabled:

```
console(config)#ip routing
```

5 BGP router 65001 is created with a router ID of 129.168.1.254:

```
console(config)#router bgp 65001
console(config-router)#bgp router-id 129.168.1.254
```

6 iBGP neighbor 129.168.0.254 is configured as a neighbor following the loopback 0 state:

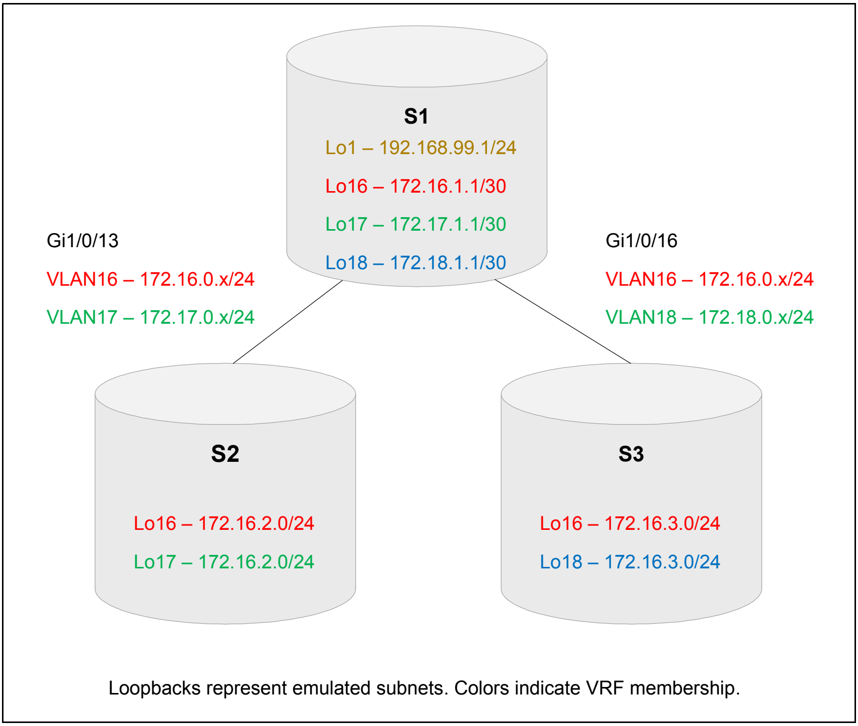
This iBGP neighbor is designated a route reflector client. Other iBGP neighbors can be configured as route reflector clients in order to reduce the explosion of neighbor configuration required to implement a full mesh iBGP network.

```
console(config-router)#neighbor 129.168.0.254 remote-as 65001
console(config-router)#neighbor 129.168.0.254 update-source
loopback 0
console(config-router)#neighbor 129.168.0.254 route-reflector-
client
console(config-router)#exit
```

Campus Network MP-BGP and OSPF Configuration

Consider the topology below, which is a subset of what might be found on a small campus. This network services three customers (Red, Green, and Blue). The internet connection to the outside world is hosted in router S1. Router S2 hosts the Red and Green network. Router S3 hosts the Red and Blue network. A common service is supplied over the 192.168.99.1/24 network.

Figure 39-4. Campus Network MP-BGP and OSPF Configuration



The IPv4 prefixes are assigned as follows:

- Red: 172.16.0.0/16
- Green: 172.17.0.0/16
- Blue: 172.18.0.0/16
- VoIP Services: 192.168.99.0/24

Four VRFs are created on S1. Each VRF is assigned a unique route distinguisher (RD). The RDs utilized here are taken from the private ASN address space. Three of the VRFs are assigned to the Red, Green, and Blue networks and the last VRF is utilized for the common service. We use a loopback on S1 to emulate the common service network instead of a VLAN and physical interface. The VRF configuration on the loopback is identical to the case of a VLAN and physical interface. Interfaces need not be assigned to a VRF before configuring the IP address. IP addresses are migrated when an interface is assigned to a VRF.

VRF Blue will not participate in the common service; however VRFs Red and Green will participate. Route target import and export statements set the appropriate distribution of routes.

1 Create VLANs for the VRFs.

```
console#configure
console(config)#vlan 16-18
console(config-vlan)#exit
```

2 Set the hostname.

```
console(config)#hostname "S1"
```

3 Set a loopback for the common service. This could be replaced with a VLAN and physical interface.

```
S1(config)#interface loopback 1
S1(config-if-loopback1)#ip vrf forwarding Shared
S1(config-if-loopback1)#ip address 192.168.99.1 255.255.255.0
S1(config-if-loopback1)#exit
```

4 Create VRF Blue.

```
S1(config)#ip vrf Blue
S1(config-ip-vrf-Blue)#rd 65000:3
S1(config-ip-vrf-Blue)#exit
```

5 Create VRF Green, import the common service, and export the Green network.

```
S1(config)#ip vrf Green
S1(config-ip-vrf-Green)#rd 65000:2
S1(config-ip-vrf-Green)#route-target export 65000:2
S1(config-ip-vrf-Green)#route-target import 65000:99
S1(config-ip-vrf-Green)#exit
```

6 Create VRF Red, import the common service, and export the Red network.

```
S1(config)#ip vrf Red
S1(config-ip-vrf-Red)#rd 65000:1
S1(config-ip-vrf-Red)#route-target export 65000:1
S1(config-ip-vrf-Red)#route-target import 65000:99
S1(config-ip-vrf-Red)#exit
```

7 Create VRF Shared, import the Red and Green network, and export the common service.

```
S1(config)#ip vrf Shared
S1(config-vrf-Shared)#rd 65000:99
S1(config-vrf-Shared)#route-target import 65000:1
S1(config-vrf-Shared)#route-target import 65000:2
S1(config-vrf-Shared)#route-target export 65000:99
S1(config-vrf-Shared)#exit
S1(config)#no ip domain-lookup
```

8 Enable IPv4 routing.

```
S1(config)#ip routing
```

9 Set a loopback for Red network interface access. This could be replaced with a VLAN and physical interface.

```
S1(config)#interface loopback 16
S1(config-if-loopback16)#ip vrf forwarding Red
S1(config-if-loopback16)#ip address 172.16.1.1 255.255.255.252
S1(config-if-loopback16)#exit
```

10 Set a loopback for Green network interface access. This could be replaced with a VLAN and physical interface.

```
S1(config)#interface loopback 17
S1(config-if-loopback17)#ip vrf forwarding Green
S1(config-if-loopback17)#ip address 172.17.1.1 255.255.255.252
S1(config-if-loopback17)#exit
```

11 Set a loopback for Blue network interface access. This could be replaced with a VLAN and physical interface.

```
S1(config)#interface loopback 18
S1(config-if-loopback18)#ip vrf forwarding Blue
S1(config-if-loopback18)#ip address 172.18.1.1 255.255.255.252
S1(config-if-loopback18)#exit
```

12 Associate the Red VRF with a VLAN routed interface.

```
S1(config)#interface vlan 16
S1(config-if-vlan16)#ip vrf forwarding Red
S1(config-if-vlan16)#ip address 172.16.0.1 255.255.255.0
S1(config-if-vlan16)#exit
```

13 Associate the Green VRF with a VLAN routed interface.

```
S1(config)#interface vlan 17
S1(config-if-vlan17)#ip vrf forwarding Green
S1(config-if-vlan17)#ip address 172.17.0.1 255.255.255.0
S1(config-if-vlan17)#exit
```

14 Associate the Blue VRF with a VLAN routed interface.

```
S1(config)#interface vlan 18
S1(config-if-vlan18)#ip vrf forwarding Blue
S1(config-if-vlan18)#ip address 172.18.0.1 255.255.255.0
S1(config-if-vlan18)#exit
```

We can display the VRF associated RDs and interfaces as follows:

```
S1(config)#show ip vrf

Number of VRFs................... 4

Name                 Identifier      Route Distinguisher
-------------------- --------------- ---------------------
Blue                 1               65000:3
Green                2               65000:2
Red                  3               65000:1
Shared               4               65000:99


S1(config)#show ip vrf interface

VRF Name              Interface    State  IP Address      IP Mask         Method
--------------------  ----------   -----  --------------- --------------- -------
Blue                  Vl18         Up     172.18.0.1      255.255.255.0   Manual
Blue                  Lo18         Up     172.18.1.1      255.255.255.252 Manual
Green                 Vl17         Up     172.17.0.1      255.255.255.0   Manual
Green                 Lo17         Up     172.17.1.1      255.255.255.252 Manual
Red                   Vl16         Up     172.16.0.1      255.255.255.0   Manual
Red                   Lo16         Up     172.16.1.1      255.255.255.252 Manual
Shared                Lo1          Up     192.168.99.1    255.255.255.0   Manual
```

Next, configure OSPF to exchange routes with the other routers. OSPF runs in the VRFs and area 0 is used within each VRF. Each VRF is configured to redistribute BGP subnets advertised by S1.

1 Configure router Blue.

```
S1(config)#router ospf vrf "Blue"
```

2 A router ID is required.

```
S1(config-router-vrf-Blue)#router-id 172.18.0.1
```

3 Configure network as 'don't care'. A non-zero IP address is required.

```
S1(config-router-vrf-Blue)#network 172.18.0.0 255.255.255.255
area 0
```

4 Redistribute BGP subnets.

```
S1(config-router-vrf-Blue)#redistribute bgp subnets
S1(config-router-vrf-Blue)#exit
```

5 VRF Green and Red are nearly identical to VRF Blue.

```
S1(config)#router ospf vrf "Green"
S1(config-router-vrf-Green)#router-id 172.17.0.1
S1(config-router-vrf-Green)#network 172.17.0.0 255.255.255.255
area 0
S1(config-router-vrf-Green)#redistribute bgp subnets
S1(config-router-vrf-Green)#exit

S1(config)#router ospf vrf "Red"
S1(config-router-vrf-Red)#router-id 172.16.0.1
S1(config-router-vrf-Red)#network 172.16.0.0 255.255.255.255
area 0
S1(config-router-vrf-Red)#redistribute bgp subnets
S1(config-router-vrf-Red)#exit
```

Next, assign the VRF associated VLANs to the interfaces connected to the rest of the Red, Green, and Blue networks:

1 Configure the S1-S2 trunk.

```
S1(config)#interface Gi1/0/13
S1(config-if-Gi1/0/13)#switchport mode trunk
S1(config-if-Gi1/0/13)#switchport trunk allowed vlan 1,16-17
S1(config-if-Gi1/0/13)#exit
```

2 Configure the S1-S3 trunk.

```
S1(config)#interface Gi1/0/16
S1(config-if-Gi1/0/16)#switchport mode trunk
S1(config-if-Gi1/0/16)#switchport trunk allowed vlan 1,16-18
S1(config-if-Gi1/0/16)#exit
```

Routers S2 and S3 require VRFs to be created and OSPF to be configured. The RDs must match the RDs configured for the VRFs on S1.

1 Configure S2.

```
console#configure
console(config)#vlan 16-17
console(config-vlan)#exit
```

2 Set the hostname.

```
console(config)#hostname "S2"
```

3 Create VRF Green. Same RD as on S1.

```
S2(config)#ip vrf Green
S2(config-vrf-Green)#rd 65000:2
S2(config-vrf-Green)#exit
```

4 Create VRF Red. Same RD as on S1.

```
S2(config)#ip vrf Red
S2(config-vrf-Red)#rd 65000:1
S2(config-vrf-Red)#exit
S2(config)#no ip domain-lookup
```

5 Enable IP routing.

```
S2(config)#ip routing
```

6 Emulate a network in the Red VRF. The loopback subnet can be replaced with a VLAN-routed interface.

```
S2(config)#interface loopback 16
S2(config-if-loopback16)#ip vrf forwarding Red
S2(config-if-loopback16)#ip address 172.16.2.1 255.255.255.0
S2(config-if-loopback16)#exit
```

7 Emulate a network in the Green VRF. The loopback network can be replaced with a VLAN-routed interface.

```
S2(config)#interface loopback 17
S2(config-if-loopback17)#ip vrf forwarding Green
S2(config-if-loopback17)#ip address 172.17.2.1 255.255.255.0
S2(config-if-loopback17)#exit
```

8 Create a VLAN routed interface to router S1 for VRF Red.

```
S2(config)#interface vlan 16
S2(config-if-vlan16)#ip vrf forwarding Red
S2(config-if-vlan16)#ip address 172.16.0.2 255.255.255.0
S2(config-if-vlan16)#exit
```

9 Create a VLAN routed interface to router S1 for VRF Green.

```
S2(config)#interface vlan 17
S2(config-if-vlan17)#ip vrf forwarding Green
S2(config-if-vlan17)#ip address 172.17.0.2 255.255.255.0
S2(config-if-vlan17)#exit
```

10 Assign VLANs to a trunk.

```
S2(config)#interface Gi1/0/13
S2(config-if-Gi1/0/13)#switchport mode trunk
S2(config-if-Gi1/0/13)#switchport trunk allowed vlan 1,16-17
S2(config-if-Gi1/0/13)#exit
```

S3 is configured almost identically to S2 with the exception that VRF Green is not configured on S2.

1 Create VLANs and configure the hostname for S3.

```
console#configure
console(config)#vlan 16,18
console(config-vlan)#exit
console(config)#hostname "S3"
```

2 Configure VRF Blue. Same RD as on S1 and S2.

```
S3(config)#ip vrf Blue
S3(config-vrf-Blue)#rd 65000:3
S3(config-vrf-Blue)#exit
```

3 Configure VRF Red. Same RD as on S1 and S2.

```
S3(config)#ip vrf Red
S3(config-vrf-Red)#rd 65000:1
S3(config-vrf-Red)#exit
S3(config)#no ip domain-lookup
```

4 Enable routing.

```
S3(config)#ip routing
```

5 Emulate the Red network using a loopback.

```
S3(config)#interface loopback 16
S3(config-if-loopback16)#ip vrf forwarding Red
S3(config-if-loopback16)#ip address 172.16.3.1 255.255.255.0
S3(config-if-loopback16)#exit
```

6 Emulate the Blue network using a loopback.

```
S3(config)#interface loopback 18
S3(config-if-loopback18)#ip vrf forwarding Blue
S3(config-if-loopback18)#ip address 172.18.3.1 255.255.255.0
S3(config-if-loopback18)#exit
```

7 Assign VLANs to the VRFs.

```
S3(config)#interface vlan 16
S3(config-if-vlan16)#ip vrf forwarding Red
S3(config-if-vlan16)#ip address 172.16.0.3 255.255.255.0
S3(config-if-vlan16)#exit

S3(config)#interface vlan 18
S3(config-if-vlan18)#ip vrf forwarding Blue
S3(config-if-vlan18)#ip address 172.18.0.3 255.255.255.0
S3(config-if-vlan18)#exit
```

8 Assign the VLANs to a physical interface.

```
S3(config)#interface Gi1/0/16
S3(config-if-Gi1/0/16)#switchport mode trunk
S3(config-if-Gi1/0/16)#switchport trunk allowed vlan 1,16,18
S3(config-if-Gi1/0/16)#exit
```

This is a very simple OSPF configuration for each of the routers. In this case, a loopback is used to emulate an OSPF connected interface. If an actual VLAN-routed interface is used, declare it a passive interface in the OSPF configuration.

For router S2, VRF Green and Red are configured.

1 Create an OSPF instance for VRF Green

```
S2(config)#router ospf vrf "Green"
```

2 Router ID is required.

```
S2(config-router-vrf-Green)#router-id 172.17.0.99
```

3 Network is all 'don't care'.

```
S2(config-router-vrf-Green)#network 172.17.0.0 255.255.255.255
area 0
```

4 Redistribute connected routes to S1.

```
S2(config-router-vrf-Green)#redistribute connected
S2(config-router-vrf-Green)#exit
```

5 Create an OSPF instance for VRF Red.

```
S2(config)#router ospf vrf "Red"
```

6 Router ID is required.

```
S2(config-router-vrf-Red)#router-id 172.16.0.99
```

7 Network is all 'don't care'.

```
S2(config-router-vrf-Red)#network 172.16.0.0 255.255.255.255
area 0
```

8 Redistribute connected routes.

```
S2(config-router-vrf-Red)#redistribute connected
S2(config-router-vrf-Red)#exit
```

9 Allow both VRF Red and Green access to router S1 over a physical interface.

```
S2(config)#interface Gi1/0/13
S2(config-if-Gi1/0/16)#switchport mode trunk
S2(config-if-Gi1/0/16)#switchport trunk allowed vlan 1,16-17
S2(config-if-Gi1/0/16)#exit
```

OSPF on S3 is configured similarly to S2 with VRF Red and Blue:

1 Create OSPF sessions in each VRF. Assign area 0. Router ID assignment is required.

```
S3(config)#router ospf vrf "Blue"
S3(config-router-vrf-Blue)#router-id 172.18.0.99
S3(config-router-vrf-Blue)#network 172.18.0.0 255.255.255.255
area 0
S3(config-router-vrf-Blue)#exit

S3(config)#router ospf vrf "Red"
S3(config-router-vrf-Red)#router-id 172.16.0.98
S3(config-router-vrf-Red)#network 172.16.0.0 255.255.255.255
area 0
S3(config-router-vrf-Red)#exit
```

2 Assign the VLANs to the physical interface connected to S1.

```
S3(config)#interface Gi1/0/16
S3(config-if-Gi1/0/16)#switchport mode trunk
S3(config-if-Gi1/0/16)#switchport trunk allowed vlan 1,16,18
S3(config-if-Gi1/0/16)#exit
```

Examine the OSPF adjacencies on router S1. Every router is connected to every other router.

```
S1#show ip ospf neighbor vrf Red

   Router ID     Priority IP Address      Interface   State            Dead
                                                                       Time
---------------- -------- --------------- ----------- ---------------- ------
172.16.0.99      1        172.16.0.2      Vl16        Full/BACKUP-DR   37
172.16.0.98      1        172.16.0.3      Vl16        Full/BACKUP-DR   33

S1#show ip ospf neighbor vrf Green

   Router ID     Priority IP Address      Interface   State            Dead
                                                                       Time
---------------- -------- --------------- ----------- ---------------- ------
172.17.0.99      1        172.17.0.2      Vl17        Full/BACKUP-DR   33


S1#show ip ospf neighbor vrf Blue

   Router ID     Priority IP Address      Interface   State            Dead
                                                                       Time
---------------- -------- --------------- ----------- ---------------- ------
172.18.0.99      1        172.18.0.3      Vl18        Full/BACKUP-DR   35
```

The VRFs should all have full connectivity.

```
S1#show ip route vrf Red

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route


 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *172.16.0.0/24 [0/0] directly connected,   Vl16
C      *172.16.1.0/30 [0/0] directly connected,   Lo16
O      *172.16.2.0/24 [110/11] via 172.16.0.2,   Vl16
O      *172.16.3.0/24 [110/11] via 172.16.0.3,   Vl16

S1#show ip route vrf Green

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *172.17.0.0/24 [0/0] directly connected,   Vl17
C      *172.17.1.0/30 [0/0] directly connected,   Lo17
O      *172.17.2.0/24 [110/11] via 172.17.0.2,   Vl17


S1#show ip route vrf Blue

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *172.18.0.0/24 [0/0] directly connected,   Vl18
C      *172.18.1.0/30 [0/0] directly connected,   Lo18
O      *172.18.3.0/24 [110/11] via 172.18.0.3,   Vl18
```

To provision MPBGP to distribute routes for the shared service, on S1 configure a loopback to emulate the common service network:

1 Set a loopback for the BGP router.

```
S1(config)#interface loopback 0
S1(config-if-loopback0)#ip address 192.0.2.1 255.255.255.255
S1(config-if-loopback0)#exit
```

Next, configure a BGP router and allow route redistribution to occur. Configuration of the router ID is required.

2 Configure a BGP router.

```
S1(config)#router bgp 65000
S1(config-router)#bgp log-neighbor-changes
S1(config-router)#bgp router-id 192.0.2.1
```

3 Add the Blue VRF address family and allow redistribution of OSPF and connected origin routes.

```
S1(config-router)#address-family ipv4 vrf Blue
S1(config-router-af)#redistribute connected
S1(config-router-af)#redistribute ospf
S1(config-router-af)#exit
```

4 Add the Green VRF address family and allow redistribution of OSPF and connected origin routes.

```
S1(config-router)#address-family ipv4 vrf Green
S1(config-router-af)#redistribute connected
S1(config-router-af)#redistribute ospf
S1(config-router)#exit
```

5 Add the Red VRF address family and allow redistribution of OSPF and connected origin routes.

```
S1(config-router)#address-family ipv4 vrf Red
S1(config-router-af)#redistribute connected
S1(config-router-af)#redistribute ospf
S1(config-router-af)#exit
```

6 Add the Shared VRF and allow redistribution of OSPF and connected origin routes.

```
S1(config-router)#address-family ipv4 vrf Shared
S1(config-router-af)#redistribute connected
S1(config-router-af)#redistribute ospf
S1(config-router-af)#exit
S1(config-router)#exit
```

Verify that BGP maintains routes for each of the VRFs. The common service VRF "Shared" is exported via the route-target 65000:99 and imported into the Red and Green VRFs.

```
S1(config-router)#show ip bgp vpnv4 all

BGP table version is 0, local router ID is 192.0.2.1
Status codes: s suppressed, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

    Network             Next Hop         Metric     LocPref Path
    ------------------- ---------------- ---------- ------- ----
    Route Distinguisher : 65000:3  for VRF Blue
*>i 172.18.0.0/24       ::
                        0.0.0.0                         100  ?
*>i 172.18.1.0/30       ::
                        0.0.0.0                         100  ?
*>i 172.18.3.0/24       172.18.0.3                      100  ?
    Route Distinguisher : 65000:2  for VRF Green
*>i 172.17.0.0/24       ::
                        0.0.0.0                         100  ?
*>i 172.17.1.0/30       ::
                        0.0.0.0                         100  ?
*>i 172.17.2.0/24       172.17.0.2                      100  ?
*>i 192.168.99.0/24     ::
                        0.0.0.0                         100  ?
    Route Distinguisher : 65000:1  for VRF Red
*>i 172.16.0.0/24       ::
                        0.0.0.0                         100  ?
*>i 172.16.1.0/30       ::
                        0.0.0.0                         100  ?
*>i 172.16.2.0/24       172.16.0.2                      100  ?
*>i 172.16.3.0/24       172.16.0.3                      100  ?
*>i 192.168.99.0/24     ::
                        0.0.0.0                         100  ?
    Route Distinguisher : 65000:99  for VRF Shared
*>i 172.16.0.0/24       ::
                        0.0.0.0                         100  ?
*>i 172.17.0.0/24       ::
                        0.0.0.0                         100  ?
*>i 172.16.1.0/30       ::
                        0.0.0.0                         100  ?
*>i 172.17.1.0/30       ::
                        0.0.0.0                         100  ?
*>i 172.16.2.0/24       172.16.0.2                      100  ?
*>i 172.17.2.0/24       172.17.0.2                      100  ?
*>i 172.16.3.0/24       172.16.0.3                      100  ?
*>i 192.168.99.0/24     ::
                        0.0.0.0                         100  ?
```

The best routes are placed into the route table in each of the VRFs. VRF Blue does not import or export any routes and does not have access to the common services.

```
S1#show ip route vrf Shared

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
L      *172.16.0.0/24 [200/0] via 0.0.0.0,   Vl16
L      *172.16.1.0/30 [200/0] via 0.0.0.0,   Lo16
L      *172.16.2.0/24 [200/0] via 172.16.0.2,   Vl16
L      *172.16.3.0/24 [200/0] via 172.16.0.3,   Vl16
L      *172.17.0.0/24 [200/0] via 0.0.0.0,   Vl17
L      *172.17.1.0/30 [200/0] via 0.0.0.0,   Lo17
L      *172.17.2.0/24 [200/0] via 172.17.0.2,   Vl17
C      *192.168.99.0/24 [0/0] directly connected,   Lo1


S1#show ip route vrf Red

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *172.16.0.0/24 [0/0] directly connected,   Vl16
C      *172.16.1.0/30 [0/0] directly connected,   Lo16
O      *172.16.2.0/24 [110/11] via 172.16.0.2,   Vl16
O      *172.16.3.0/24 [110/11] via 172.16.0.3,   Vl16
L      *192.168.99.0/24 [200/0] via 0.0.0.0,   Lo1


S1#show ip route vrf Green

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *172.17.0.0/24 [0/0] directly connected,   Vl17
C      *172.17.1.0/30 [0/0] directly connected,   Lo17
O      *172.17.2.0/24 [110/11] via 172.17.0.2,   Vl17
L      *192.168.99.0/24 [200/0] via 0.0.0.0,   Lo1
```

```
S1#show ip route vrf Blue

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *172.18.0.0/24 [0/0] directly connected,   Vl18
C      *172.18.1.0/30 [0/0] directly connected,   Lo18
O      *172.18.3.0/24 [110/11] via 172.18.0.3,   Vl18
```

The routes are propagated via OSPF to the S2 and S3 routers.

```
S2#show ip route vrf Red

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *172.16.0.0/24 [0/0] directly connected,   Vl16
O      *172.16.1.0/30 [110/11] via 172.16.0.1,   Vl16
C      *172.16.2.0/24 [0/0] directly connected,   Lo16
O      *172.16.3.0/24 [110/11] via 172.16.0.3,   Vl16
O E2   *192.168.99.0/24 [110/1] via 172.16.0.1,   Vl16


S2#show ip route vrf Green

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *172.17.0.0/24 [0/0] directly connected,   Vl17
O      *172.17.1.0/30 [110/11] via 172.17.0.1,   Vl17
C      *172.17.2.0/24 [0/0] directly connected,   Lo17
O E2   *192.168.99.0/24 [110/1] via 172.17.0.1,   Vl17


S3#show ip route vrf Red

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route
```

```
 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *172.16.0.0/24 [0/0] directly connected,   Vl16
O      *172.16.1.0/30 [110/11] via 172.16.0.1,   Vl16
O      *172.16.2.0/24 [110/11] via 172.16.0.2,   Vl16
C      *172.16.3.0/24 [0/0] directly connected,   Lo16
O E2   *192.168.99.0/24 [110/1] via 172.16.0.1,   Vl16


S3#show ip route vrf Blue

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *172.18.0.0/24 [0/0] directly connected,   Vl18
O      *172.18.1.0/30 [110/11] via 172.18.0.1,   Vl18
C      *172.18.3.0/24 [0/0] directly connected,   Lo18
```

Configuring MP-eBGP and Extended Communities

In this configuration, router R1 is connected to router R2 (via VLAN 100 on Gi1/0/13) and router R3 (via VLAN 200 in Gi1/0/16). Router R1 (AS 5500) and R2 (AS 6500) communicate via MP-eBGP. Router R1 and R3 are both in AS 5500 and for an iBGP relationship. R3's purpose in this configuration is to show that routes received from R2 are redistributed within the IGP and to inject routes into the IGP.

Router R1 Configuration

Router R1 does not have any VRFs and does not send the extended community attribute.

1 Create two VLANs for connection to R2 (VLAN 100) and R3 (VLAN 200).

```
console#configure
console(config)#vlan 100,200
console(config-vlan)#exit
console(config)#hostname "R1"
```

2 Disable domain lookup and enable IP routing.

```
R1(config)#no ip domain-lookup
R1(config)#ip routing
```

3 Create a loopback for the router ID.

```
R1(config)#interface loopback 0
R1(config-if-loopback0)#ip address 10.10.10.10 255.255.255.255
R1(config-if-loopback0)#exit
```

4 Configure the two IP routing VLANs for the connections to R1 and R2.

```
R1(config)#interface vlan 100
R1(config-if-vlan100)#ip address 172.16.10.1 255.255.255.0
R1(config-if-vlan100)#exit

R1(config)#interface vlan 200
R1(config-if-vlan200)#ip address 192.168.100.10 255.255.255.0
R1(config-if-vlan200)#exit
```

5 Assign the R1 physical interface.

```
R1(config)#interface Gi1/0/13
R1(config-if-Gi1/0/13)#switchport access vlan 100
R1(config-if-Gi1/0/13)#exit
```

6 Assign the R2 physical interface.

```
R1(config)#interface Gi1/0/16
```

```
R1(config-if-Gi1/0/16)#switchport access vlan 200
R1(config-if-Gi1/0/16)#exit
```

7 Configure the BGP router.

```
R1(config)#router bgp 5500
R1(config-router)#bgp log-neighbor-changes
```

8 Configure the router ID.

```
R1(config-router)#bgp router-id 10.10.10.10
```

9 This router advertises the 192.168.100.0/24 network.

```
R1(config-router)#network 192.168.100.0 mask 255.255.255.0
```

10 Redistribute connected routes (10.10.10.10/32).

```
R1(config-router)#redistribute connected
```

11 Configure the R2 neighbor.

```
R1(config-router)#neighbor 172.16.10.2 remote-as 6500
```

12 Configure the R1 neighbor.

```
R1(config-router)#neighbor 192.168.100.11 remote-as 5500
R1(config-router)#address-family vpnv4 unicast
R1(config-router-af)#exit
R1(config-router)#exit
```

Router R2 Configuration

Router R2 has a VRF WAN with route distinguisher 2020:1. This attribute is sent in the UPDATE message in the MP_REACH_NLRI path attribute. Router R2 exhibits an MP-BGP capability toward router R1. The administrator for R2 can implement route maps to control distribution of VRF route information to R1 by matching on the extended community attribute.

1 Configure a VLAN for the R1 neighbor.

```
console#configure
console(config)#vlan 100
console(config-vlan100)#exit
console(config)#hostname "R2"
```

2 Create a VRF.

```
R2(config)#ip vrf WAN
R2(config-vrf-WAN)#rd 2020:1
R2(config-vrf-WAN)#route-target export 2020:1
R2(config-vrf-WAN)#route-target import 2020:1
R2(config-vrf-WAN)#exit
```

3 Disable domain lookup and enable IP routing.

```
R2(config)#no ip domain-lookup
R2(config)#ip routing
```

4 Create a loopback for the BGP router.

```
R2(config)#interface loopback 0
R2(config-if-loopback0)#ip address 20.20.20.20 255.255.255.255
R2(config-if-loopback0)#exit
```

5 Create a loopback to emulate a subnet in the VRF. This could be assigned to a real VLAN.

```
R2(config)#interface loopback 1
R2(config-if-loopback1)#ip vrf forwarding WAN
R2(config-if-loopback1)#ip address 30.30.30.30 255.255.255.0
R2(config-if-loopback1)#exit
```

6 VLAN 100 is connected to R1.

```
R2(config)#interface vlan 100
R2(config-if-vlan100)#ip address 172.16.10.2 255.255.255.0
R2(config-if-vlan100)#exit
```

7 Assign the physical connection to R1.

```
R2(config)#interface Gi1/0/13
R2(config-if-Gi1/0/13)#switchport access vlan 100
R2(config-if-Gi1/0/13)#exit
```

8 Configure a BGP router with as 6500.

```
R2(config)#router bgp 6500
R2(config-router)#bgp log-neighbor-changes
```

9 Use the loopback for the router ID. The router ID is required.

```
R2(config-router)#bgp router-id 20.20.20.20
```

10 Redistribute connected subnets.

```
R2(config-router)#redistribute connected
```

11 R1 is an eBGP connection.

```
R2(config-router)#neighbor 172.16.10.1 remote-as 5500
R2(config-router)#neighbor 172.16.10.1 send-community
```

12 Advertise the IPv4 routes in VRF WAN (20.20.20.20/32 and 172.16.10.0/24).

```
R2(config-router)#address-family ipv4 vrf WAN
R2(config-router-af)#neighbor 172.16.10.1 remote-as 5500
R2(config-router-af)#redistribute connected
```

```
R2(config-router-af)#redistribute static
R2(config-router-af)#exit
```

13 Advertise the VPNv4 routes (30.30.30.0/24). These routes are transmitted with the extended community attribute (2020:1).

```
R2(config-router)#address-family vpnv4 unicast
R2(config-router-af)#neighbor 172.16.10.1 send-community both
R2(config-router-af)#neighbor 172.16.10.1 activate
R2(config-router-af)#exit
R2(config-router)#exit
R2(config)#exit
```

Router R3 Configuration

1 Configure a VLAN for connection to R1.

```
console#configure
console(config)#vlan 200
console(config-vlan200)#exit
console(config)#hostname "R3"
```

2 Create a loopback for the BGP router ID.

```
R3(config)#interface loopback 0
R3(config-if-loopback0)#ip address 11.11.11.11 255.255.255.255
R3(config-if-loopback0)#exit
```

3 Disable domain lookup.

```
R3(config)#no ip domain-lookup
```

4 Create a routed VLAN for connection to R1.

```
R3(config)#interface vlan 200
R3(config-if-vlan200)#ip address 192.168.100.11 255.255.255.0
R3(config-if-vlan200)#exit
```

5 Attach a physical interface to the VLAN.

```
R3(config)#interface Gi1/0/16
R3(config-if-Gi1/0/16)#switchport access vlan 200
R3(config-if-Gi1/0/16)#exit
```

6 Create an iBGP router.

```
R3(config)#router bgp 6500
R3(config-router)#bgp log-neighbor-changes
```

7 Setting the router ID is mandatory.

```
R3(config-router)#bgp router-id 11.11.11.11
```

8 Identify the BGP neighbor.

```
R3(config-router)#neighbor 192.168.100.10 remote-as 5500
```

9 Redistribute connected and static routes.

```
R3(config-router)#redistribute connected
R3(config-router)#redistribute static
R3(config-router)#exit
R3(config)#exit
R3#exit
```

Discussion

Verify that the routes on R2 are being distributed to R1 and R3. This shows the R2 BGP and routing tables.

```
R2#show ip route vrf WAN

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *30.30.30.0/24 [0/0] directly connected,   Lo1

R2#show ip bgp vpnv4 all

BGP table version is 24, local router ID is 20.20.20.20
Status codes: s suppressed, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

    Network             Next Hop         Metric     LocPref Path
    ------------------- ---------------- ---------- ------- ----
    Route Distinguisher : 2020:1  for VRF WAN
*>i 30.30.30.0/24       ::
                        0.0.0.0                         100  ?
```

R2 shows routes from R1 and R3 in the IPv4 address family.

```
R2#show ip route

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
B      *10.10.10.10/32 [20/0] via 172.16.10.1,   Vl100
B      *11.11.11.11/32 [20/0] via 172.16.10.1,   Vl100
C      *20.20.20.20/32 [0/0] directly connected,   Lo0
C      *172.16.10.0/24 [0/0] directly connected,   Vl100
```

```
B      *192.168.100.0/24 [20/0] via 172.16.10.1,   Vl100
```

This is the resulting R1 routing table.

```
R1#show ip route

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
C      *10.10.10.10/32 [0/0] directly connected,   Lo0
B      *11.11.11.11/32 [200/0] via 192.168.100.11,   Vl200
B      *20.20.20.20/32 [20/0] via 172.16.10.2,   Vl100
C      *172.16.10.0/24 [0/0] directly connected,   Vl100
B      *30.30.30.30/32 [20/0] via 172.16.10.2,   Vl100

C      *192.168.100.0/24 [0/0] directly connected,   Vl200
```

This is the resulting R3 routing table.

```
R3#show ip route

Route Codes: R - RIP Derived, O - OSPF Derived, C - Connected, K - Kernel S - Static
       B - BGP Derived, E - Externally Derived, IA - OSPF Inter Area
       E1 - OSPF External Type 1, E2 - OSPF External Type 2
       N1 - OSPF NSSA External Type 1, N2 - OSPF NSSA External Type 2
       S U - Unnumbered Peer, L - Leaked Route

 * Indicates the best (lowest metric) route for the subnet.

No default gateway is configured.
B      *10.10.10.10/32 [200/0] via 192.168.100.10,   Vl200
C      *11.11.11.11/32 [0/0] directly connected,   Lo0
B      *20.20.20.20/32 [200/0] via 192.168.100.10,   Vl200
B      *172.16.10.0/24 [200/0] via 192.168.100.10,   Vl200
B      *30.30.30.30/24 [200/0] via 192.168.100.10,   Vl200
C      *192.168.100.0/24 [0/0] directly connected,   Vl200
```

On R1, the result of the routing decision process can be shown for routes coming from R2. Use the **received-routes** option to display routes received from R2.

```
R1#show ip bgp neighbors 172.16.10.2 received-routes

Local router ID is 10.10.10.10
Origin codes: i - IGP, e - EGP, ? - incomplete

Network             Next Hop         Metric     LocPref    Path          Origin
------------------- ---------------- ---------- ---------- ------------- ------
172.16.10.0/24      172.16.10.2                            6500          ?
```

```
20.20.20.20/32      172.16.10.2                                     6500          ?
30.30.30.0/24       172.16.10.2                                     6500          ?
```

Use the **routes** option to display routes received from R2.

```
R1#show ip bgp neighbors 172.16.10.2 routes

Local router ID is 10.10.10.10
Origin codes: i - IGP, e - EGP, ? - incomplete

Network             Next Hop         Metric     LocPref    Path          Origin
------------------- ---------------- ---------- ---------- ------------- ------
172.16.10.0/24      172.16.10.2                            6500          ?
20.20.20.20/32      172.16.10.2                            6500          ?
30.30.30.0/24       172.16.10.2                            6500          ?
```

Use the **rejected-routes** option to display routes received from R2 which are not matched by a policy.

```
R1#show ip bgp neighbors 172.16.10.2 rejected-routes

Local router ID is 10.10.10.10
Origin codes: i - IGP, e - EGP, ? - incomplete

Network             Next Hop         Metric     LocPref    Path          Origin
------------------- ---------------- ---------- ---------- ------------- ------
```

40

Bidirectional Forwarding Detection

Dell Networking N3000 and N4000 Series Switches

NOTE: This feature is not available on Dell Networking N1500 and N2000 Series switches.

Bidirectional Forwarding Detection (BFD) provides a lightweight fast failure detection mechanism to verify bidirectional connectivity between forwarding engines, which may be a single hop or multiple hops away from each other.

The topics covered in this chapter include:

- Overview
- BFD Operational Modes
- Limitations
- BFD Example

Overview

BFD only supports notification of failures to the BGP and OSPF protocols. The BFD protocol is designed to work over any underlying transmission mechanism and protocol layer with a wide range of detection times, especially in scenarios where fast failure detection is required in the data plane level for multiple concurrent sessions. For example, the BGP keepalive timer is 60 seconds and the hold timer is 180 seconds. A typical BFD session might use a 200-millisecond min_rx interval with a multiplier of 5. This means that BGP would detect a connectivity failure in 180 seconds, whereas BFD would report the same failure in approximately one second.

In Dell Networking N-Series switches, BFD is presented as a service that reports on the session status to its client applications.

BFD uses a simple 'hello' mechanism that is similar to the neighbor detection components of some well-known protocols. It establishes an operational session between a pair of network devices to verify a two-way communication path between them and serves information regarding the connectivity status to the applications. The pair of devices transmits BFD packets between them periodically and, if one stops receiving peer packets within the detection time limit, it considers the bidirectional path to have failed. It then notifies the application protocol of this failure.

BFD allows each device to estimate how quickly it can send and receive BFD packets to agree with its neighbor upon how fast detection of failure may be performed.

BFD operates between two devices on top of any underlying data protocol (network layer, link layer, tunnels, etc.) as payload of any encapsulating protocol appropriate for the transmission medium. The Dell Networking implementation of BFD works with IP networks (v4 and v6) and supports IPv4/v6 address-based encapsulations.

BFD is standardized in RFC 5880.

BFD Operational Modes

BFD implements two main operational modes, as well as an additional capability that may be used in combination with either of the two modes. The two modes are Asynchronous mode and Demand mode, and the additional capability is the Echo function.

Asynchronous Mode

This is the nominal operating mode for BFD. In this mode, the pair of devices periodically sends BFD control packets to one another and, if a consecutive number of those packets are not received by the other device, the session is declared down.

The asynchronous mode is advantageous as it requires half the number of packets to achieve a particular detection time as does to the echo function.

Demand Mode

In demand mode, it is assumed that a device has an independent way of verifying that it has connectivity to the other system. Once a BFD session is established, a demand mode device may ask the other to stop sending BFD control packets, except when the device needs to verify the connectivity explicitly. In this case, a short sequence of BFD Control packets, known as the Poll Sequence, is exchanged to ascertain the connectivity. Demand mode may operate independently in either direction.

Demand mode is advantageous in cases when the overhead of a periodic protocol appears burdensome on a device, e.g., a router with a large number of BFD sessions running.

Dell Networking BFD does not support demand mode.

Echo Function

Echo mode is an auxiliary operation that may be used with either BFD mode. When the echo function is active, a stream of BFD echo packets is transmitted in such a way that the other system loops them back through its forwarding path. If a configured number of consecutive packets of the echoed data stream are not received, the session is declared to be down. Since the echo function is handling the task of neighbor detection, the rate of periodic transmission of BFD control packets may be reduced (in the case of asynchronous mode) or eliminated completely (in the case of demand mode).

The echo function has the advantage of testing the forwarding path on the remote system. This may reduce round-trip jitter and, thus, allow more aggressive detection times, and can potentially catch some classes of failure that might not otherwise be detected.

Limitations

- Dell Networking BFD does not support demand mode.
- Dell Networking BFD does not support authentication.
- The BFD feature provides notification to BGP or OSPF when an interface is detected to not be in a forwarding state. No other routing protocols are supported.
- BFD is supported in the default VRF only.
- BFD should be configured on routed interfaces only. BFD should not be configured mirrored ports or on interfaces enabled for IEEE 802.1x.
- BFD is supported across link aggregation groups, but does not detect individual LAG member link failure.
- BFD does not operate on the out-of-band interface.

BFD Example

This example configures BFD for a BGP peer session. BFD is only supported in conjunction with BGP. The BGP configuration is taken from BGP Redistribution of OSPF Example in the BGP Configuration Examples section and is not explained further here. The fast-external-fallover is not enabled in this example, as BFD will provide failure detection.

1 Enable the BFD feature. This step is mandatory before configuring or enabling BFD:

```
console#config
console(config)#feature bfd
console(config)#interface te1/0/1
console(config-if-Te1/0/1)#switchport mode trunk
console(config-if-Te1/0/1)#switchport trunk native vlan 100
console(config-if-Te1/0/1)#exit

console(config)#ip routing
```

2 Configure a VLAN routing interface and enable notification to BGP on routing connectivity failure.

   (Optional) Configure BFD sessions parameter on the BGP peer link.

   BFD echo mode is enabled first. Then the BFD control packet interval is set to 1 second. This configuration will send echo packets every 100 ms. If three consecutive control or echo packets are missed, the interface is declared down. BGP fast external failover is disabled on the peer interface as BFD will notify BGP if the routing peer is not reachable:

```
console(config)#vlan 10
console(config-vlan10)#exit

console(config)#interface vlan 10
console(config-if-vlan10)#ip address 172.16.65.1 /30
console(config-if-vlan10)#ip bgp fast-external-fallover deny
console(config-if-vlan10)#bfd echo
console(config-if-vlan10)#bfd slow-timer 1000
console(config-if-vlan10)#exit

console(config)#ip route 0.0.0.0 0.0.0.0 172.16.65.2 name
'Default-Route'
console(config)#router bgp 3434
console(config-router)#bgp router-id 172.16.64.1
```

```
console(config-router)#neighbor 216.31.219.19 remote-private-
as 1402
console(config-router)#redistribute static
console(config-router)#redistribute ospf match external 1
console(config-router)#redistribute ospf match external 2
```

3 Enable a BFD session on the BGP peer link:

```
console(config-router)#neighbor 216.31.219.19 bfd fast-
fallover
console(config-router)#exit
```

41

IPv6 Routing

Dell Networking N3000 and N4000 Series Switches

NOTE: This feature is not available on Dell Networking N1500 and N2000 Series switches.

This chapter describes how to configure general IPv6 routing information on the switch, including global routing settings and IPv6 static routes. The topics covered in this chapter include:

- IPv6 Routing Overview
- Default IPv6 Routing Values
- Configuring IPv6 Routing Features (Web)
- Configuring IPv6 Routing Features (CLI)
- IPv6 Static Reject and Discard Routes
- IPv6 Router Advertisement Guard

The Dell Networking N-Series switches support additional features to help manage IPv6 networks, including OSPFv3, DHCPv6, and IPv6 multicast. For information about OSPFv3, see "OSPF and OSPFv3 " on page 1181. For information about DHCPv6, see "DHCPv6 Server and Relay Settings " on page 1425. For information about IPv6 multicast, see "IPv4 and IPv6 Multicast " on page 1509.

For configuration examples that include IPv6 interface configuration, see "OSPF Configuration Examples " on page 1247

IPv6 Routing Overview

IPv6 is the next generation of the Internet Protocol. With 128-bit addresses, versus 32-bit addresses for IPv4, IPv6 solves the address depletion issues seen with IPv4 and removes the requirement for Network Address Translation (NAT), which is used in IPv4 networks to reduce the number of globally unique IP addresses required for a given network.

On the Dell Networking N1500, N2000, N3000, and N4000 Series switches, IPv6 coexists with IPv4. As with IPv4, IPv6 routing can be enabled on loopback and VLAN interfaces. Each L3 routing interface can be used for IPv4, IPv6, or both. IP protocols running over L3 (for example, UDP and TCP) are common to both IPv4 and IPv6.

How Does IPv6 Compare with IPv4?

There are many conceptual similarities between IPv4 and IPv6 network operation. Addresses still have a network prefix portion (network) and a device interface specific portion (host). While the length of the network portion is still variable, most users have standardized on using a network prefix length of 64 bits. This leaves 64 bits for the interface specific portion, called an Interface ID in IPv6. Depending upon the underlying link addressing, the Interface ID can be automatically computed from the link (e.g., MAC address). Such an automatically computed Interface ID is called an EUI-64 identifier, which is the interface MAC address with ff:fe inserted in the middle.

IPv6 packets on the network are of an entirely different format than traditional IPv4 packets and are also encapsulated in a different EtherType (86DD rather than 0800 which is used with IPv4). The details for encapsulating IPv6 in Ethernet frames are described in RFC2462.

Unlike IPv4, IPv6 does not have broadcasts. There are two types of IPv6 addresses — unicast and multicast. Unicast addresses allow direct one-to-one communication between two hosts, whereas multicast addresses allow one-to-many communication. Multicast addresses are used as destinations only. Unicast addresses will have 00 through fe in the most significant octets and multicast addresses will have ff in the most significant octets.

How Are IPv6 Interfaces Configured?

The basic IPv6 protocol specifies two classes of PDU options, both of which are supported: hop-by-hop options and destination. Although new options may be defined in the future, the following are currently supported: routing (for source routing), fragment, router alert, and pad. IPv6 jumbograms (RFC 2675) are not supported. In IPv6, only source nodes fragment. ICMPv6 support of path MTU discovery is therefore supported. IPv6 forwarded or routed packets are never fragmented by the switch. IPv6 flow labels are ignored.

Neighbor Discovery (ND) protocol is the IPv6 replacement for Address Resolution Protocol (ARP) in IPv4. The IPv6 Neighbor Discovery protocol is described in detail in RFC7048. Dell Networking IPv6 supports neighbor advertise and solicit, duplicate address detection, and unreachability detection. Router advertisement is part of the Neighbor Discovery process and is required for IPv6. As part of router advertisement, Dell Networking N1500, N2000, N3000, and N4000 Series switches software supports stateless auto configuration of end nodes. The switch supports both EUI-64 interface identifiers and manually configured interface IDs.

For ICMPv6, error PDU generation is supported, as are path MTU, echo, and redirect.

While optional in IPv4, router advertisement is mandatory in IPv6. Router advertisements specify the network prefix(es) on a link which can be used by receiving hosts, in conjunction with an EUI-64 identifier, to autoconfigure a host's address. Routers have their network prefixes configured and may use EUI-64 or manually configured interface IDs. In addition to a single global address and a single unique local address in the fc00::/7 range, each IPv6 interface also has an autoconfigured "link-local" address, which is:

- fe80::/10, with the EUI-64 address in the least significant bits.
- Reachable only on the local VLAN — link-local addresses are never routed.
- Not globally unique

Next hop addresses computed by routing protocols are usually link-local addresses.

During the period of transitioning the Internet to IPv6, a global IPv6 Internet backbone may not be available. One transition mechanism is to tunnel IPv6 packets inside IPv4 to reach remote IPv6 islands. When a packet is sent over such a link, it is encapsulated in IPv4 in order to traverse an IPv4 network and has the IPv4 headers removed at the other end of the tunnel.

Default IPv6 Routing Values

IPv6 is disabled by default on the switch and on all interfaces.

Table 41-1 shows the default values for the IP routing features this chapter describes.

Table 41-1. IPv6 Routing Defaults

Parameter	Default Value
IPv6 Unicast Routing Mode	Disabled
IPv6 Hop Limit	Unconfigured
ICMPv6 Rate Limit Error Interval	1000 milliseconds
ICMPv6 Rate Limit Burst Size	100
Interface IPv6 Mode	Disabled
IPv6 Router Route Preferences	Local—0 Static—1 OSPFv3 Intra—110 OSPFv3 Inter—110 OSPFv3 External—110
IPv6 Router Advertisement Guard	Disabled

Table 41-2 shows the default IPv6 interface values after a VLAN routing interface has been created.

Table 41-2. IPv6 Interface Defaults

Parameter	Default Value
IPv6 Mode	Disabled
DHCPv6 Client Mode	Disabled
Stateless Address AutoConfig Mode	Disabled
Routing Mode	Enabled
Interface Maximum Transmit Unit	1500
Router Duplicate Address Detection Transmits	1

Table 41-2. IPv6 Interface Defaults (Continued)

Parameter	Default Value
Router Advertisement NS Interval	Not configured
Router Lifetime Interval	1800 seconds
Router Advertisement Reachable Time	0 seconds
Router Advertisement Interval	600 seconds
Router Advertisement Managed Config Flag	Disabled
Router Advertisement Other Config Flag	Disabled
Router Advertisement Suppress Flag	Disabled
IPv6 Destination Unreachables	Enabled

Configuring IPv6 Routing Features (Web)

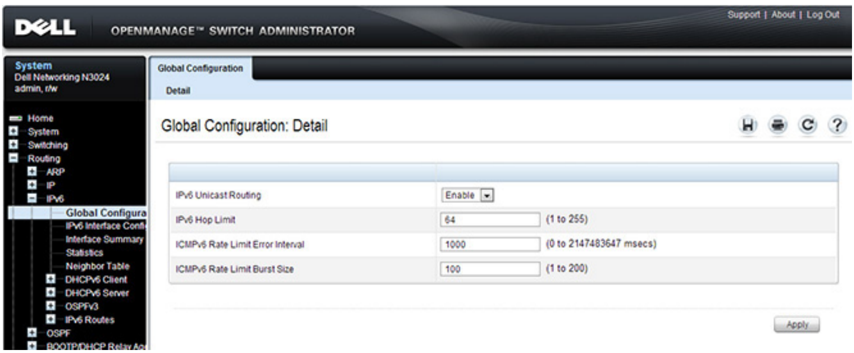
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring IPv6 unicast routing features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click  at the top of the page.

Global Configuration

Use the **Global Configuration** page to enable IPv6 forwarding on the router, enable the forwarding of IPv6 unicast datagrams, and configure global IPv6 settings.

To display the page, click **Routing** → **IPv6** → **Global Configuration** in the navigation panel.

Figure 41-1. IPv6 Global Configuration

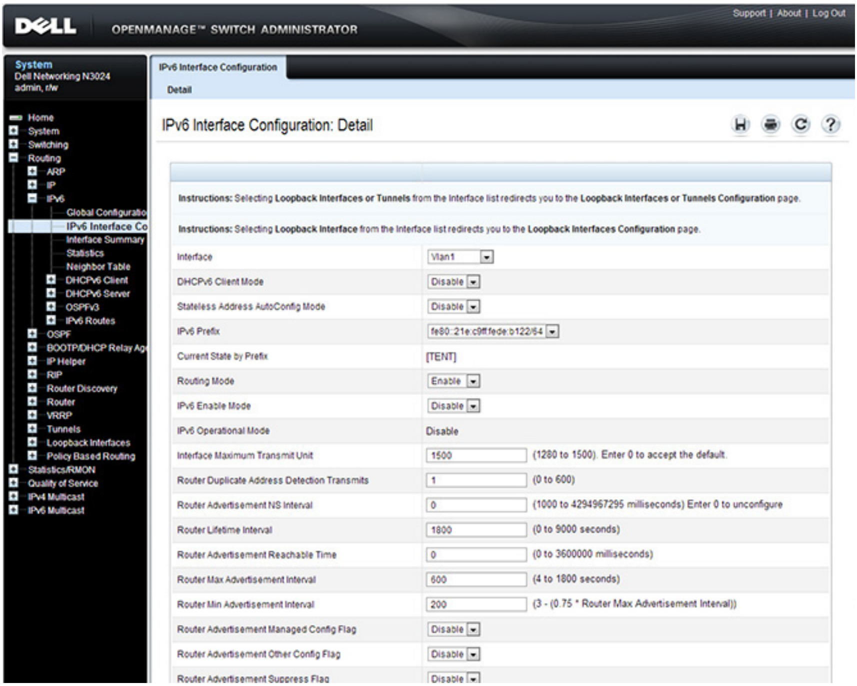


Interface Configuration

Use the **Interface Configuration** page to configure IPv6 interface parameters. This page has been updated to include the IPv6 Destination Unreachables field.

To display the page, click **Routing** → **IPv6** → **Interface Configuration** in the navigation panel.

Figure 41-2. IPv6 Interface Configuration

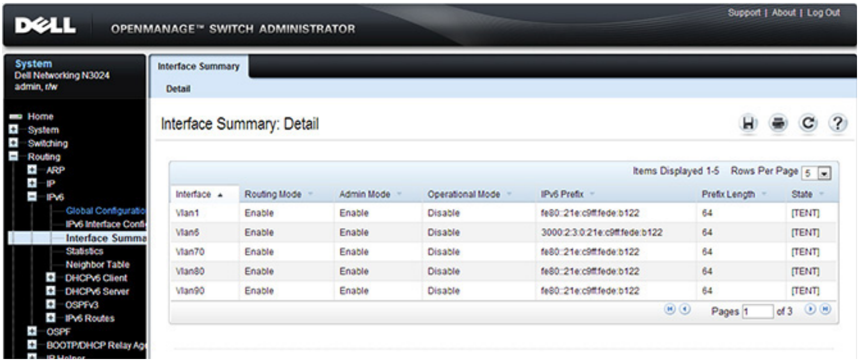


Interface Summary

Use the **Interface Summary** page to display settings for all IPv6 interfaces.

To display the page, click **Routing** → **IPv6** → **Interface Summary** in the navigation panel.

Figure 41-3. IPv6 Interface Summary

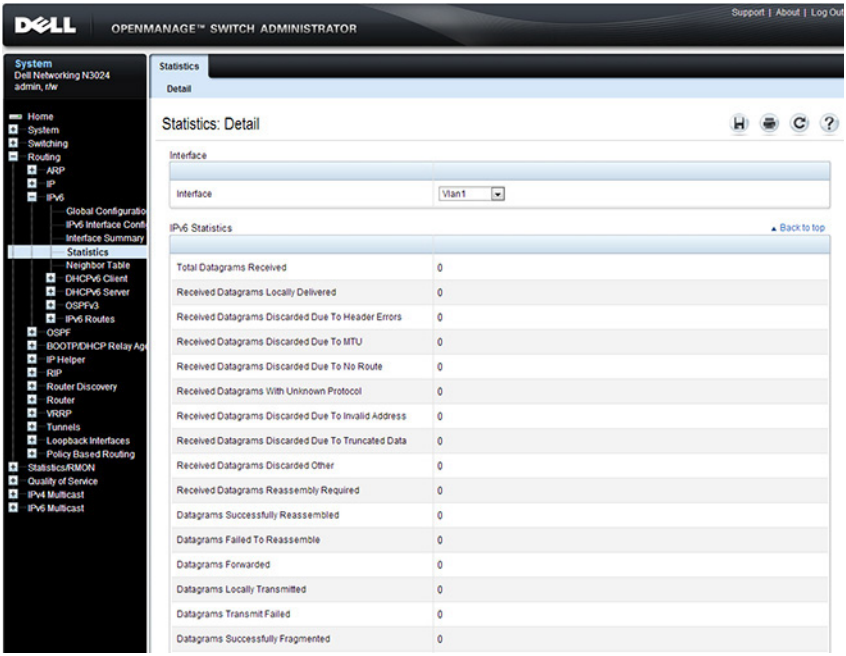


IPv6 Statistics

Use the **IPv6 Statistics** page to display IPv6 traffic statistics for one or all interfaces.

To display the page, click **Routing** → **IPv6** → **IPv6 Statistics** in the navigation panel.

Figure 41-4. IPv6 Statistics

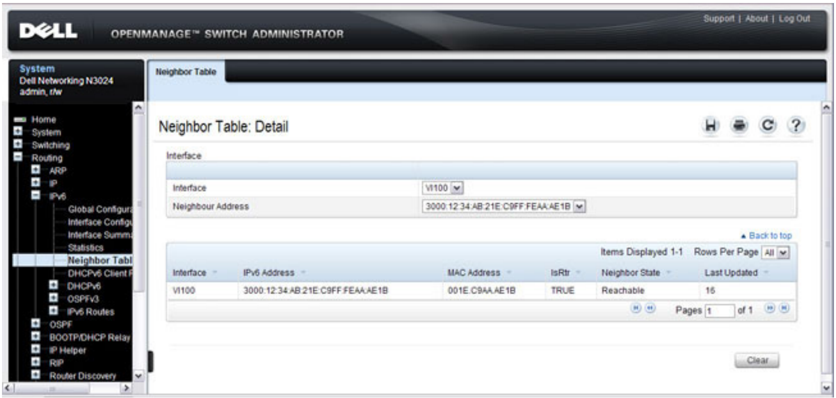


IPv6 Neighbor Table

Use the **IPv6 Neighbor Table** page to display IPv6 neighbor details for a specified interface.

To display the page, click **IPv6 → IPv6 Neighbor Table** in the navigation panel.

Figure 41-5. IPv6 Neighbor Table

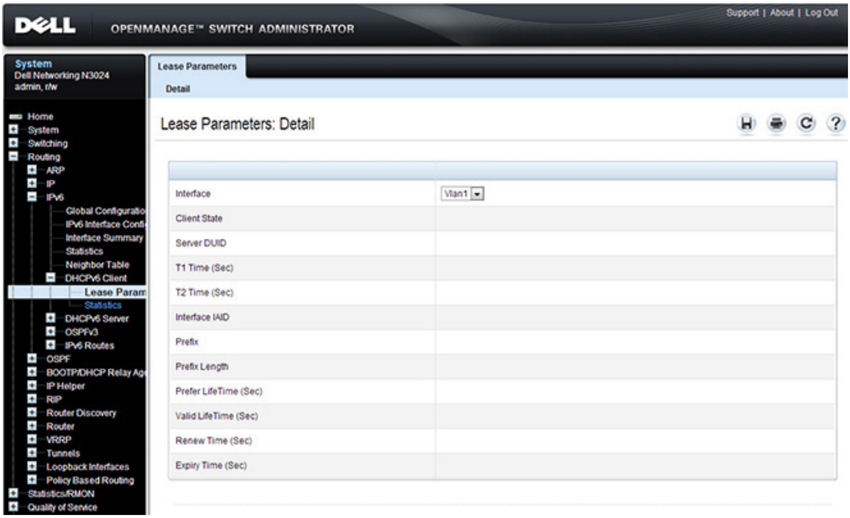


DHCPv6 Client Parameters

Use the **DHCPv6 Client Parameters** page to view information about the network information automatically assigned to an interface by the DHCPv6 server. This page displays information only if the DHCPv6 client has been enabled on an IPv6 routing interface.

To display the page, click **Routing** → **IPv6** → **DHCPv6 Client** → **Lease Parameters** in the navigation panel.

Figure 41-6. DHCPv6 Lease Parameters

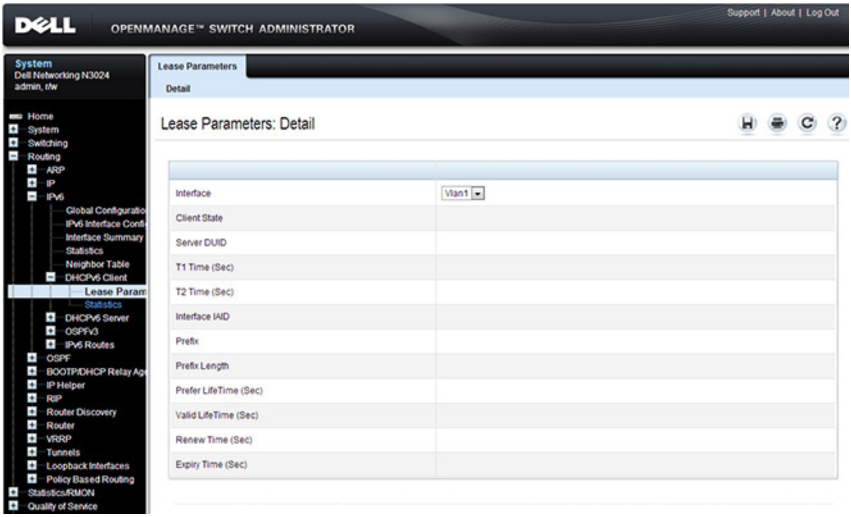


DHCPv6 Client Statistics

Use the **DHCPv6 Client Statistics** page to view information about DHCPv6 packets received and transmitted on a DHCPv6 client interface.

To display the page, click **Routing** → **IPv6** → **DHCPv6 Client** → **Statistics** in the navigation panel.

Figure 41-7. DHCPv6 Lease Parameters

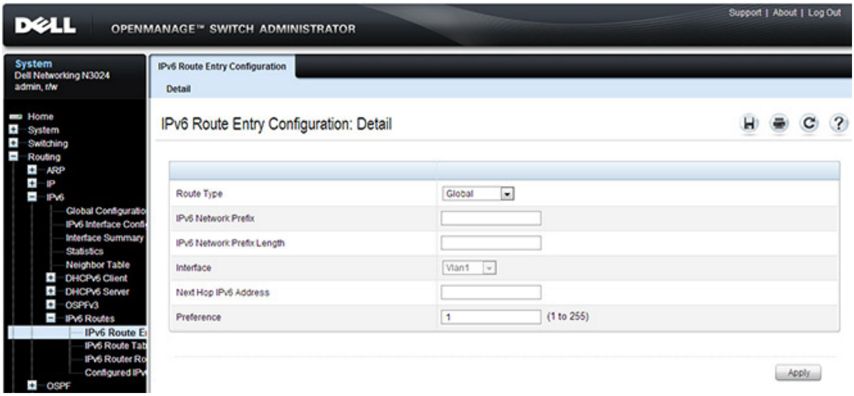


IPv6 Router Entry Configuration

Use the **IPv6 Route Entry Configuration** page to configure information for IPv6 routes.

To display the page, click **Routing** → **IPv6** → **IPv6 Routes** → **IPv6 Route Entry Configuration** in the navigation panel.

Figure 41-8. IPv6 Route Entry Configuration

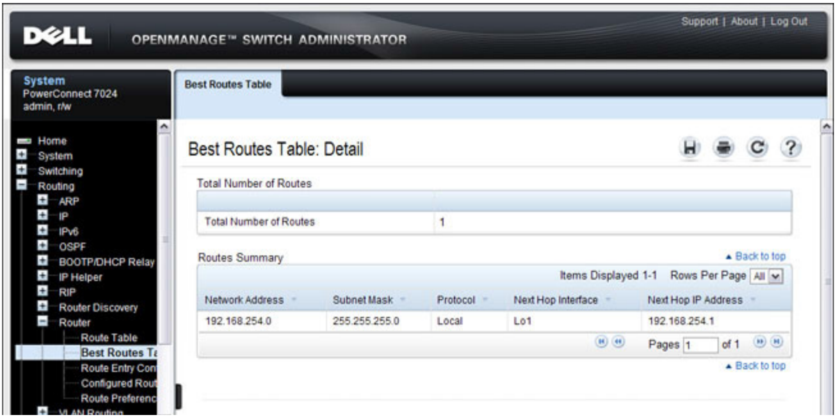


IPv6 Route Table

Use the **IPv6 Route Table** page to display all active IPv6 routes and their settings.

To display the page, click **Routing** → **IPv6** → **IPv6 Routes** → **IPv6 Route Table** in the navigation panel.

Figure 41-9. IPv6 Route Table

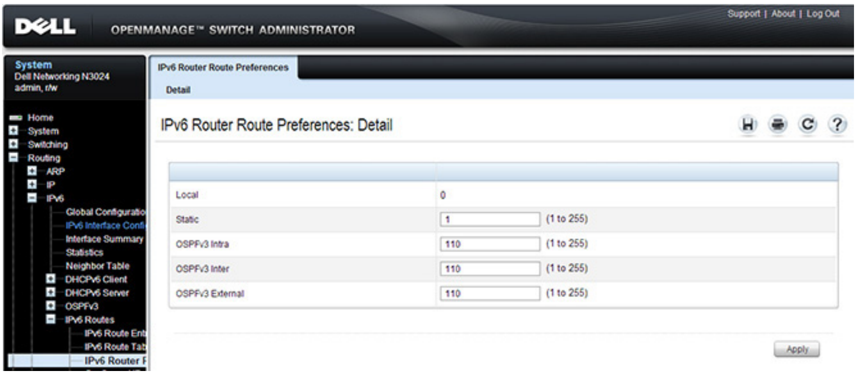


IPv6 Route Preferences

Use the **IPv6 Route Preferences** page to configure the default preference for each protocol. These values are arbitrary values in the range of 1 to 255 and are independent of route metrics. Most routing protocols use a route metric to determine the shortest path known to the protocol, independent of any other protocol. The best route to a destination is chosen by selecting the route with the lowest preference value. When there are multiple routes to a destination, the preference values are used to determine the preferred route. If there is still a tie, the route with the best route metric is chosen. To avoid problems with mismatched metrics, you must configure different preference values for each of the protocols.

To display the page, click **Routing** → **IPv6** → **IPv6 Routes** → **IPv6 Route Preferences** in the navigation panel.

Figure 41-10. IPv6 Route Preferences



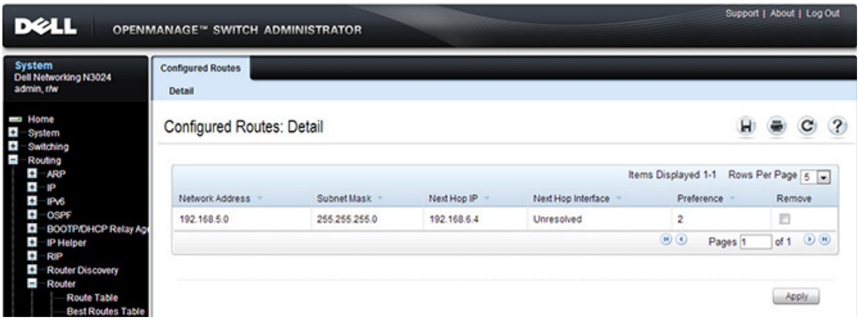
Configured IPv6 Routes

Use the **Configured IPv6 Routes** page to display selected IPv6 routes.

NOTE: For a static reject route, the next hop interface value is Null0. Packets to the network address specified in static reject routes are intentionally dropped.

To display the page, click **Routing** → **IPv6** → **IPv6 Routes** → **Configured IPv6 Routes** in the navigation panel.

Figure 41-11. Configured IPv6 Routes



To remove a configured route, select the check box in the **Delete** column of the route to remove, and click **Apply**.

Configuring IPv6 Routing Features (CLI)

This section provides information about the commands used for configuring IPv6 routing on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Global IP Routing Settings

Beginning in Privileged EXEC mode, use the following commands to configure various global IP routing settings for the switch.

Command	Purpose
configure	Enter global configuration mode.
sdm prefer dual-ipv4-and-ipv6 default	Select a Switch Database Management (SDM) template to enable support for both IPv4 and IPv6. Changing the SDM template requires a system reload.
ipv6 unicast-routing	Globally enable IPv6 routing on the switch.
ipv6 hop-limit *limit*	Set the TTL value for the router. The valid range is 0 to 255.
ipv6 icmp error-interval *burst-interval* [*burst-size*]	Limit the rate at which IPv4 ICMP error messages are sent. • *burst-interval* — How often the token bucket is initialized (Range: 0–2147483647 milliseconds). • *burst-size* — The maximum number of messages that can be sent during a burst interval (Range: 1–200).
exit	Exit to Privileged EXEC mode.

Configuring IPv6 Interface Settings

Beginning in Privileged EXEC mode, use the following commands to configure IPv6 settings for VLAN, tunnel, or loopback interfaces.

Command	Purpose
configure	Enter Global Configuration mode.
interface {**vlan** | **tunnel** | **loopback**} *interface-id*	Enter Interface Configuration mode for the specified VLAN, tunnel, or loopback interface.
ipv6 enable	Enable IPv6 on the interface. Configuring an IPv6 address will automatically enable IPv6 on the interface.
ipv6 address {**autoconfig** | **dhcp** | *prefix/prefix-length* [**eui64**]}	Configure the IPv6 address and network prefix length. Setting an IPv6 address enables IPv6 on the interface. The **ipv6 enable command** can be used to enable IPv6 on the interface without setting an address. Link-local, multicast, IPv4-compatible, and IPv4-mapped addresses are not allowed to be configured. Multiple globally unique unicast addresses (2001::/23) with non-overlapping subnets and one unique local (fc00::/7) address may be configured in addition to the link local address. Include the **EUI-64** keyword to have the system add the 64-bit interface ID to the address. You must use a network prefix length of 64 in this case. For VLAN interfaces, use the **dhcp** keyword to enable the DHCPv6 client and obtain an IP address form a network DHCPv6 server.
ipv6 traffic-filter *ACL name*	Add an access-list filter to this interface.
ipv6 unreachables	(VLAN interfaces only) Allow the interface to send ICMPv6 Destination Unreachable messages. The **no ipv6 unreachables** command suppresses the ICMPv6 unreachable messages for this interface.
exit	Exit the interface configuration mode.

Configuring IPv6 Neighbor Discovery

Use the following commands to configure IPv6 Neighbor Discovery settings.

Command	Purpose
ipv6 nd prefix *prefix/prefix-length* [{*valid-lifetime*| **infinite**} {*preferred-lifetime*| **infinite**}] [**no-autoconfig**] [**off-link**]	Configure parameters associated with network prefixes that the router advertises in its Neighbor Discovery advertisements. • ipv6-prefix—IPv6 network prefix. • *prefix-length*—IPv6 network prefix length. • *valid-lifetime*—Valid lifetime of the router in seconds. (Range: 0–4294967295 seconds.) • **infinite**—Indicates lifetime value is infinite. • *preferred-lifetime*—Preferred-lifetime of the router in seconds. (Range: 0–4294967295 seconds.) • **no-autoconfig**—Do not use the prefix for auto configuration. • **off-link**—Do not use the prefix for onlink determination.
ipv6 nd ra-interval *maximum minimum*	Set the transmission interval between router Neighbor Discovery advertisements. • *maximum* — The maximum interval duration (Range: 4–1800 seconds). • *minimum* — The minimum interval duration (Range: 3 – (0.75 * maximum) seconds).
ipv6 nd ra-lifetime *seconds*	Set the value that is placed in the Router Lifetime field of the router Neighbor Discovery advertisements sent from the interface. The *seconds* value must be zero, or it must be an integer between the value of the router advertisement transmission interval and 9000 seconds. A value of zero means this router is not to be used as the default router. (Range: 0-9000).
ipv6 nd suppress-ra	Suppress router advertisement transmission on an interface.
ipv6 nd dad attempts *value*	Set the number of duplicate address detection probes transmitted while doing Neighbor Discovery. The range for *value* is 0–600.

Command	Purpose
ipv6 nd ns-interval *milliseconds*	Set the interval between router advertisements for advertised neighbor solicitations. The range is 1000 to 4294967295 milliseconds.
ipv6 nd other-config-flag	Set the *other stateful configuration* flag in router advertisements sent from the interface.
ipv6 nd managed-config-flag	Set the *managed address configuration* flag in router advertisements. When the value is true, end nodes use DHCPv6. When the value is false, end nodes automatically configure addresses.
ipv6 nd reachable-time *milliseconds*	Set the router advertisement time to consider a neighbor reachable after neighbor discovery confirmation.

Configuring IPv6 Route Table Entries and Route Preferences

Beginning in Privileged EXEC mode, use the following commands to configure IPv6 Static Routes.

Command	Purpose
configure	Enter global configuration mode.
ipv6 route *ipv6-prefix/prefix-length* {*next-hop-address* | *interface-type interface-number next-hop-address* } [*preference*]	Configure a static route.Use the keyword **null** instead of the next hop router IP address to configure a static reject route. • *prefix/prefix-length*—The IPv6 network prefix and prefix length that is the destination of the static route. Use the ::/0 form (unspecified address and zero length prefix) to specify a default route. • *interface-type interface-number*—Must be specified when using a link-local address as the next hop. The interface-type can be **vlan** or **tunnel**. • *next-hop-address* —The IPv6 address of the next hop that can be used to reach the specified network. A link-local next hop address must have a prefix length of 128. The next hop address cannot be an unspecified address (all zeros), a multicast address, or a loopback address. If a link local next hop address is specified, the interface (VLAN or tunnel), must also be specified. • *preference*—Also known as Administrative Distance, a metric the router uses to compare this route with routes from other route sources that have the same network prefix. (Range: 1-255). Lower values have precedence over higher values. The default preference for static routes is 1. Routes with a preference of 255 are considered as "disabled" and will not be used for forwarding. Routes with a preference metric of 254 are used by the local router but will never be advertised to other neighboring routers.
ipv6 route *ipv6-prefix/prefix-length* **null** [*preference*]	Configure a static reject route. IPv6 packets matching the reject route will be silently discarded.

Command	Purpose
ipv6 route distance *integer*	Set the default distance (preference) for static IPv6 routes. Lower route preference values are preferred when determining the best route. The default distance (preference) for static routes is 1.
exit	Exit to Global Config mode.

IPv6 Show Commands

Use the following commands in Privileged EXEC mode to view IPv6 configuration status and related data.

Command	Purpose
show sdm prefer	Show the currently active SDM template.
show sdm prefer dual-ipv4-and-ipv6 default	Show parameters for the SDM template.
show ipv6 dhcp interface vlan *vlan-id*	View information about the DHCPv6 lease acquired by the specified interface.
show ipv6 interface {**vlan** | **tunnel** | **loopback**} *interface-id*	View the IP interface configuration information for the specified IPv6 routing interface.
show ipv6 brief	View the global IPv6 settings for the switch.
show ipv6 route [*ipv6-address* | *ipv6-prefix/prefix-length* | *protocol* | *interface-type interface-number*] [**best**]	View the routing table. • *ipv6-address*—Specifies an IPv6 address for which the best-matching route would be displayed. • *protocol*—Specifies the protocol that installed the routes. Is one of the following keywords: connected, ospf, static. • *ipv6-prefix/ prefix-length*—Specifies an IPv6 network for which the matching route would be displayed. • *interface-type interface-number*—Valid IPv6 interface. Specifies that the routes with next-hops on the selected interface be displayed. • **best**—Specifies that only the best routes are displayed. If the connected keyword is selected for protocol, the best option is not available because there are no best or non-best connected routes.
show ipv6 route summary	View summary information about the IPv6 routing table.
show ipv6 route preferences	View detailed information about the IPv6 route preferences.

IPv6 Static Reject and Discard Routes

A static configured route with a next-hop of "null" causes any packet matching the route to disappear or vanish from the network. This type of route is called a "Discard" route if the router returns an ICMP "network-unreachable" message, or is called a "Reject" route if no ICMP message is returned. The Dell Networking N-Series switches support "Reject" routes, where any packets matching the route network prefix silently disappear.

A common use of a Reject route is to quickly discard packets that cannot be delivered because a valid route to the destination is not known. Without the Reject route, these undeliverable packets will continue to circulate through the network, following the default routes, until their TTL expires. Forwarding packets that cannot be delivered wastes bandwidth, particularly on expensive WAN connections. The Reject route will also suppress a type of "Denial of Service" (DoS) attack where an internal host sends large numbers of packets to unknown destinations, causing congestion of the WAN links.

- ipv6 route ::/0 null 254

  Use this in all routers except the ones with direct Internet connectivity. Routers with direct Internet connectivity should advertise a default route. The effect of this route is that when a router does not have connectivity to the Internet, the router will quickly discard packets that it cannot deliver.

  If the router learns a default route from another router, the learned route will have a lower distance metric and therefore a higher preference. Routes that are more specific (have more bits in the prefix) will have precedence over less specific routes. This will cause packets destined for non-existent networks to be quickly discarded. Also, because of the high distance metric (254), this route will never be advertised to any neighbor routers.

- ipv6 route fc00::/7 null 254

  This route covers the entire ULA (IPv6 private) address space. If you have networks configured in this address space, you will have more specific routes for those networks. The more specific routes (more bits of prefix) will have precedence over this route. Any destinations in this range not known via another, more specific route do not exist. The effect is that packets destined for private networks that do not exist in your network will be quickly discarded instead of being forwarded to the default route.

- ipv6 route 2001::/16 null 254
  ipv6 route 2002::/16 null 254

  These address ranges are reserved and not reachable in the Internet. If for some reason you have local networks in this range, a more specific route will have precedence.

Another use for the Reject route is to prevent internal hosts from communication with specific addresses or ranges of addresses. The effect is the same as an outgoing access-list with a “deny” statement. A route is generally more efficient than an access-list that performs the same function. If you need more fine-grained filtering, such as protocols or port numbers, use the access-list instead.

IPv6 Router Advertisement Guard

Dell Networking N-Series switches support IPv6 Router Advertisement Guard (RA-Guard) to protect against attacks via rogue Router Advertisements in accordance with RFC 6105. Dell Networking RA-Guard supports Stateless RA-Guard, where the administrator can configure the interface to allow received router advertisements and router redirect message to be processed/forwarded or dropped.

By default, RA-Guard is not enabled on any interfaces. RA-Guard is enabled/disabled on physical interfaces or port-channels. RA-Guard does not require IPv6 routing to be enabled. This allows VLANs to span interfaces connected to routers and hosts, while allowing configuration such that router advertisements or redirect messages received from connected hosts are dropped (L2 configuration). L3 configuration of RA-Guard on IPv6 routing interfaces is also supported.

Dell Networking supports a single unnamed RA-Guard policy that blocks all incoming IPv6 router advertisements and IPv6 router redirect messages. The single unnamed policy is pre-configured and may not be renamed or removed.

The following example configures the unnamed RA-Guard policy to drop all RA advertisements and router redirect messages on host connected routed interface Gi1/0/1. In the example, routed VLAN 10 is isolated to physical interface Gi1/0/1 connected to a host. IPv6 routing is enabled on VLAN 10 and IPv6 unicast routing is enabled globally. Interface gi1/0/1 is placed into access mode, meaning untagged incoming and outgoing packets are processed on VLAN 10. RA-Guard is enabled on interface Gi1/0/1 and then the configuration is verified with the show command.

```
console(config)#vlan 10
console(config-vlan10)#exit
console(config)#interface vlan 10
console(config-if-vlan10)#ipv6 enable
console(config-if-vlan10)#exit
console(config)#ipv6 unicast-routing
console(config)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport mode access
console(config-if-Gi1/0/1)#switchport access vlan 10
console(config-if-Gi1/0/1)#exit
console(config)#interface gi1/0/1
console(config-if-Gi1/0/1)#ipv6 nd raguard attach-policy
console(config-if-Gi1/0/1)#show ipv6 nd raguard policy

Ipv6 RA-Guard Configured Interfaces

Interface         Role
---------------   -------
Gi1/0/1           Host
```

The following example configures the unnamed RA-Guard policy to drop all RA advertisements and router redirect messages on host connected interface Gi1/0/1. In the example, switched VLAN 10 spans physical interface Gi1/0/1 connected to a host as well as interface Te1/0/1 connected to a router. Interface gi1/0/1 is placed into access mode, meaning untagged incoming and outgoing packets are processed on VLAN 10. RA-Guard is enabled on interface Gi1/0/1 and then the configuration is verified with the show command.

```
console(config)#vlan 10
console(config-vlan10)#exit
console(config)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport mode access
console(config-if-Gi1/0/1)#switchport access vlan 10
console(config-if-Gi1/0/1)#exit
console(config)#interface te1/0/1
console(config-if-Te1/0/1)#switchport mode trunk
console(config-if-Te1/0/1)#exit
console(config)#interface gi1/0/1
```

```
console(config-if-Gi1/0/1)#ipv6 nd raguard attach-policy
console(config-if-Gi1/0/1)#show ipv6 nd raguard policy

Ipv6 RA-Guard Configured Interfaces

Interface        Role
---------------  -------
Gi1/0/1          Host
```

42

DHCPv6 Server and Relay Settings

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

NOTE: The DHCPv6 Server is not available on the Dell Networking N1500 Series switches.

This chapter describes how to configure the switch to dynamically assign network information to IPv6 hosts by using the Dynamic Host Configuration Protocol for IPv6 (DHCPv6).

The topics covered in this chapter include:

- DHCPv6 Overview
- Default DHCPv6 Server and Relay Values
- Configuring the DHCPv6 Server and Relay (Web)
- Configuring the DHCPv6 Server and Relay (CLI)
- DHCPv6 Configuration Examples

DHCPv6 Overview

DHCP is a protocol that is generally used between clients and servers for the purpose of assigning IP addresses, gateways, and other networking definitions such as Domain Name System (DNS) and Network Time Protocol (NTP) parameters. However, IPv6 natively provides IP address auto configuration through IPv6 Neighbor Discovery Protocol (NDP) and through the use of Router Advertisement messages. Thus, the role of DHCPv6 within the network is different than that of DHCPv4 because DHCPv6 is not the primary source for IP address assignment.

DHCPv6 server and client interactions are described by RFC 3315. There are many similarities between DHCPv6 and DHCPv4 interactions and options, but there are enough differences in the messages and option definitions that there is no DHCPv4 to DHCPv6 migration or interoperability.

What Is a DHCPv6 Pool?

DHCPv6 pools are used to specify information for DHCPv6 server to distribute to DHCPv6 clients. These pools are shared between multiple interfaces over which DHCPv6 server capabilities are configured.

What Is a Stateless Server?

DHCPv6 incorporates the notion of the *stateless* server, where DHCPv6 is not used for IP address assignment to a client; rather, it provides other networking information such as DNS or NTP information. The stateless server behavior is described by RFC 3736, which simply contains descriptions of the portions of RFC 3315 that are necessary for stateless server behavior. In order for a router to drive a DHCPv6 client to utilize stateless DHCPv6, the *other stateful configuration* option must be configured for neighbor discovery on the corresponding IPv6 router interface. This, in turn, causes DHCPv6 clients to send the DHCPv6 Information Request message in response. A DHCPv6 server then responds by providing only networking definitions such as DNS domain name and server definitions, NTP server definitions, or SIP definitions.

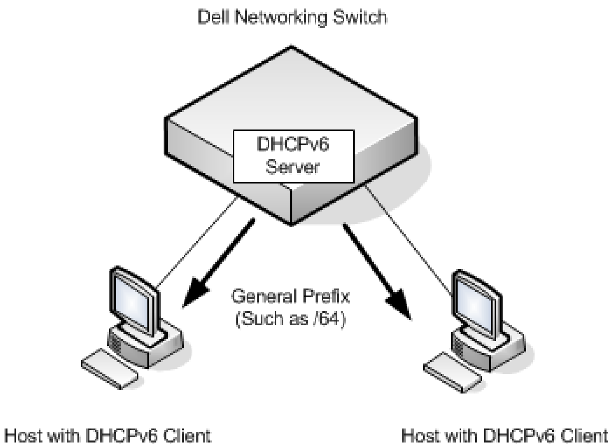
What Is the DHCPv6 Relay Agent Information Option?

The DHCPv6 Relay Agent Information Option allows for various sub-options to be attached to messages that are being relayed by the local router to a DHCPv6 server. The DHCPv6+ server may in turn use this information in determining an address to assign to a DHCPv6 client.RFC 3315 also describes DHCPv6 Relay Agent interactions, which are very much like DHCPv4 Relay Agents. Additionally, there is a DHCPv6 Relay Agent Option described in RFC 4649, which employs very similar capabilities as those described by the DHCPv4 Relay Agent Option in RFC 2132.

What Is a Prefix Delegation?

With the larger address space inherent to IPv6, addresses within a network can be allocated more effectively in a hierarchical fashion. DHCPv6 introduces the notion of prefix delegation as described in RFC 3633 as a way for routers to centralize and delegate IP address assignment. Figure 42-1 depicts a typical network scenario where prefix delegation is used.

Figure 42-1. DHCPv6 Prefix Delegation Scenario



In Figure 42-1, the Dell Networking switch acts as the Prefix Delegation (PD) server and defines one or more *general* prefixes to allocate and assign addresses to hosts that may be utilizing IPv6 auto-address configuration or acting as DHCPv6 clients.

DHCPv6 clients may request multiple IPv6 prefixes. Also, DHCPv6 clients may request specific IPv6 prefixes. If the configured DHCPv6 pool contains the specific prefix that a DHCPv6 client requests, then that prefix will be delegated to the client. Otherwise, the first available IPv6 prefix within the configured pool will be delegated to the client.

Default DHCPv6 Server and Relay Values

By default, the DHCPv6 server is disabled, and no address pools are configured. VLAN routing interfaces are not configured to perform DHCPv6 server or DHCPv6 relay functions.

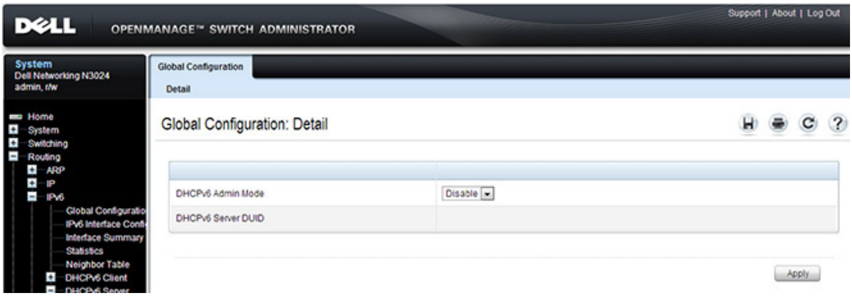
Configuring the DHCPv6 Server and Relay (Web)

This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring the DHCPv6 server on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

DHCPv6 Global Configuration

Use the **Global Configuration** page to configure DHCPv6 global parameters. To display the page, click **Routing** → **IPv6** → **DHCPv6 Server** → **Global Configuration** in the navigation panel.

Figure 42-2. DHCPv6 Global Configuration

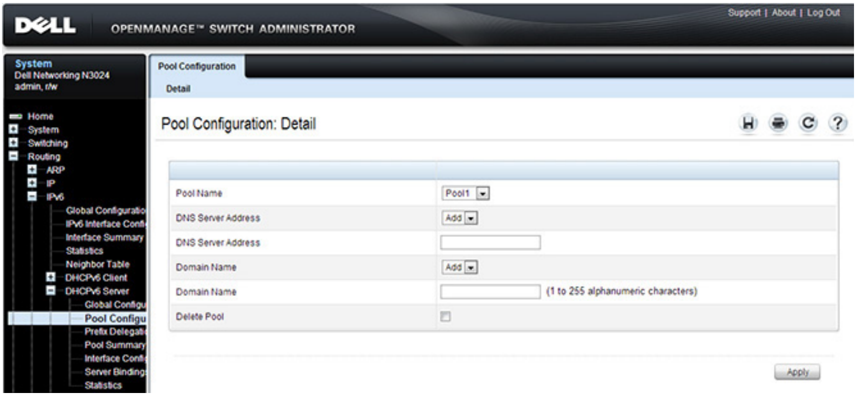


DHCPv6 Pool Configuration

Use the **Pool Configuration** page to set up a pool of DHCPv6 parameters for DHCPv6 clients. The pool is identified with a pool name and contains IPv6 addresses and domain names of DNS servers.

To display the page, click **Routing** → **IPv6** → **DHCPv6** → **Pool Configuration** in the navigation panel. Figure 42-3 shows the page when no pools have been created. After a pool has been created, additional fields display.

Figure 42-3. Pool Configuration

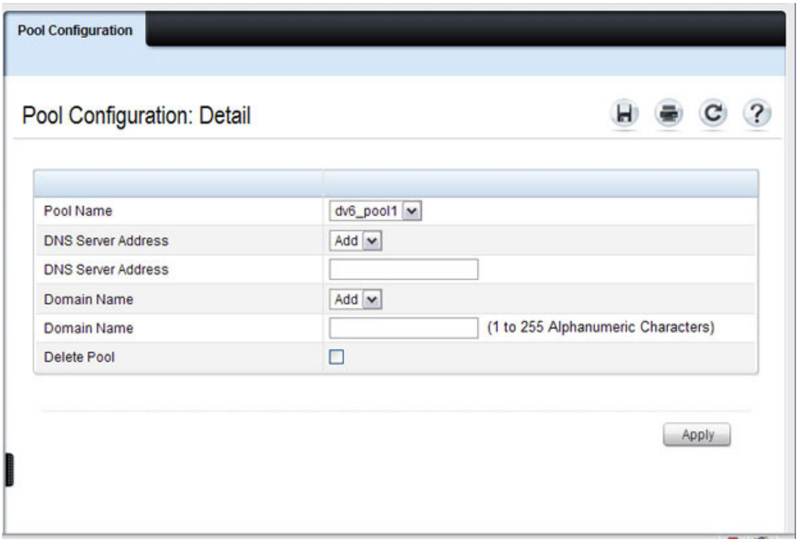


Configuring a DHCPv6 Pool

To configure the pool:

1. Open the **Pool Configuration** page.
2. Select **Create** from the **Pool Name** menu and type a name in the **Pool Name** text box.
3. Click **Apply**.

Figure 42-4. Pool Configuration



4 From the **DNS Server Address** menu, select an existing DNS Server Address to associate with this pool, or select **Add** and specify a new server to add.

5 From the **Domain Name** menu, select an existing domain name to associate with this pool, or select **Add** and specify a new domain name.

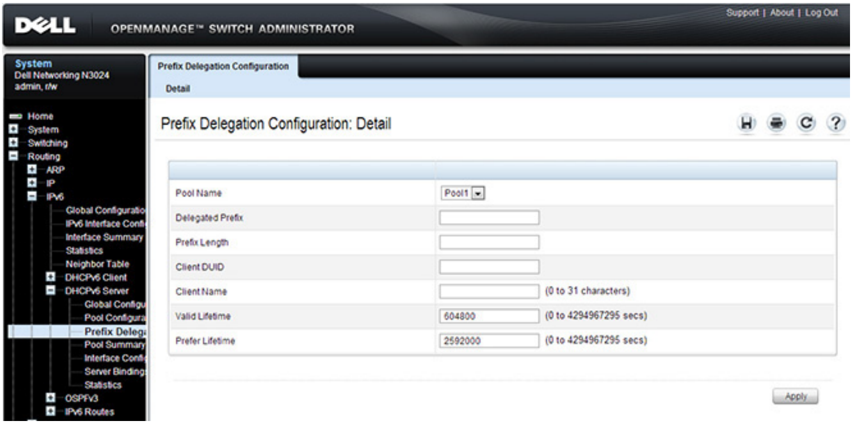
6 Click **Apply**.

Prefix Delegation Configuration

Use the **Prefix Delegation Configuration** page to configure a delegated prefix for a pool. At least one pool must be created using DHCPv6 Pool Configuration before a delegated prefix can be configured.

To display the page, click **Routing** → **IPv6** → **DHCPv6** → **Prefix Delegation Configuration** in the navigation panel.

Figure 42-5. Prefix Delegation Configuration

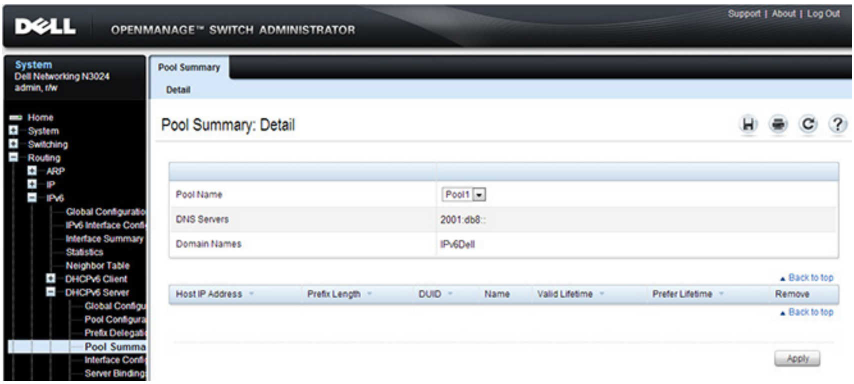


DHCPv6 Pool Summary

Use the **Pool Summary** page to display settings for all DHCPv6 Pools. At least one pool must be created using DHCPv6 Pool Configuration before the Pool Summary displays.

To display the page, click **Routing** → **IPv6** → **DHCPv6** → **Pool Summary** in the navigation panel.

Figure 42-6. Pool Summary



DHCPv6 Interface Configuration

Use the **DHCPv6 Interface Configuration** page to configure a DHCPv6 interface.

To display the page, click **Routing → IPv6 → DHCPv6 → Interface Configuration** in the navigation panel. The fields that display on the page depend on the selected interface mode.

Figure 42-7. DHCPv6 Interface Configuration

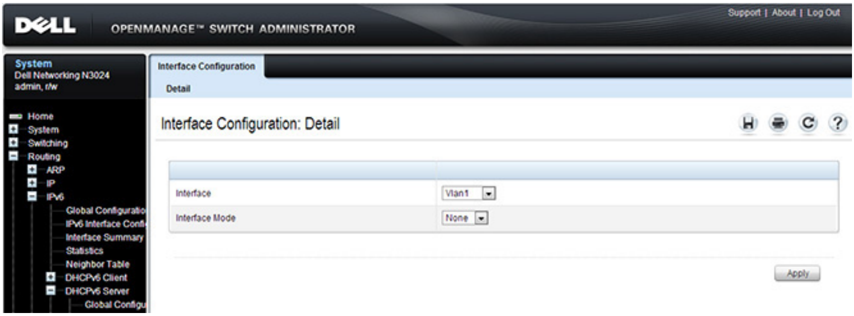


Figure 42-8 shows the screen when the selected interface mode is Server.

Figure 42-8. DHCPv6 Interface Configuration - Server Mode

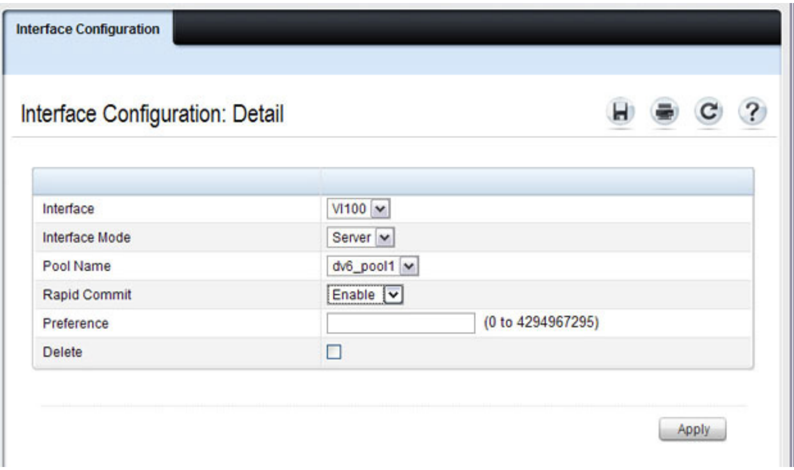
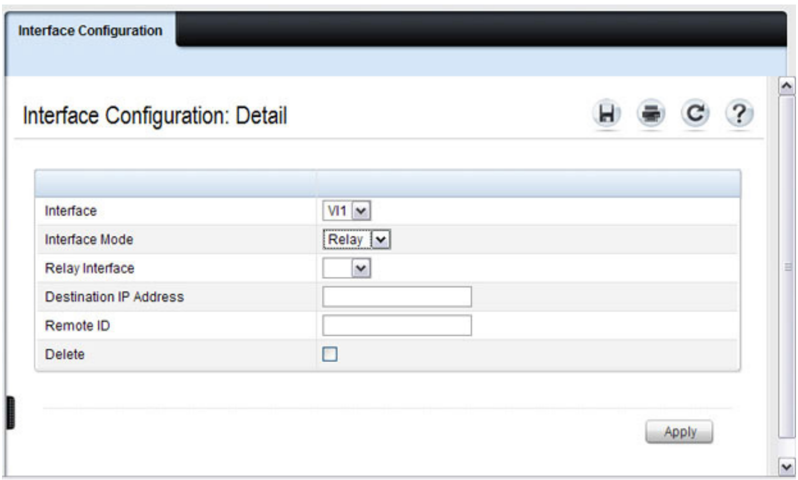


Figure 42-9 shows the screen when the selected interface mode is Relay.

Figure 42-9. DHCPv6 Interface Configuration - Relay Mode

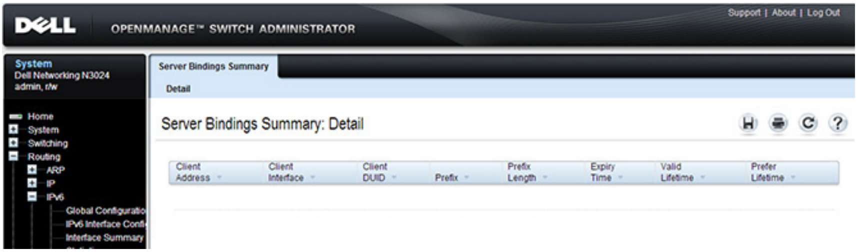


DHCPv6 Server Bindings Summary

Use the **Server Bindings Summary** page to display all DHCPv6 server bindings.

To display the page, click **Routing** → **IPv6** → **DHCPv6** → **Bindings Summary** in the navigation panel.

Figure 42-10. Server Bindings Summary

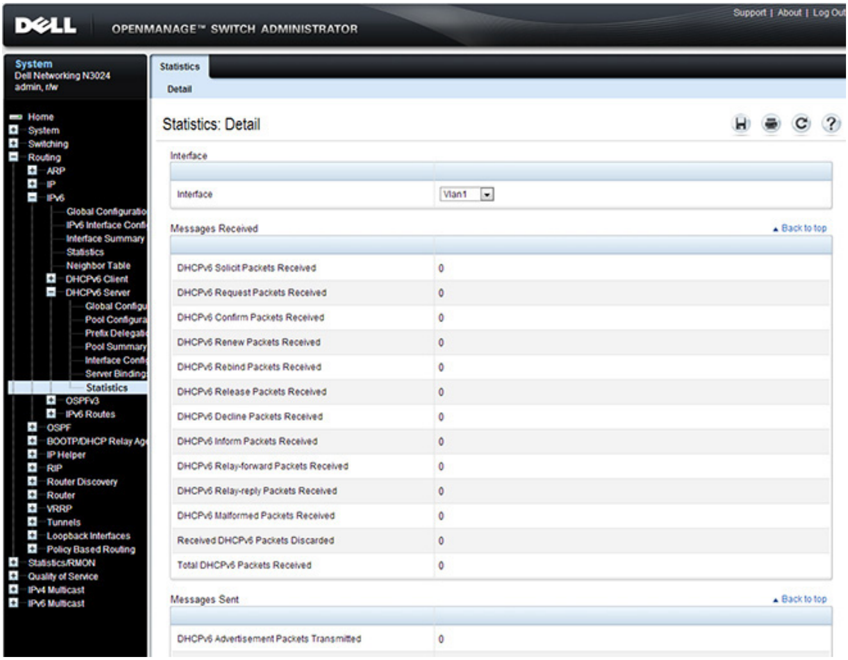


DHCPv6 Statistics

Use the **DHCPv6 Statistics** page to display DHCPv6 statistics for one or all interfaces.

To display the page, click **Routing** → **IPv6** → **DHCPv6** → **Statistics** in the navigation panel.

Figure 42-11. DHCPv6 Statistics



Configuring the DHCPv6 Server and Relay (CLI)

This section provides information about the commands used for configuring and monitoring the DHCP server and address pools. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring Global DHCP Server and Relay Agent Settings

Beginning in Privileged EXEC mode, use the following commands to configure settings for the DHCPv6 server.

Command	Purpose
configure	Enter Global Configuration mode.
service dhcpv6	Enable the DHCPv6 server.
ipv6 dhcp relay-agent-info-opt *option*	Configure a number to represent the DHCPv6 Relay Agent Information Option. The *option* parameter is an integer from 54–65535.
ipv6 dhcp relay-agent-info-remote-id-subopt *suboption*	Configure a number to represent the DHCPv6 remote-ID sub-option The *suboption* parameter is an integer from 1–65535.
exit	Exit to Privileged EXEC mode.
show ipv6 dhcp	Verify the global DHCPv6 server configuration.

Configuring a DHCPv6 Pool for Stateless Server Support

Beginning in Privileged EXEC mode, use the following commands to create a pool and configure pool parameters for DHCPv6 clients that obtain IPv6 network information dynamically.

Command	Purpose
configure	Enter Global Configuration mode.
ipv6 dhcp pool *name*	Create a DHCPv6 pool and enter DHCPv6 pool configuration mode.
dns-server *ipv6-address*	Set up to eight IPv6 DNS server addresses to provide to a DHCPv6 client by the DHCPv6 server.

Command	Purpose
domain-name *domain*	Set up to five DNS domain names to provide to a DHCPv6 client by the DHCPv6 server.
CTRL + Z	Exit to Privileged EXEC mode.
show ipv6 dhcp pool [*name*]	View the settings for all DHCPv6 pools or for the specified pool.

Configuring a DHCPv6 Pool for Specific Hosts

Beginning in Privileged EXEC mode, use the following commands to create a pool and/or configure pool parameters for specific DHCPv6 clients.

Command	Purpose
configure	Enter Global Configuration mode.
ipv6 dhcp pool *name*	Create a DHCPv6 pool and enter DHCPv6 pool configuration mode.
prefix-delegation *ipv6-prefix/prefix-length client-DUID* [**name** *hostname*] [**valid-lifetime** {*valid-lifetime* | **infinite**}] [**preferred-lifetime** {*preferred-lifetime* | **infinite**}]	Define an IPv6 prefixes within a pool for distributing to specific DHCPv6 Prefix delegation clients. • *prefix/prefix-length*—Delegated IPv6 prefix. • *client-DUID*—DHCP Unique Identifier for the client (e.g. 00:01:00:09:f8:79:4e:00:04:76:73:43:76'). • *hostname*—Client hostname used for logging and tracing. (Range: 0-31 characters.) The command allows spaces in the host name. • *valid-lifetime*—Valid lifetime for delegated prefix. (Range: 0-4294967295 seconds) or use the keyword infinite. • *preferred-lifetime*—Preferred lifetime for delegated prefix. (Range: 0-4294967295 seconds) or use the keyword infinite.
CTRL + Z	Exit to Privileged EXEC mode.
show ipv6 dhcp pool	View information about the DHCPv6 pools configured on the switch.

Configuring DHCPv6 Interface Information

Beginning in Privileged EXEC mode, use the following commands to configure an interface as a DHCPv6 server or a DHCPv6 relay agent. The server and relay functionality are mutually exclusive. In other words, a VLAN routing interface can be configured as a DHCPv6 server or a DHCPv6 relay agent, but not both.

Command	Purpose
configure	Enter Global Configuration mode.
interface {**tunnel** *tunnel-id* | **vlan** *vlan-id*}	Enter interface configuration mode for a tunnel or VLAN routing interface to configure as a DHCPv6 relay agent.
ipv6 dhcp relay {**destination** *relay-address* [**interface vlan** *vlan-id*] | **interface vlan** *vlan-id*} [**remote-id** {**duid-ifid** | *user-defined-string*}]	Configure the interface for DHCPv6 relay functionality. • **destination** — Keyword that sets the relay server IPv6 address. • *relay-address* — An IPv6 address of a DHCPv6 relay server. • **interface** — Sets the relay server interface. • *vlan-id* — A valid VLAN ID. • [remote-id {duid-ifid | *user-defined-string*}] — The Relay Agent Information Option “remote ID” sub-option to be added to relayed messages. This can either be the special keyword duid-ifid, which causes the “remote ID” to be derived from the DHCPv6 server DUID and the relay interface number, or it can be specified as a user-defined string.
exit	Exit to Global Configuration Mode
interface {**tunnel** *tunnel-id* | **vlan** *vlan-id*}	Enter interface configuration mode for a tunnel or VLAN routing interface to configure with DHCPv6 server functionality.

Command	Purpose
ipv6 dhcp server *pool-name* [**rapid-commit**] [**preference** *pref-value*]	Configure DHCPv6 server functionality on the interface. • *pool-name* — The name of the DHCPv6 pool containing stateless and/or prefix delegation parameters • **rapid-commit** — Is an option that allows for an abbreviated exchange between the client and server. • *pref-value* — Preference value—used by clients to determine preference between multiple DHCPv6 servers. (Range: 0-4294967295)
CTRL + Z	Exit to Privileged Exec Mode.
show ipv6 dhcp interface [**tunnel** *tunnel-id* | **vlan** *vlan-id*]	View DHCPv6 information for all interfaces or for the specified interface.

Monitoring DHCPv6 Information

Beginning in Privileged EXEC mode, use the following commands to view bindings, and statistics, and to clear the information.

Command	Purpose
show ipv6 dhcp binding [*address*]	View the current binding information in the DHCP server database. Specify the IP address to view a specific binding.
show ipv6 dhcp statistics	View DHCPv6 server and relay agent statistics.
clear ipv6 dhcp statistics	Reset all DHCPv6 server and relay agent statistics to zero.

DHCPv6 Configuration Examples

This section contains the following examples:

- Configuring a DHCPv6 Stateless Server
- Configuring the DHCPv6 Server for Prefix Delegation
- Configuring an Interface as a DHCPv6 Relay Agent

Configuring a DHCPv6 Stateless Server

This example configures a DHCPv6 pool that will provide information for the DHCPv6 server to distribute to DHCPv6 clients that are members of VLAN 100. To define stateless information for the DHCPv6 server to distribute, multiple DNS domain names and DNS server addresses are defined within the pool.

VLAN routing interface 100 is configured as a DHCPv6 server. Setting NDP on the interface to send the other-config-flag option allows the interface to prompt DHCPv6 clients to request only stateless server information.

To configure the switch:

1 Enable the DHCPv6 feature.

```
console#configure
console(config)#service dhcpv6
```

2 Create the DHCPv6 pool and configure stateless information.

```
console(config)#ipv6 dhcp pool my-pool
console(config-dhcp6s-pool)#domain-name pengo.dell.com
console(config-dhcp6s-pool)#domain-name dell.com
console(config-dhcp6s-pool)#dns-server 2001:DB8:A328:22C::1
console(config-dhcp6s-pool)#dns-server 2001:DB8:A328:22C::2
```

3 Configure VLAN 100 as a routing interface and assign a globally unique IPv6 address.

```
console(config)#interface vlan 100
console(config-if-vlan100)#ipv6 address
2001:DB8:A328:34B::11/32
```

4 Configure the DHCPv6 server functionality on VLAN 100. Clients can use the preference value to determine which DHCPv6 server to use when multiple servers exist.

```
console(config-if-vlan100)#ipv6 dhcp server my-pool preference
10
```

```
console(config-if-vlan100)#ipv6 nd other-config-flag
console(config-if-vlan100)#exit
```

Configuring the DHCPv6 Server for Prefix Delegation

In this example, VLAN routing interface 200 is configured to delegate specific prefixes to certain DHCPv6 clients. The prefix-to-DUID mapping is defined within the DHCPv6 pool.

To configure the switch:

1 Create the DHCPv6 pool and specify the domain name and DNS server information.

```
console(config)#ipv6 dhcp pool my-pool2
console(config-dhcp6s-pool)#domain-name dell.com
console(config-dhcp6s-pool)#dns-server 2001:DB8:A328:22C::1
```

2 Specify the prefix delegations for specific clients. The first two commands provide multiple prefixes to the same client.

```
console(config-dhcp6s-pool)#prefix-delegation
2001:DB8:1000::/32 00:01:00:09:f8:79:4e:00:04:76:73:43:76
valid-lifetime 600 preferred-lifetime 400

console(config-dhcp6s-pool)#prefix-delegation
2001:DB8:1001::/32 00:01:00:09:f8:79:4e:00:04:76:73:43:76
valid-lifetime 600 preferred-lifetime 400

console(config-dhcp6s-pool)#prefix-delegation
2001:DB8:1002::/32 00:01:00:09:f8:79:4e:00:04:76:73:43:76
valid-lifetime 600 preferred-lifetime 400

console(config-dhcp6s-pool)#exit
```

3 Configure the DHCPv6 server functionality on VLAN 200 and specify the pool to use for DHCPv6 clients.

```
console(config)#interface vlan 200
console(config-if-vlan200)#ipv6 dhcp server my-pool2
preference 20
```

Configuring an Interface as a DHCPv6 Relay Agent

This example configures a VLAN routing interface as a DHCPv6 Relay. The command defines the destination address of the relay server and the interface used for reachability to the relay server.

To configure the switch:

1 Create VLAN 300 and define its IPv6 address.

```
console(config)#interface vlan 300
console(config-if-vlan300)#ipv6 address 2001:DB8:03a::14/64
```

2 Configure the interface as a DHCPv6 relay agent and specify the IPv6 address of the relay server. The command also specifies that the route to the server is through the VLAN 100 routing interface.

```
console(config-if-vlan300)#ipv6 dhcp relay destination
FE80::250:A2FF:FEBF:A056 interface vlan 100
console(config-if-vlan300)#exit
console(config)#exit
```

3 View the DHCPv6 configuration for VLAN 300.

```
console#show ipv6 dhcp interface vlan 300

IPv6 Interface......................... ...Vl300
Mode....................................... Relay
Relay Address........... FE80::250:A2FF:FEBF:A056
Relay Interface Number....................Vl100
Relay Remote ID...........................
Option Flags..............................
```

43

Differentiated Services

Dell Networking N2000, N3000, and N4000 Series Switches

NOTE: Diffserv is not available on the Dell Networking N1500 Series switches.

This chapter describes how to configure the Differentiated Services (DiffServ) feature. DiffServ enables traffic to be classified into streams and given certain QoS treatment in accordance with defined per-hop behaviors.

The Diffserv feature is not supported on the Dell Networking N1500 Series switches.

The topics covered in this chapter include:

- DiffServ Overview
- Default DiffServ Values
- Configuring DiffServ (Web)
- Configuring DiffServ (CLI)
- DiffServ Configuration Examples

DiffServ Overview

Standard IP-based networks are designed to provide "best effort" data delivery service. Best effort service implies that the network delivers the data in a timely fashion, although there is no guarantee that it will. During times of congestion, packets may be delayed, sent sporadically, or dropped. For typical Internet applications, such as email and file transfer, a slight degradation in service is acceptable and in many cases unnoticeable. Conversely, any degradation of service has undesirable effects on applications with strict timing requirements, such as voice or multimedia.

How Does DiffServ Functionality Vary Based on the Role of the Switch?

How you configure DiffServ support in Dell Networking N2000, N3000, and N4000 Series switches software varies depending on the role of the switch in your network:

- **Edge device**: An edge device handles ingress traffic, flowing towards the core of the network, and egress traffic, flowing away from the core. An edge device segregates inbound traffic into a small set of traffic classes, and is responsible for determining a packet's classification. Classification is primarily based on the contents of the layer-3 and layer-4 headers, and is recorded in the Differentiated Services Code Point (DSCP) added to a packet's IP header.
- **Interior node**: A switch in the core of the network is responsible for forwarding packets, rather than for classifying them. It decodes the DSCP in an incoming packet, and provides buffering and forwarding services using the appropriate queue management algorithms.

Before configuring DiffServ on Dell Networking N-Series switches, you must determine the QoS requirements for the network as a whole. The requirements are expressed in terms of rules, which are used to classify inbound or outbound traffic on a particular interface.

What Are the Elements of DiffServ Configuration?

During configuration, you define DiffServ rules in terms of classes, policies, and services:

- **Class**: A class consists of a set of rules that identify which packets belong to the class. Inbound traffic is separated into traffic classes based on layer-2, layer-3, and layer-4 header data. The class type **All** is supported; this specifies that every match criterion defined for the class must be true for a match to occur.
- **Policy**: A policy defines the QoS attributes for one or more traffic classes. An attribute identifies the action taken when a packet matches a class rule. An example of an attribute is to mark a packet. The switch supports the ability to assign traffic classes to output CoS queues, and to mirror incoming packets in a traffic stream to a specific egress interface (physical port or LAG).

Dell Networking N2000, N3000, and N4000 Series switches software supports the **Traffic Conditioning Policy** type which is associated with an inbound traffic class and specifies the actions to be performed on packets meeting the class rules:

  – Marking the packet with a given DSCP, IP precedence, or CoS value. Traffic to be processed by the DiffServ feature requires an IP header if the system uses IP Precedence or IP DSCP marking.
  – Policing packets by dropping or re-marking those that exceed the class's assigned data rate.
  – Counting the traffic within the class.

- **Service:** Assigns a policy to an interface for inbound traffic.

NOTE: An 802.1X authenticator or RADIUS server can be used to dynamically assign DiffServ policy to ports when a host connects to a port and authenticates by using 802.1X. For more information, see "How Does the Authentication Server Assign DiffServ Policy? " on page 286

Default DiffServ Values

Table 43-1 shows the global default values for DiffServ.

Table 43-1. DiffServ Global Defaults

Parameter	Default Value
DiffServ	Enabled
Classes	None configured
Policies	None configured
Services	None configured

Configuring DiffServ (Web)

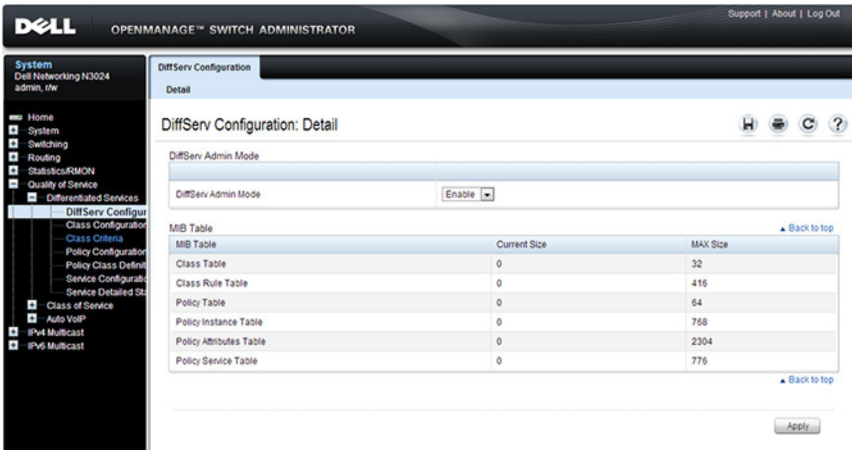
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring DiffServ features on a Dell Networking N2000, N3000, and N4000 Series switches. For details about the fields on a page, click at the top of the page.

DiffServ Configuration

Use the **DiffServ Configuration** page to display the DiffServ administrative mode setting as well as the current and maximum number of rows in each of the main DiffServ private MIB tables.

To display the page, click **Quality of Service** → **Differentiated Services** → **DiffServ Configuration** in the navigation panel.

Figure 43-1. DiffServ Configuration

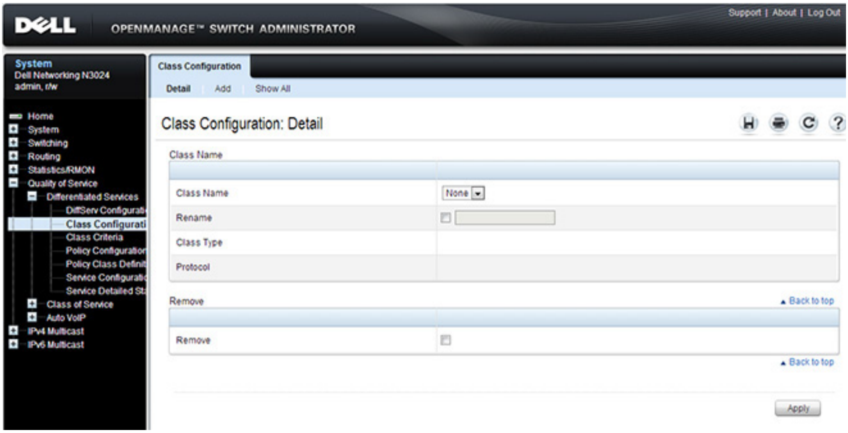


Class Configuration

Use the **DiffServ Class Configuration** page to add a new DiffServ class name, or to rename or delete an existing class.

To display the page, click **Quality of Service** → **Differentiated Services** → **Class Configuration** in the navigation panel.

Figure 43-2. DiffServ Class Configuration

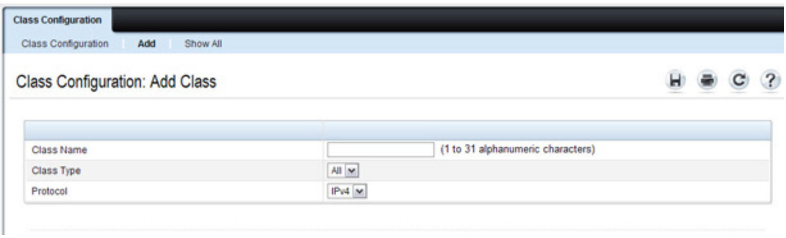


Adding a DiffServ Class

To add a DiffServ class:

1 From the **DiffServ Class Configuration** page, click **Add** to display the **Add Class** page.

Figure 43-3. Add DiffServ Class

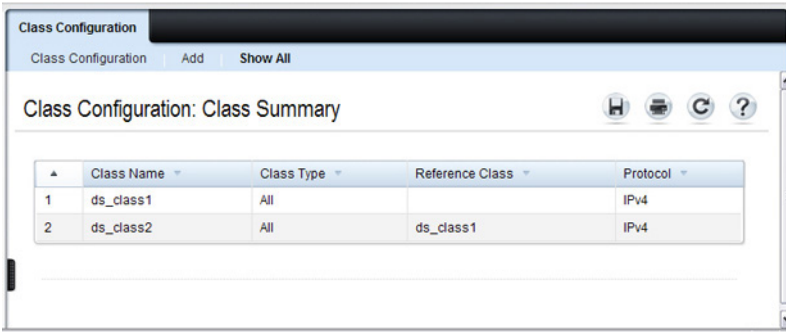


2 Enter a name for the class and select the protocol to use for class match criteria.

3 Click **Apply** to add the new class.

4 To view a summary of the classes configured on the switch, click **Show All**.

Figure 43-4. View DiffServ Class Summary

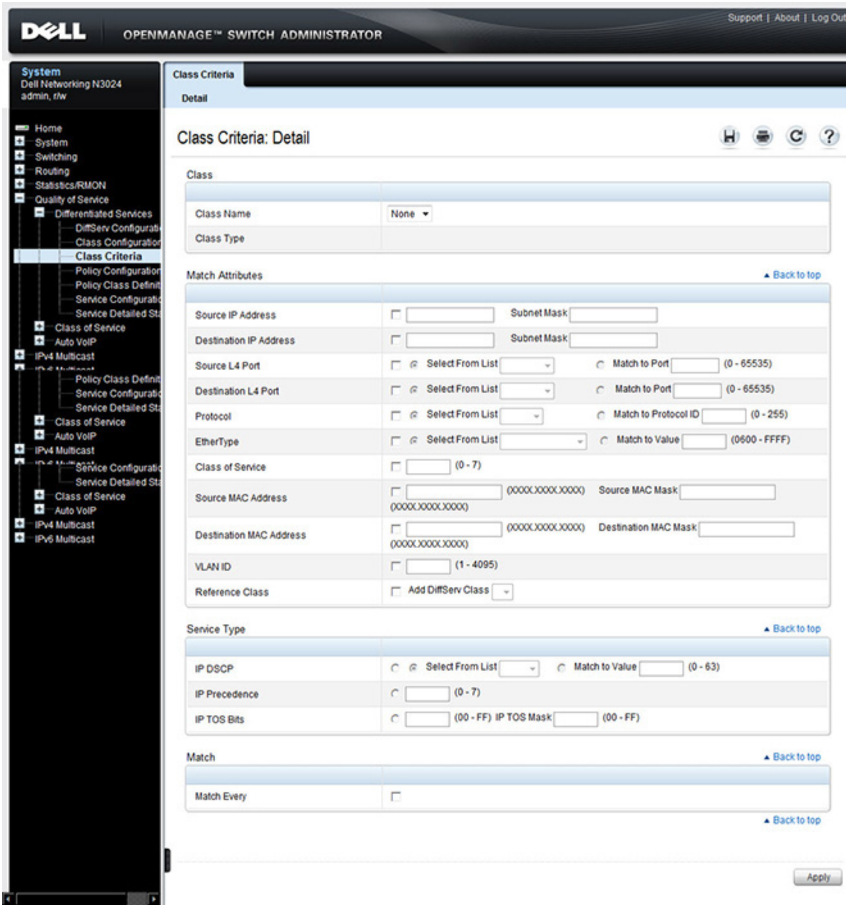


Class Criteria

Use the **DiffServ Class Criteria** page to define the criteria to associate with a DiffServ class. As packets are received, these DiffServ classes are used to identify packets.

To display the page, click **Quality of Service** → **Differentiated Services** → **Class Criteria** in the navigation panel.

Figure 43-5. DiffServ Class Criteria

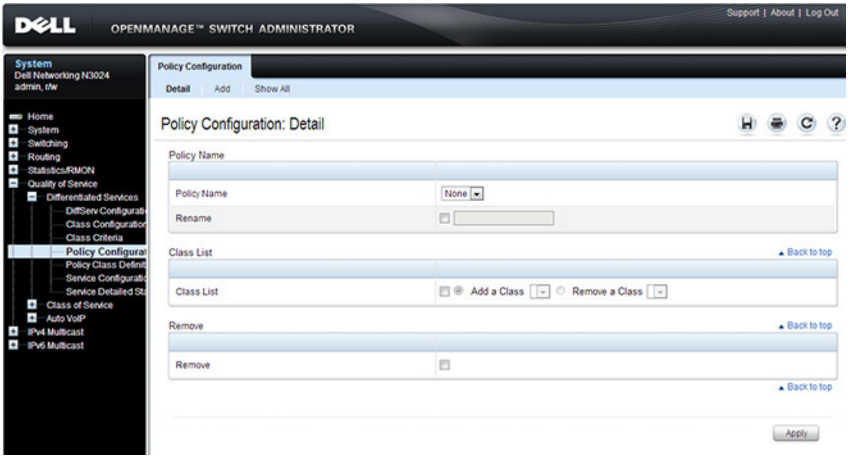


Policy Configuration

Use the **DiffServ Policy Configuration** page to associate a collection of classes with one or more policy statements.

To display the page, click **Quality of Service** → **Differentiated Services** → **Policy Configuration** in the navigation panel.

Figure 43-6. DiffServ Policy Configuration

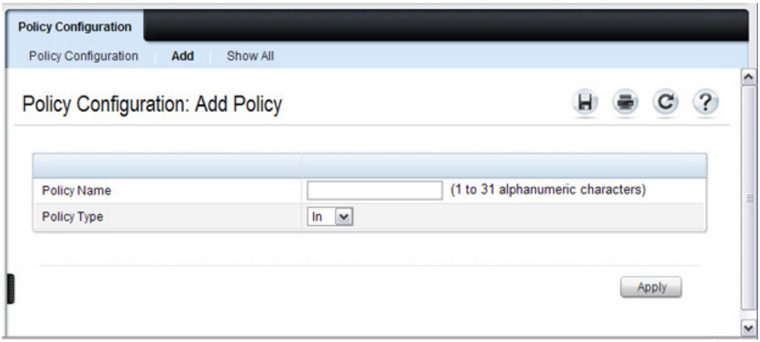


Adding a New Policy Name

To add a policy:

1 From the **DiffServ Policy Configuration** page, click **Add** to display the **Add Policy** page.

Figure 43-7. Add DiffServ Policy

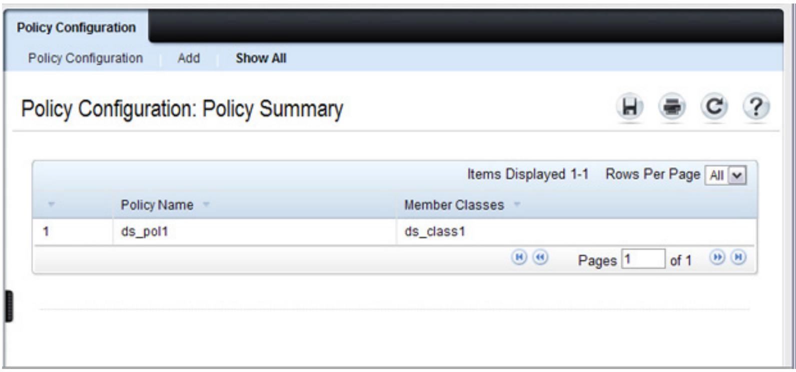


2 Enter the new **Policy Name**.

3 Click **Apply** to save the new policy.

4 To view a summary of the policies configured on the switch, click **Show All**.

Figure 43-8. View DiffServ Policies

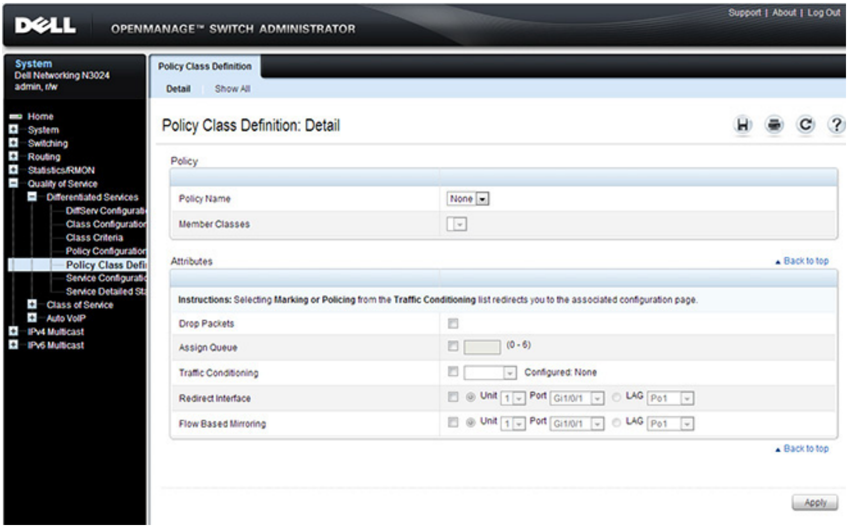


Policy Class Definition

Use the **DiffServ Policy Class Definition** page to associate a class to a policy, and to define attributes for that policy-class instance.

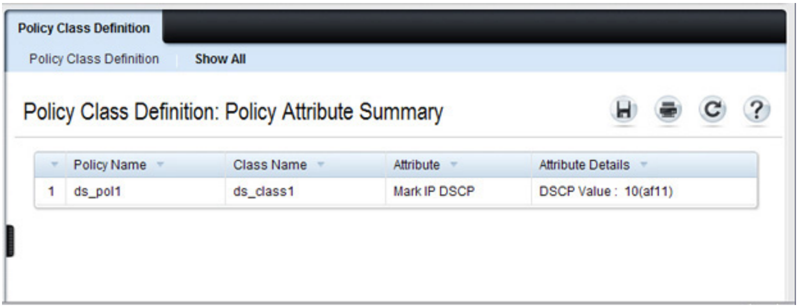
To display the page, click **Quality of Service** → **Differentiated Services** → **Policy Class Definition** in the navigation panel.

Figure 43-9. DiffServ Policy Class Definition



To view a summary of the policy attributes, click **Show All**.

Figure 43-10. Policy Class Definition



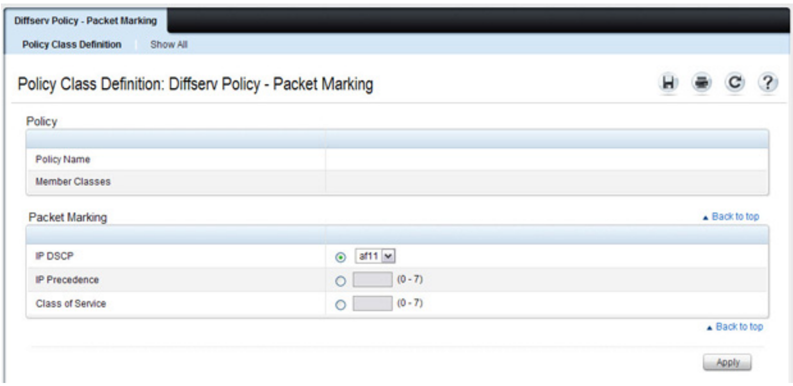
Packet Marking Traffic Condition

Follow these steps to have packets that match the class criteria for this policy marked with a marked with either an IP DSCP, IP precedence, or CoS value:

1 Select Marking from the **Traffic Conditioning** drop-down menu on the **DiffServ Policy Class Definition** page.

The **Packet Marking** page displays.

Figure 43-11. Policy Class Definition - Attributes



2 Select **IP DSCP**, **IP Precedence**, or **Class of Service** to mark for this policy-class.

3 Select or enter a value for this field.

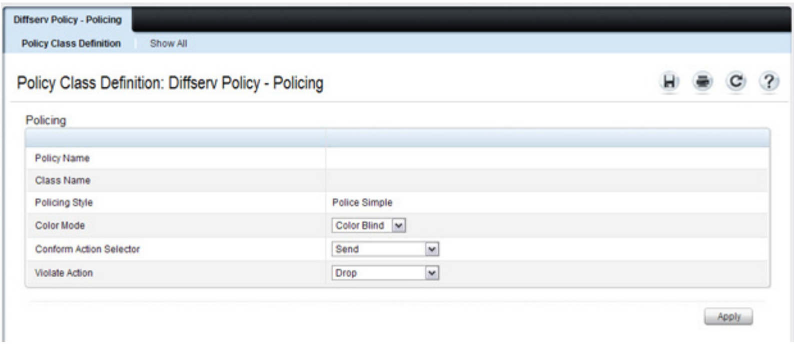
4 Click **Apply** to define the policy-class.

Policing Traffic Condition

Follow these steps to perform policing on the packets that match this policy class:

1 Select **Policing** from the **Traffic Conditioning** drop-down menu on the **DiffServ Policy Class Definition** page to display the **DiffServ Policy - Policing** page.

Figure 43-12. Policy Class Definition - Policing



The **DiffServ Policy - Policing** page displays the **Policy Name**, **Class Name**, and **Policing Style**.

Select a value for the following fields:

- **Color Mode** — The type of color policing used: Color Blind or Color Aware.
- **Conform Action Selector** — The action taken on packets that are considered conforming (below the police rate). Options are Send, Drop, Mark CoS, Mark IP DSCP, Mark IP Precedence.
- **Violate Action** — The action taken on packets that are considered non-conforming (above the police rate). Options are Send, Drop, Mark CoS, Mark IP DSCP, Mark IP Precedence.

2 Click **Apply**.

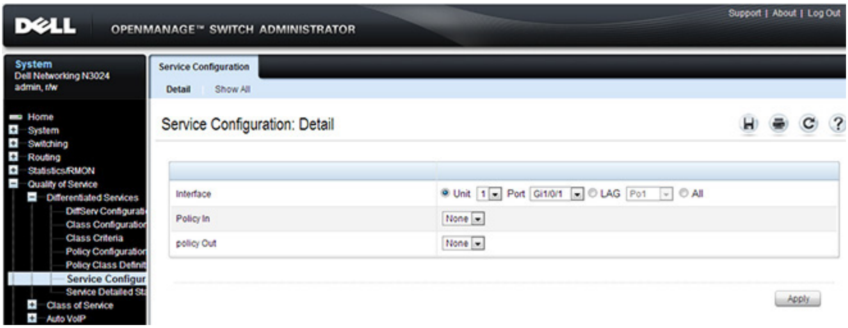
The policy-class is defined, and the device is updated.

Service Configuration

Use the **DiffServ Service Configuration** page to activate a policy on a port.

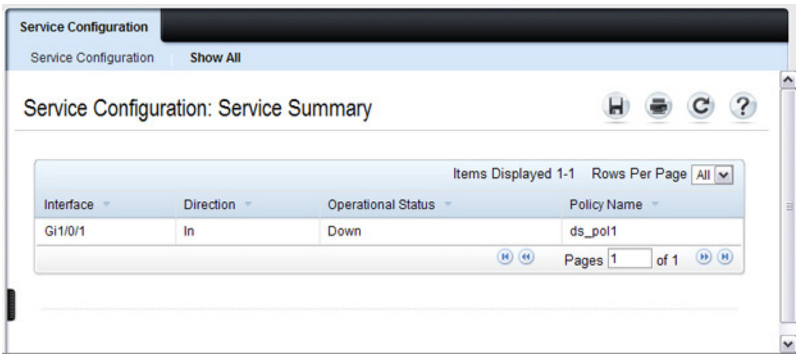
To display the page, click **Quality of Service → Differentiated Services → Service Configuration** in the navigation panel.

Figure 43-13. DiffServ Service Configuration



To view a summary of the services configured on the switch, click **Show All**.

Figure 43-14. DiffServ Service Summary

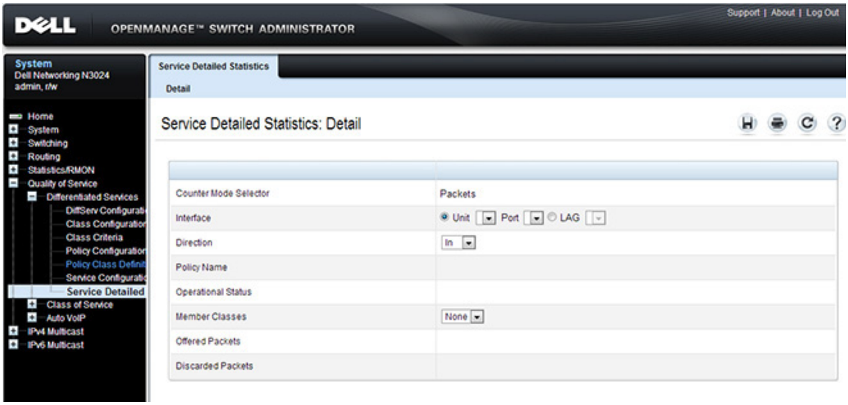


Service Detailed Statistics

Use the **DiffServ Service Detailed Statistics** page to display packet details for a particular port and class.

To display the page, click **Quality of Service** → **Differentiated Services** → **Service Detailed Statistics** in the navigation panel.

Figure 43-15. DiffServ Service Detailed Statistics

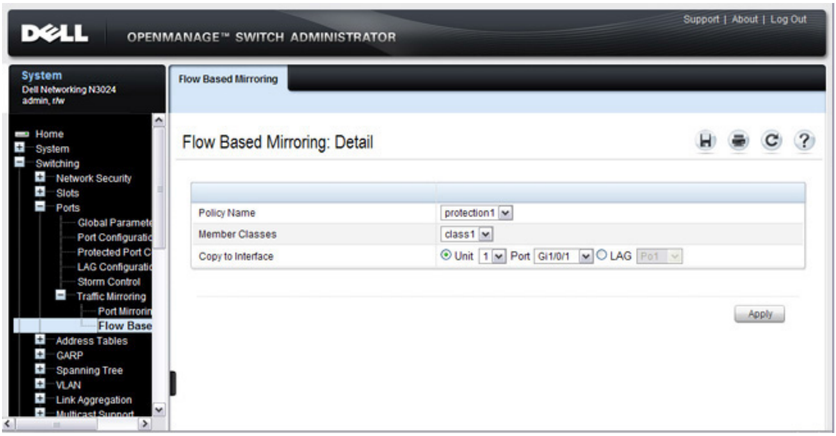


Flow-Based Mirroring

Use the **Flow-Based Mirroring** page to create a mirroring session in which the traffic that matches the specified policy and member class is mirrored to a destination port.

To display the **Flow-Based Mirroring** page, click **Switching → Ports → Traffic Mirroring → Flow-Based Mirroring** in the navigation panel.

Figure 43-16. Flow-Based Mirroring



Configuring DiffServ (CLI)

This section provides information about the commands used for configuring DiffServ settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

DiffServ Configuration (Global)

Beginning in Privileged Exec mode, use the following commands in to configure the global DiffServ mode and view related settings.

CLI Command	Description
configure	Enter global configuration mode.
diffserv	Set the DiffServ operational mode to active.
exit	Exit to Privileged EXEC mode.
show diffserv	Display the DiffServ general information, which includes the current administrative mode setting as well as the current and maximum number of DiffServ components.

DiffServ Class Configuration for IPv4

Beginning in Privileged Exec mode, use the following commands to configure DiffServ classes for IPv4 and view related information.

CLI Command	Description
configure	Enter global configuration mode.
class-map match-all *class-map-name*	Define a new DiffServ class and enter Class-Map Configuration mode for the specified class. **NOTE:** To enter Class-Map Configuration mode for a class that has already been created, use the **class-map** *class-map-name* command.
match any	Configure a match condition for all the packets.
match class-map	Add to the specified class definition the set of match conditions defined for another class.

CLI Command	Description
match cos	Add to the specified class definition a match condition for the Class of Service value.
match destination-address mac	Add to the specified class definition a match condition based on the destination MAC address of a packet.
match dstip	Add to the specified class definition a match condition based on the destination IP address of a packet.
match dstl4port	Add to the specified class definition a match condition based on the destination layer-4 port of a packet using a single keyword, or a numeric notation.
match ethertype	Add to the specified class definition a match condition based on the value of the ethertype.
match ip dscp	Add to the specified class definition a match condition based on the value of the IP DiffServ Code Point (DSCP) field in a packet.
match ip precedence	Add to the specified class definition a match condition based on the value of the IP.
match ip tos	Add to the specified class definition a match condition based on the value of the IP TOS field in a packet.
match protocol	Add to the specified class definition a match condition based on the value of the IP Protocol field in a packet using a single keyword notation or a numeric value notation.
match secondary-cos	Configure a match condition based on a secondary CoS value.
match secondary-vlan	Configure a match condition based on a secondary VLAN value.
match source-address mac	Add to the specified class definition a match condition based on the source MAC address of the packet.

CLI Command	Description
match srcip	Add to the specified class definition a match condition based on the source IP address of a packet.
match srcl4port	Add to the specified class definition a match condition based on the source layer-4 port of a packet using a single keyword, a numeric notation, or a numeric range notation.
match vlan	Add to the specified class definition a match condition based on the value of the layer-2 VLAN Identifier field.

DiffServ Class Configuration for IPv6

Beginning in Privileged Exec mode, use the following commands to configure DiffServ classes for IPv6 and view related information.

CLI Command	Description
configure	Enter global configuration mode.
class-map match-all *class-map-name* **ipv6**	Define a new DiffServ class.
match any	Configure a match condition for all the packets.
match class-map	Add to the specified class definition the set of match conditions defined for another class.
match dstip6	Add to the specified class definition a match condition based on the destination IPv6 address of a packet.
match dstl4port	Add to the specified class definition a match condition based on the destination layer-4 port of a packet using a single keyword, or a numeric notation.
match ip6flowlbl	Add to the specified class definition a match condition based on the IPv6 flow label of a packet.
match ip dscp	Add to the specified class definition a match condition based on the value of the IP DiffServ Code Point (DSCP) field in a packet.

CLI Command	Description
match protocol	Add to the specified class definition a match condition based on the value of the IP Protocol field in a packet using a single keyword notation or a numeric value notation.
match source-address mac	Add to the specified class definition a match condition based on the source MAC address of the packet.
match srcip6	Add to the specified class definition a match condition based on the source IPv6 address of a packet.
match srcl4port	Add to the specified class definition a match condition based on the source layer-4 port of a packet using a single keyword, a numeric notation, or a numeric range notation.

DiffServ Policy Creation

Beginning in Privileged Exec mode, use the following commands to configure DiffServ policies and view related information.

CLI Command	Description
configure	Enter global configuration mode.
policy-map *policy-name* **in**	Create a new DiffServ policy for ingress traffic and enter Policy Map Configuration mode for the policy.
exit	Exit to Privilege Exec mode.
show policy-map	Displays all configuration information for the specified policy.
show policy-map *interface* **in**	Displays policy-oriented statistics information for the specified interface.

DiffServ Policy Attributes Configuration

Beginning in Privilege Exec mode, use the following commands to configure policy attributes and view related information.

CLI Command	Description
configure	Enter global configuration mode.
policy-map *policy-map-name*	Enter Policy Map Configuration mode for the specified policy.
class *class-name*	Create an instance of a class definition within the specified policy for the purpose of defining treatment of the traffic class through subsequent policy attribute statements. Also enters Policy-Class-Map Configuration mode for the policy-class-map instance.
assign-queue *queue-id*	Modify the queue ID (range: 0–6) to which the associated traffic stream is assigned.
police-simple {*datarate burstsize* **conform-action** {**drop** | **set-cos-transmit** *cos* | **set-prectransmit** *cos* | **set-dscp-transmit** *dscpval* | **transmit**} [**violateaction** {**drop** | **set-cos-transmit** *cos* | **set-prec-transmit** *cos* | **set-dscp-transmit** *dscpval* | **transmit**}]}	Establish the traffic policing style for the specified class. The simple form of the police command uses a single data rate and burst size, resulting in two outcomes: conform and nonconform. • *datarate* — Data rate in kilobits per second (kbps). (Range: 1–4294967295) • *burstsize* — Burst size in Kbps (Range: 1–128) • **conform action** — Indicates what happens when the packet is conforming to the policing rule: it could be dropped, it could have its CoS modified, it could have its IP precedence modified, or it could have its DSCP modified. The same actions are available for packets that do not conform to the policing rule. • *cos* — Class of Service value. (Range: 0–7) • *dscpval* — DSCP value. (Range: 0–63 or a keyword from this list, **af11**, **af12**, **af13**, **af21**, **af22**, **af23**, **af31**, **af32**, **af33**, **af41**, **af42**, **af43**, **be**, **cs0**, **cs1**, **cs2**, **cs3**, **cs4**, **cs5**, **cs6**, **cs7**, **ef**)

CLI Command	Description
conform-color *class-map-name* [**exceed-color** *class-map-name*]	Specify the color class for color-aware policing. The action for the policy-class-map instance must be set to **police-simple** before issuing the **conform-color** command.
drop	Specify that all packets for the associated traffic stream are to be dropped at ingress.
mark cos *cos-value*	Mark all packets for the associated traffic stream with the specified class of service value (range: 0–7) in the priority field of the 802.1p header.
mark ip-dscp *dscp-value*	Mark all packets for the associated traffic stream with the specified IP DSCP value.
mark ip-precedence *value*	Mark all packets for the associated traffic stream with the specified IP precedence value (range: 0–7).
mirror *interface* | **redirect** *interface*	Use **mirror** to mirror all packets for the associated traffic stream that matches the defined class to the specified destination port or LAG. Use **redirect** to specify that all incoming packets for the associated traffic stream are redirected to the specified destination port or LAG.
exit	Exit to Policy-Map Config mode.
exit	Exit to Global Config mode.
exit	Exit to Privilege Exec mode.
show policy-map *policy-map-name*	Displays configuration information for the specified policy.

DiffServ Service Configuration

Beginning Privilege Exec mode, use the following commands to associate a policy with an interface and view related information.

CLI Command	Description
configure	Enter Global Configuration mode.
service-policy {in | out} *policy-map-name*	Attach a policy to an interface in the inbound or outbound direction. This command can be used in either Global Configuration mode (for all system interfaces) or Interface Configuration mode (for a specific interface).
exit	Exit to Privilege Exec mode.
show diffserv service brief [in | out]	Display all interfaces in the system to which a DiffServ policy has been attached.
show diffserv service interface *interface* **{in | out}**	Display policy service information for the specified interface, where *interface* is replaced by **gigabitethernet** *unit/slot/port*, **tengigabitethernet** *unit/slot/port*, or **port-channel** *port-channel number*.
show service-policy {in | out}	Display a summary of policy-oriented statistics information for all interfaces.

DiffServ Configuration Examples

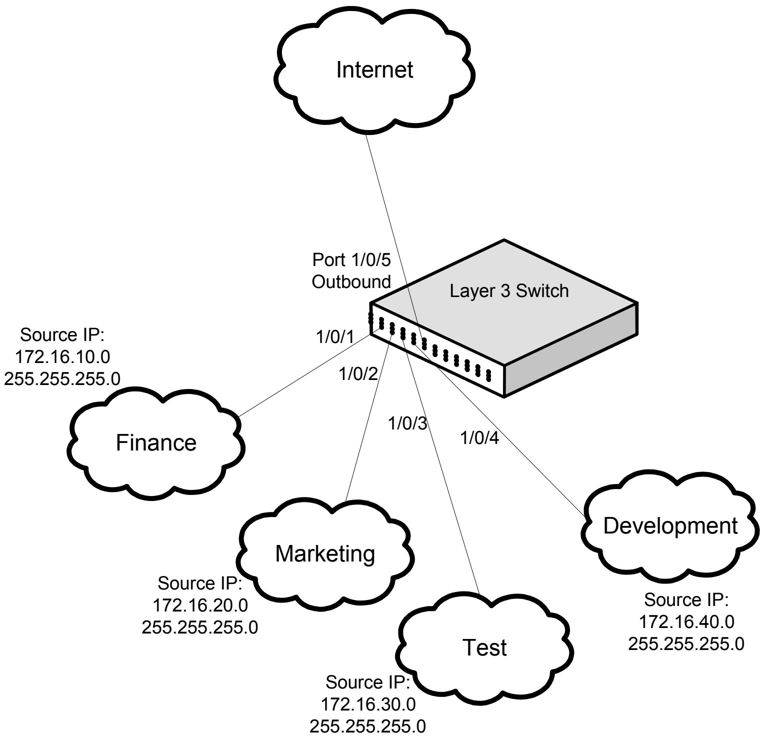
This section contains the following examples:

- Providing Subnets Equal Access to External Network
- DiffServ for VoIP

Providing Subnets Equal Access to External Network

This example shows how a network administrator can provide equal access to the Internet (or other external network) to different departments within a company. Each of four departments has its own Class B subnet that is allocated 25% of the available bandwidth on the port accessing the Internet.

Figure 43-17. DiffServ Internet Access Example Network Diagram



The following commands show how to configure the DiffServ example depicted in Figure 43-17.

1 Enable DiffServ operation for the switch.

```
console#config
console(config)#diffserv
```

2 Create a DiffServ class of type *all* for each of the departments, and name them. Also, define the match criteria—Source IP address—for the new classes.

```
console(config)#class-map match-all finance_dept
console(config-classmap)#match srcip 172.16.10.0 255.255.255.0
console(config-classmap)#exit

console(config)#class-map match-all marketing_dept
console(config-classmap)#match srcip 172.16.20.0 255.255.255.0
console(config-classmap)#exit

console(config)#class-map match-all test_dept
console(config-classmap)#match srcip 172.16.30.0 255.255.255.0
console(config-classmap)#exit

console(config)#class-map match-all development_dept
console(config-classmap)#match srcip 172.16.40.0 255.255.255.0
console(config-classmap)#exit
```

3 Create a DiffServ policy for inbound traffic named *internet_access,* adding the previously created department classes as instances within this policy. This policy uses the assign-queue attribute to put each department's traffic on a different egress queue. This is how the DiffServ inbound policy connects to the CoS queue settings established below.

```
console(config)#policy-map internet_access in
console(config-policy-map)#class finance_dept
console(config-policy-classmap)#assign-queue 1
console(config-policy-classmap)#exit

console(config-policy-map)#class marketing_dept
console(config-policy-classmap)#assign-queue 2
console(config-policy-classmap)#exit

console(config-policy-map)#class test_dept
console(config-policy-classmap)#assign-queue 3
console(config-policy-classmap)#exit
```

```
console(config-policy-map)#class development_dept
console(config-policy-classmap)#assign-queue 4
console(config-policy-classmap)#exit
console(config-policy-map)#exit
```

4 Attach the defined policy to 10-Gigabit Ethernet interfaces 1/0/1 through 1/0/4 in the inbound direction

```
console(config)#interface tengigabitethernet 1/0/1
console(config-if-Te1/0/1)#service-policy in internet_access
console(config-if-Te1/0/1)#exit

console(config)#interface tengigabitethernet 1/0/2
console(config-if-Te1/0/2)#service-policy in internet_access
console(config-if-Te1/0/2)#exit

console(config)#interface tengigabitethernet 1/0/3
console(config-if-Te1/0/3)#service-policy in internet_access
console(config-if-Te1/0/3)#exit

console(config)#interface tengigabitethernet 1/0/4
console(config-if-Te1/0/4)#service-policy in internet_access
console(config-if-Te1/0/4)#exit
```

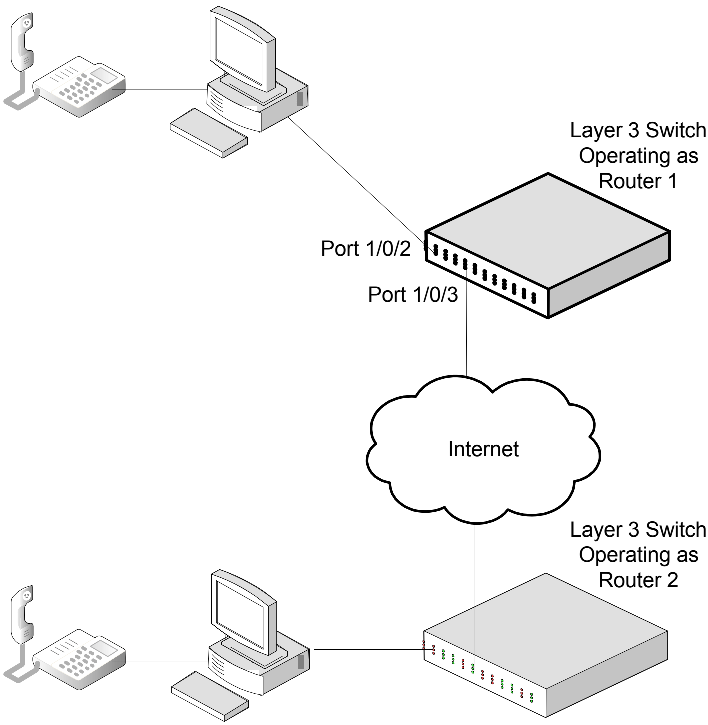
5 Set the CoS queue configuration for the (presumed) egress 10-Gigabit Ethernet interface 1/0/1 such that each of queues 1, 2, 3 and 4 get a minimum guaranteed bandwidth of 25%. All queues for this interface use weighted round robin scheduling by default. The DiffServ inbound policy designates that these queues are to be used for the departmental traffic through the assign-queue attribute. It is presumed that the switch will forward this traffic to 10-Gigabit Ethernet interface 1/0/1 based on a normal destination address lookup for internet traffic.

```
console(config)#interface tengigabitethernet 1/0/5
console(config-if-Te1/0/5)#cos-queue min-bandwidth 0 25 25 25 25 0 0
console(config-if-Te1/0/5)#exit
console(config)#exit
```

DiffServ for VoIP

One of the most valuable uses of DiffServ is to support Voice over IP (VoIP). VoIP traffic is inherently time-sensitive: for a network to provide acceptable service, a guaranteed transmission rate is vital. This example shows one way to provide the necessary quality of service: how to set up a class for UDP traffic, have that traffic marked on the inbound side, and then expedite the traffic on the outbound side. The configuration script is for Router 1 in the accompanying diagram: a similar script should be applied to Router 2.

Figure 43-18. DiffServ VoIP Example Network Diagram



The following commands show how to configure the DiffServ example depicted in Figure 43-18.

1 Set queue 6 on all ports to use strict priority mode. This queue shall be used for all VoIP packets. Activate DiffServ for the switch.

```
console#config
console(config)#cos-queue strict 6
console(config)#diffserv
```

2 Create a DiffServ classifier named *class_voip* and define a single match criterion to detect UDP packets. The class type *match-all* indicates that all match criteria defined for the class must be satisfied in order for a packet to be considered a match.

```
console(config)#class-map match-all class_voip
console(config-classmap)#match protocol udp
console(config-classmap)#exit
```

3 Create a second DiffServ classifier named class_ef and define a single match criterion to detect a DiffServ code point (DSCP) of EF (expedited forwarding). This handles incoming traffic that was previously marked as expedited elsewhere in the network.

```
console(config)#class-map match-all class_ef
console(config-classmap)#match ip dscp ef
console(config-classmap)#exit
```

4 Create a DiffServ policy for inbound traffic named pol_voip, then add the previously created classes 'class_ef' and 'class_voip' as instances within this policy. This policy handles incoming packets already marked with a DSCP value of EF (per class_ef definition), or marks UDP packets (per the class_voip definition) with a DSCP value of EF. In each case, the matching packets are assigned internally to use queue 6 of the egress port to which they are forwarded.

```
console(config)#policy-map pol_voip in
console(config-policy-map)#class class_ef
console(config-policy-classmap)#assign-queue 6
console(config-policy-classmap)#exit

console(config-policy-map)#class class_voip
console(config-policy-classmap)#mark ip-dscp ef
console(config-policy-classmap)#assign-queue 6
console(config-policy-classmap)#exit
console(config-policy-map)#exit
```

5 Attach the defined policy to an inbound service interface.

```
console(config)#interface tengigabitethernet 1/0/1
console(config-if-Te1/0/1)#service-policy in pol_voip
console(config-if-Te1/0/1)#exit
console(config)#exit
```

44

Class-of-Service

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

This chapter describes how to configure the Class-of-Service (CoS) feature. The CoS queuing feature lets you directly configure certain aspects of switch queuing. This provides the desired QoS behavior for different types of network traffic when the complexities of DiffServ are not required. The priority of a packet arriving at an interface can be used to steer the packet to the appropriate outbound CoS queue through a mapping table. CoS queue characteristics that affect queue mapping, such as minimum guaranteed bandwidth, transmission rate shaping, etc., are user-configurable at the queue (or port) level.

The topics covered in this chapter include:

- CoS Overview
- Default CoS Values
- Configuring CoS (Web)
- Configuring CoS (CLI)
- CoS Configuration Example

CoS Overview

The CoS feature lets you give preferential treatment to certain types of traffic over others. To set up this preferential treatment, configure the ingress ports, the egress ports, and individual queues on the egress ports to provide customization that suits your environment.

The level of service is determined by the egress port queue to which the traffic is assigned. When traffic is queued for transmission, the rate at which it is serviced depends on how the queue is configured and possibly the amount of traffic present in other queues for that port.

Some traffic is classified for service (i.e., packet marking) before it arrives at the switch. If you decide to use these classifications, this traffic can be mapped to egress queues by setting up a CoS Mapping table.

Each ingress port on the switch has a default priority value (set by configuring VLAN Port Priority in the Switching sub-menu) that determines the egress queue its traffic gets forwarded to. Packets that arrive without a VLAN user priority, or packets from ports you've identified as "untrusted," get forwarded according to this default.

What Are Trusted and Untrusted Port Modes?

Ports can be configured in "trusted" mode or "untrusted" mode with respect to ingress traffic.

Ports in Trusted Mode

When a port is configured in trusted mode, the system accepts at face value a priority designation encoded within packets arriving on the port. Ports can be configured to trust priority designations based on one of the following fields in the packet header:

- 802.1 Priority: values 0–7
- IP DSCP: values 0–63

A mapping table associates the designated field values in the incoming packet headers with a traffic class priority (actually a CoS traffic queue).

Ports in Untrusted Mode

If you configure an ingress port in untrusted mode, the system ignores any priority designations encoded in incoming packets, and instead sends the packets to a traffic queue based on the ingress port's default priority.

How Is Traffic Shaping Used on Egress Traffic?

For unit/slot/port interfaces, a traffic shaping rate can be specified for the port (in Kbps) for egress traffic. The traffic shaping rate specifies an upper limit of the transmission bandwidth used. Once the traffic shaping rate has been reached, frames that exceeded the limit remain queued for transmission until the next scheduling slot.

How Are Traffic Queues Defined?

For each queue, the following can be specified:

- Minimum bandwidth guarantee—A percentage of the port's maximum negotiated bandwidth reserved for the queue. Unreserved bandwidth can be utilized by lower-priority queues. If the sum of the minimum bandwidth is 100%, then there is no unreserved bandwidth and no sharing of bandwidth is possible.
- Scheduler type—strict/weighted:
  - Strict priority scheduling gives an absolute priority based on CoS queue number, with traffic in the highest numbered queue sent first, then the next lowest numbered queue, and so on. Weighted queues are serviced after all strict priority queues have been serviced.
  - Weighted scheduling selects packets for transmission with a fixed weighting equal to the CoS queue number plus one. The weighted scheduler measures bandwidth based upon bytes vs. packet counts, offering a better granularity of scheduling. For example, if CoS queues 0, 1, and 2 have an equal offered load toward a congested output port, CoS queue 2 will receive 3/6 of the bandwidth, CoS queue 1 will receive 2/6 of the bandwidth, and CoS queue 0 will receive 1/6 of the bandwidth.

The minimum bandwidth setting can be used to override the strict priority and weighted settings. The highest numbered strict priority queue will receive no more bandwidth than 100 percent minus the sum of the minimum bandwidths percentages assigned to the other queues. If used, it is recommended that minimum bandwidth percentages only be high enough to ensure a minimum level of service for any queue; i.e., the sum of the minimum bandwidth percentages is a small fraction of 100%. This ensures that the system can respond to bursts in traffic. Setting the minimum bandwidth percentages such that they sum to 100% effectively sets the scheduler such that sharing of bandwidth is disabled.

Which Queue Management Methods Are Supported?

The switch supports the following methods, configurable per-interface-queue, for determining which packets are dropped when the queue is full:

- Taildrop—Any packet forwarded to a full queue is dropped regardless of its priority.

- Weighted Random Early Detection (WRED)—Drops packets queued for transmission on an interface selectively based their drop precedence level. For each of four drop precedence levels on each WRED-enabled interface queue, the following parameters can be configured:
  – Minimum Threshold: A percentage of the interface queue size below which no packets of the selected drop precedence level are dropped.
  – Maximum Threshold: A percentage of the interface queue size above which all packets of the selected drop precedence level are dropped.
  – Drop Probability: When the queue depth is between the minimum and maximum thresholds, this value provides a scaling factor for increasing the number of packets of the selected drop precedence level that are dropped as the queue depth increases. The drop probability supports configuration in the range of 0 to 10%, and the discrete values 25%, 50%, and 75%. Values not listed are truncated to the next lower value in hardware.

The minimum and maximum WRED thresholds should be calculated to give a reasonable amount of buffering to TCP flows given the switch buffer capacity. WRED thresholds are applied individually to each physical interface. For the Dell Networking N2000/N3000 Series switches, a threshold of 100% corresponds to a buffer occupancy of 295428 bytes queued for transmission on an interface. For the Dell Networking N4000 Series switch, a threshold of 100% corresponds to a buffer occupancy of 666757 bytes queued for transmission on an interface.

NOTE: WRED is not supported on the Dell Networking N1500 Series switch.

CoS Queue Usage

CoS queue 7 is reserved by the system and is not assignable. It is generally recommended that the administrator utilize CoS queues 0 to 3, as CoS queues 4-6 may be used by the system for other types of system traffic, for example, routing protocol PDU handling.

Default CoS Values

Table 44-1 shows the global default values for CoS.

Table 44-1. CoS Global Defaults

Parameter	Default Value	
Trust Mode	802.1p User Priority	
802.1p CoS value to queue mapping	**802.1p User Priority**	**Queue**
	0, 3	1
	1, 2	0
	4, 5	2
	6, 7	3
IP DSCP value to queue mapping	**IP DSCP**	**Queue**
	0–7, 24–31	1
	8–23	0
	32–47	2
	48–63	3
Interface Shaping Rate	0 Kbps	
Minimum Bandwidth	0%	
Scheduler Type	Weighted	
Queue Management Type	Taildrop	
Drop Precedence Level	1	
WRED Decay Exponent	9	
WRED Minimum Threshold	40	
WRED Maximum Threshold	100	
WRED Drop Probability Scale	10	

Configuring CoS (Web)

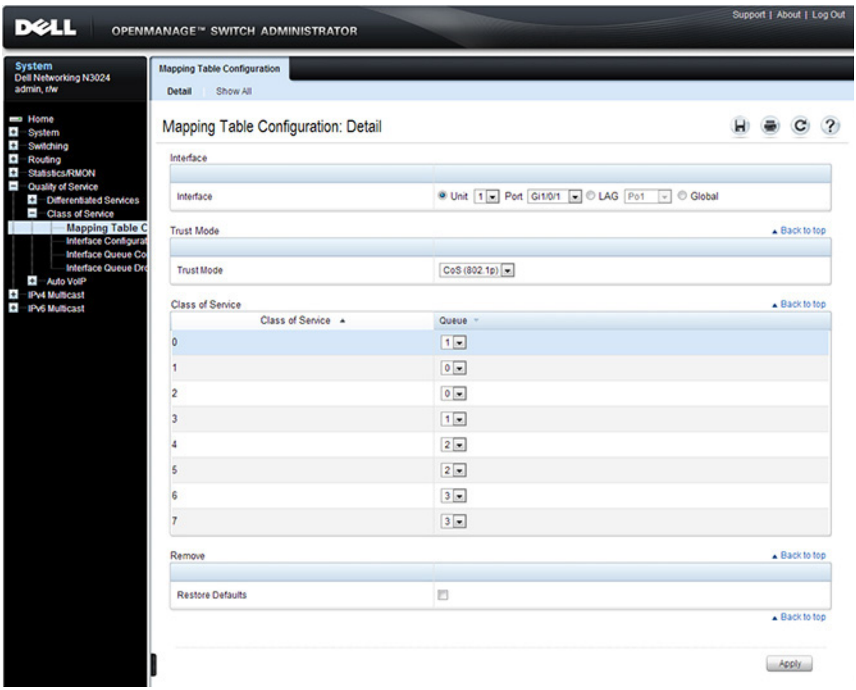
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring CoS features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

Mapping Table Configuration

Use the **Mapping Table Configuration** page to define how class of service is assigned to a packet.

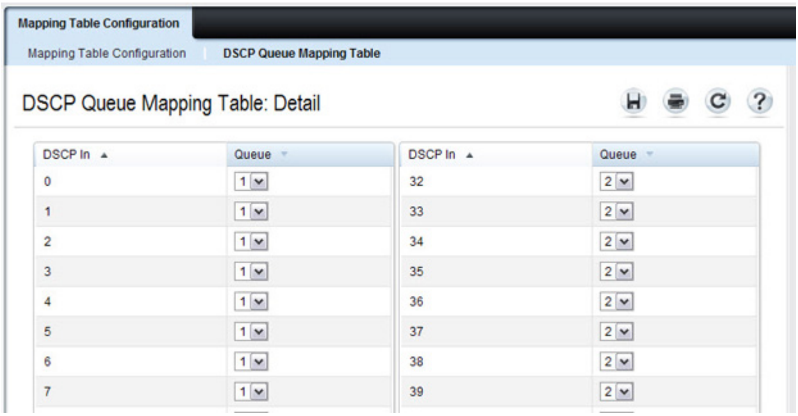
To display the page, click **Quality of Service** → **Class of Service** → **Mapping Table Configuration** in the navigation panel. CoS(802.1P) is the default mode, so this is the page that displays when **Mapping Table Configuration** is selected from the **Class of Service** menu page.

Figure 44-1. Mapping Table Configuration — CoS (802.1P)



To display the **Queue Mapping Table** for the selected Trust Mode, click the **Show All** link at the top of the page. The following figure shows the queue mapping table when CoS (802.1p) is selected as the Trust Mode.

Figure 44-2. DSCP Queue Mapping Table



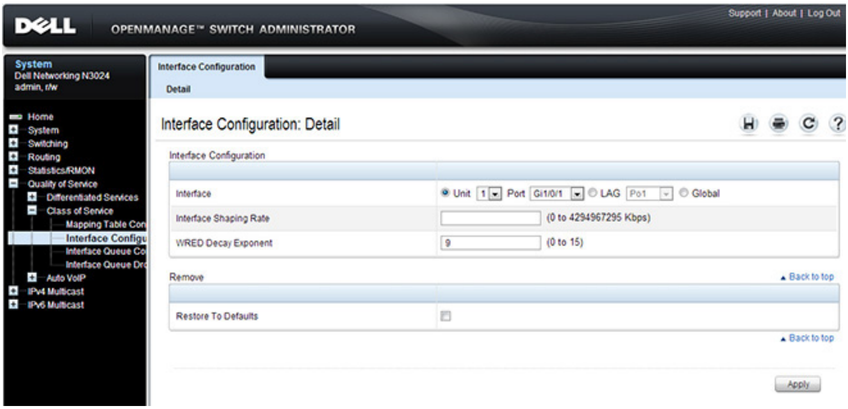
Interface Configuration

Use the **Interface Configuration** page to define the interface shaping rate for egress packets on an interface and the decay exponent for WRED queues defined on the interface.

Each interface CoS parameter can be configured globally or per-port. A global configuration change is applied to all interfaces in the system.

To display the Interface Configuration page, click **Quality of Service** → **Class of Service** → **Interface Configuration** in the navigation panel.

Figure 44-3. Interface Configuration



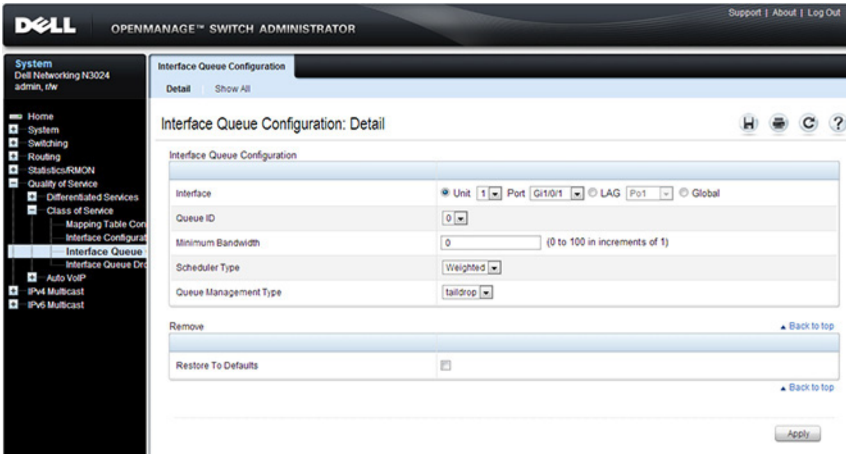
Interface Queue Configuration

Use the **Interface Queue Configuration** page to configure egress queues on interfaces. The settings you configure control the amount of bandwidth the queue uses, the scheduling method, and the queue management method.

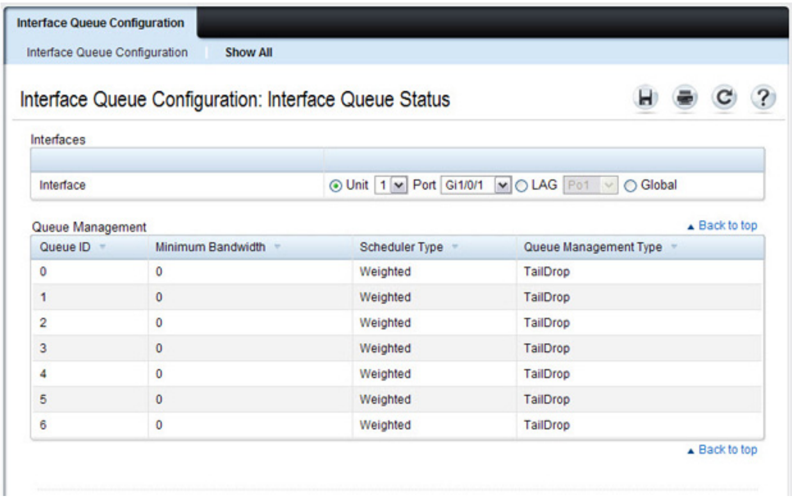
The configuration process is simplified by allowing each CoS queue parameter to be configured globally or per-port. A global configuration change is applied to the same queue ID on all ports in the system.

To display the **Interface Queue Configuration** page, click **Quality of Service** → **Class of Service** → **Interface Queue Configuration** in the navigation panel.

Figure 44-4. Interface Queue Configuration



To access the **Interface Queue Status** page, click the **Show All** link at the top of the page.



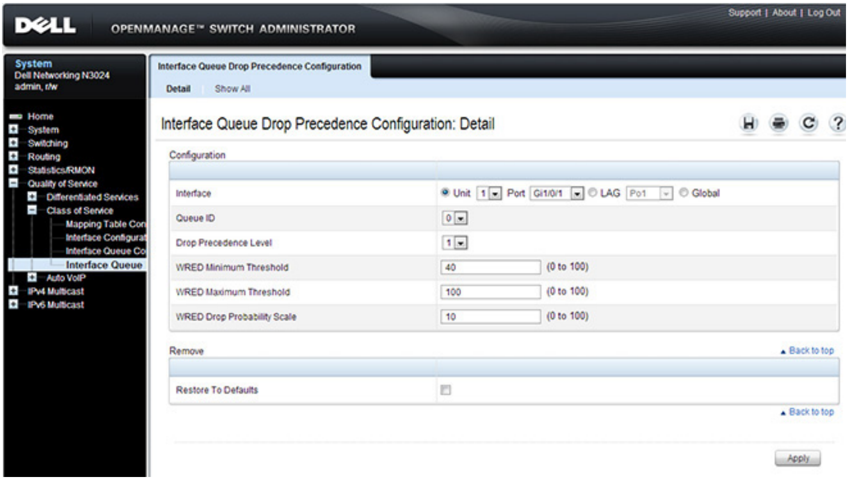
Interface Queue Drop Precedence Configuration

Use the **Interface Queue Drop Precedence Configuration** page to configure thresholds and scaling values for each of four drop precedence levels on a WRED-enabled interface queue. The settings you configure control the minimum and maximum thresholds and a drop probability scaling factor for the selected drop precedence level.

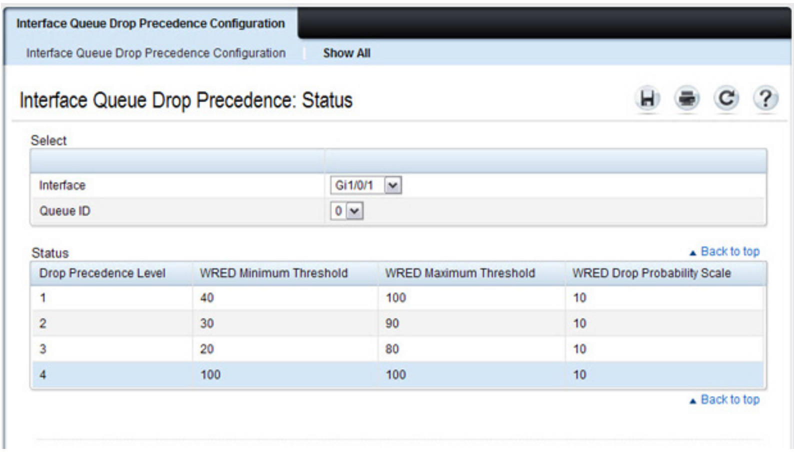
These parameters can be applied to each drop precedence level on a per-interface-queue basis, or can be set globally for the same drop precedence level and queue ID on all interfaces.

To display the **Interface Queue Drop Precedence Configuration** page, click **Quality of Service → Class of Service → Interface Queue Drop Precedence Configuration** in the navigation panel.

Figure 44-5. Interface Queue Drop Precedence Configuration



To access the **Interface Queue Drop Precedence Status** page, click the **Show All** link at the top of the page.



Drop Precedence Level	WRED Minimum Threshold	WRED Maximum Threshold	WRED Drop Probability Scale
1	40	100	10
2	30	90	10
3	20	80	10
4	100	100	10

Configuring CoS (CLI)

This section provides information about the commands used for configuring CoS settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Mapping Table Configuration

Beginning in Privileged Exec mode, use the following commands in to configure the CoS mapping tables.

CLI Command	Description
configure	Enter Global Configuration mode.
interface *interface*	Enter Interface Configuration mode, where *interface* is replaced by **gigabitethernet** *unit/slot/port*, **tengigabitethernet** *unit/slot/port*, or **port-channel** *port-channel number*.
classofservice dotlp-mapping *priority*	Map an 802.1p user priority to an internal traffic class (CoS queue) for a switch. This command can also be used in Global Configuration mode to configure the same mappings on all interfaces.
classofservice trust {dot1p | ip-dscp | untrusted}	Set the class of service trust mode of an interface.
exit	Exit to Global Config mode.
exit	Exit to Privilege Exec mode.
show classofservice dotlp-mapping	Display the current Dot1p (802.1p) priority mapping to internal traffic classes for a specific interface.
show classofservice ip-dscp-mapping	Display the current IP DSCP mapping to internal traffic classes for a specific interface.
show classofservice trust	Display the current trust mode setting for a specific interface.

CoS Interface Configuration Commands

Beginning in Privileged Exec mode, use the following commands in to configure the traffic shaping and WRED exponent values for an interface.

CLI Command	Description
configure	Enter Global Configuration mode.
interface *interface*	Enter Interface Configuration mode, where *interface* is replaced by **gigabitethernet** *unit/slot/port*, **tengigabitethernet** *unit/slot/port*, or **port-channel** *port-channel number*.
traffic-shape *bw* **kbps**	Sets the upper limit on how much traffic can leave a port. The *bw* variable represents the shaping bandwidth value from 64 to 4294967295 kbps.
random-detect exponential-weighting-constant *exponent*	Configure the WRED decay exponent (range: 0–15) for the interface. The weighting constant exponent determines how much of the previous average queue length sample is added to the current average queue length. A value of 0 indicates that no weight is given to the previous sample and only the instantaneous rate is used. A value of 1 indicates that 1/2 of the difference between the instantaneous value and the previous value is added to the current value; a value of 2 implies that 1/4 of the difference is added, 3 implies 1/8 of the difference is added, etc.

Interface Queue Configuration

Beginning in Privileged Exec mode, use the following commands in to configure and view CoS interface queue settings.

CLI Command	Description
configure	Enter Global Configuration mode.
interface *interface*	Enter Interface Configuration mode, where *interface* is replaced by **gigabitethernet** *unit/slot/port*, **tengigabitethernet** *unit/slot/port.*, or **port-channel** *port-channel number.*

CLI Command	Description
cos-queue min-bandwidth *bw*	Specify the minimum transmission bandwidth (range: 0-100% in 1% increments) for each interface queue. The sum of the configured minimum bandwidths should be less than 100% to allow for buffering of bursty traffic.
cos-queue strict *queue-id*	Activate the strict priority scheduler mode for each specified queue. The *queue-id* value ranges from 0 to 6.
cos-queue random-detect *queue-id*	Set the queue management type for the specified queue to WRED. The **no** version of this command resets the value to taildrop.
exit	Exit to Global Config mode.
exit	Exit to Privilege Exec mode.
show interfaces cos-queue	Display the class-of-service queue configuration for a specified interface or all interfaces.

Configuring Interface Queue Drop Probability

Beginning in Privileged Exec mode, use the following commands in to configure characteristics of the drop probability and view related settings. The drop probability supports configuration in the range of 0 to 10%, and the discrete values 25%, 50%, and 75%. Values not listed are truncated to the next lower value in hardware.

CLI Command	Description
configure	Enter Global Configuration mode.
interface *interface*	Enter Interface Configuration mode, where *interface* is replaced by **gigabitethernet** *unit/slot/port*, **tengigabitethernet** *unit/slot/port*, or **port-channel** *port-channel number*.
random-detect queue-parms *queue-id* [*queue-id*...] **min-thresh** *min1 min2 min3 min4* **max-thresh** *max1 max2 max3 max4* **drop-prob** *prob1 prob2 prob3 prob4*	Configure the maximum and minimum thresholds for one or more queue IDs on a WRED-enabled interface queue. This command can also be used in Global Configuration mode to configure the same parameters for one or more queues on all interfaces.
exit	Exit to Global Config mode.
exit	Exit to Privilege Exec mode.
show interfaces random-detect	Display WRED parameters for an interface or all interfaces.

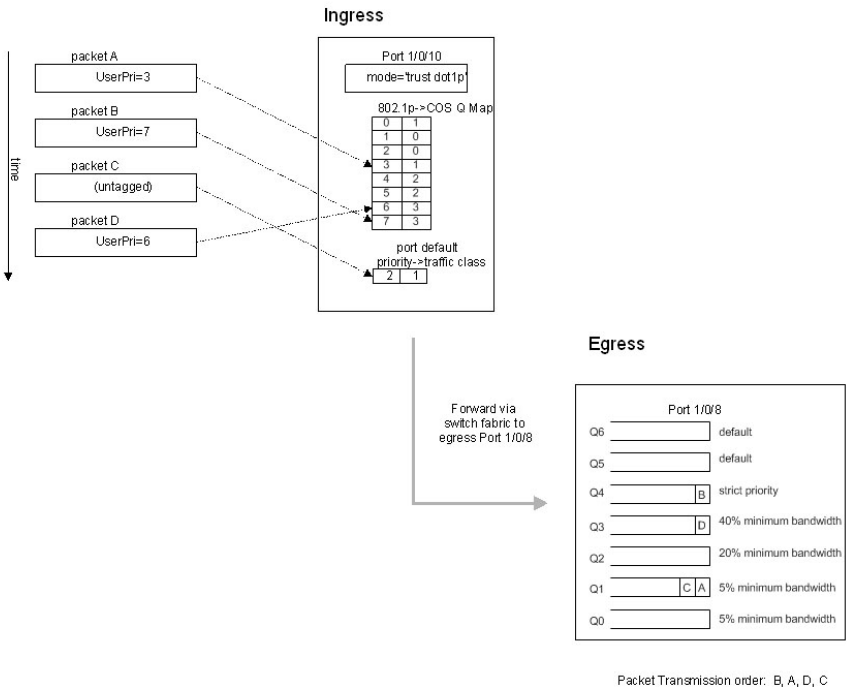
CoS Configuration Example

Figure 44-6 illustrates the network operation as it relates to CoS mapping and queue configuration.

Four packets arrive at the ingress port te1/0/10 in the order A, B, C, and D. port te1/0/10 is configured to trust the 802.1p field of the packet, which serves to direct packets A, B, and D to their respective queues on the egress port. These three packets utilize the 802.1p to CoS Mapping Table for port te1/0/10.

In this example, the 802.1p user priority 3 is configured to send the packet to queue 3 instead of the default queue 2. Since packet C does not contain a VLAN tag, the 802.1p user priority does not exist, so port te1/0/10 relies on its default port priority (2) to direct packet C to egress queue 0.

Figure 44-6. CoS Mapping and Queue Configuration



Continuing this example, the egress port te1/0/8 is configured for strict priority on queue 4, and a weighted scheduling scheme is configured for queues 3-0. Assuming queue 3 has a higher minimum bandwidth than queue 1 (relative bandwidth values are shown as a percentage, with 0% indicating the bandwidth is shared according to the default weighting), the queue service order, when congested, is 4 followed by 3 followed by 1. Assuming each queue transmits all packets shown in the diagram, the packet transmission order as seen on the network out of port te1/0/8 is B, D, A, C. Thus, packet B, with its strict priority scheduling, is transmitted ahead of the other packets at the egress port.

The following commands configure port 10 (ingress interface) and port 8 (egress interface).

1 Configure the Trust mode for port 10.

```
console#config
console(config)#interface tengigabitethernet 1/0/10
console(config-if-Te1/0/10)#classofservice trust dot1p
```

2 For port 10, configure the 802.1p user priority 7 to send the packet to queue 4 instead of the default queue (queue 3).

```
console(config-if-Te1/0/10)#classofservice dot1p-mapping 7 4
```

3 For port 10, specify that untagged VLAN packets should have a default priority of 2, which maps to queue 0.

```
console(config-if-Te1/0/10)#vlan priority 2
console(config-if-Te1/0/10)#exit
```

4 For port 8, the egress port, configure a weighted scheduling scheme for queues 3–0.

```
console(config)#interface tengigabitethernet 1/0/8
console(config-if-Te1/0/8)#cos-queue min-bandwidth 5 5 20 40 0
0 0
```

5 Configure port 8 to have strict priority on queue 4:

```
console(config-if-Te1/0/8)#cos-queue strict 4
```

To configure the CoS queues for lossless traffic when transporting iSCSI traffic, set the lossless traffic class to have a one-to-one mapping with the priority value. The following example illustrates how to change the dot1p mapping from the switch defaults to support lossless[1] transport of frames on CoS queue 4, with a 50% minimum bandwidth guarantee. Lossless traffic

1. Lossless behavior is guaranteed only when configured in conjunction with a congestion control mechanism such as PFC.

classes generally use the default WRR scheduling mode as opposed to strict priority, to avoid starving other traffic. For example, the following commands assign 802.1p user priority 4 to CoS queue 4 and reserves 50% of the scheduler time slices to CoS queue 4. This implies that, when the switch is congested, the scheduler will service CoS queue 4 fifty percent of the time to the exclusion of all other CoS queues, including higher-priority CoS queues.

```
classofservice dot1p-mapping 4 4
cos-queue min-bandwidth 0 0 0 0 50 0 0
```

WRED

NOTE: WRED is not supported on the Dell Networking N1500 Series switch.

WRED Processing

Traffic ingressing the switch can be assigned to one of four drop probabilities based on a set of matching criteria. There are three drop probabilities for TCP traffic (green, yellow, and red) and one drop probability for non-TCP traffic (all colors). Users may configure the congestion thresholds at which packets queued for transmission are dropped for each color.

WRED is intended to provide early feedback to protocols (e.g., TCP) that depend on packet drop to adjust their transmission rate. WRED packet drops only occur when the system is congested within the ranges specified. If congestion exceeds the upper limit, packets will be dropped at the rate of traffic ingressing the system, e.g., 100%. If the congestion is less than the lower limit, no packets will be dropped.

WRED Drop Probabilities

Between the minimum and maximum thresholds, the drop probability is divided into eight discrete levels of increasing probability of packet drop. The levels are as follows:

0 – 6.25% of maximum drop probability

1 – 18.75% of maximum drop probability

2 – 30.25% of maximum drop probability

3 – 43.75% of maximum drop probability

4 – 56.25% of maximum drop probability

5 – 68.75% of maximum drop probability

6 – 81.25% of maximum drop probability

7 – 92.75% of maximum drop probability

As an example, with a drop probability of 50%, a minimum threshold of 10% and a maximum threshold of 90%, the drop probability from 10% to 20% congestion is 3.125%, from 21% to 30% congestion is 9.375%.

Exponential Weighting Constant

The degree of congestion is determined by sampling the egress queue depth and calculating an average queue size. The exponential weighting constant smooths the result of the average queue depth calculation by the function:

*average depth = (previous queue depth * (1-1/2 ^n)) + (current queue depth * 1/2 ^n)*

The average queue depth is used to select the drop probability for packets queued for egress. Because the instantaneous queue depth fluctuates rapidly, larger values of the weighting constant will cause the average queue depth value to respond to changes more slowly than smaller values.

WRED Color-Aware Processing

Packets may be assigned to different colors using an ingress policing policy. Each color has a different profile of WRED drop probabilities. This capability allows the operator to configure different WRED drop policies based on the incoming packet's CoS, secondary CoS, IP DSCP, or IP precedence values.

To assign a color to incoming packets, define a class map with the desired matching criteria, assign the class map to a policy map, and set a metering rule in the policy map. Dell Networking switches implement three policing meters: a simple two-color meter, a single-rate three-color meter, and a two-rate three-color meter.

Each of these may be configured in color-aware or color-blind mode.

The **conform-color** command is used to enable color-aware mode. If the **conform-color** command is not used, packets are processed in color-blind mode. Color-blind mode means that all packets are initially colored green and are assigned their final color by the meter. Color-aware mode means that packets are pre-colored prior to entering the meter that assigns the final packet color. Packets that match the conform-color class are pre-colored green.

Packets that match the exceed color class are pre-colored yellow. Packets that do not match the conform color or exceed color class are pre-colored red, that is, they will always have a final color of red and will be processed accordingly.

- Packets that are pre-colored green and exceed the CIR will be colored yellow. Those that exceed the PIR will be colored red.

- Packets that are pre-colored yellow and exceed the PIR will be colored red. This does not apply to the simple algorithm since there is no yellow pre-coloring.
- Packets that are pre-colored red remain colored red.

Refer to RFC 2697 and RFC 2698 for further detail on color-aware and color-blind processing.

Simple Meter Implementation

The simple algorithm meters a traffic stream and colors packets red or green according to two parameters, the Committed Information Rate (CIR) and the Committed Burst Size (CBS). If the CIR is violated, the offending packets are colored red, otherwise, they are colored green.

In color-aware mode, packets may also be pre-colored red based on user-defined criteria. Packets that are pre-colored red are considered for discard as if the simple rate meter had colored them red as a result of exceeding the meter. Pre-colored packets may not be re-colored to green by the meter.

It is recommended that the CBS parameter be set to the largest IP packet size that can be legally transported.

Single Rate Meter Implementation

The police-one-rate algorithm implements a single-rate Three Color Marker (srTCM) per RFC 2697. The srTCM algorithm is useful in situations where the length of the burst is the distinguishing factor for determining service eligibility and the peak rate is not considered. A srTCM meters a traffic stream per the Committed Burst Size (CBS) and Excess Burst Size (EBS) parameters and colors packets according to the Committed Information Rate (CIR), the CBS and the EBS. The PIR must be greater than or equal to the CIR. At least one of the CBS or EBS must be greater than 0. It is recommended that when the CBS or EBS parameters are larger than 0, they be configured to be greater than or equal to the largest IP packet size that can be legally transported. In color-blind mode, a packet is colored green if it does not exceed the CBS, yellows if it exceeds the CBS but not the EBS, and red if it exceeds both.

In color-aware mode, packets may be pre-colored yellow or red based on user-defined criteria prior to being processed by the meter. Packets that are pre-colored yellow or red are considered for discard as if the meter had colored them as a result of exceeding the meter. Pre-colored packets are not re-colored to green or yellow by the meter. Yellow packets may be colored red as a result of exceeding the meter.

Refer to RFC 2697 for further details.

Two-Rate Meter Implementation

The police-two-rate algorithm implements a two-rate Three-Color Marker (trTCM) per RFC 2698. The trTCM algorithm is useful in situations where a peak rate needs to be enforced separately from a committed rate. A trTCM meters a traffic stream per the Committed Burst Size (CBS) and Peak Burst Size (PBS) parameters and colors packets according to the Committed Information Rate (CIR) and the Peak Information Rate (PIR). The PIR must be greater than or equal to the CIR. The PBS must be greater than or equal to the CBS. It is recommended that the CBS parameter be set to the largest IP packet size that can be legally transported. In color-blind mode, a packet is colored red if it exceeds the PIR, yellow if it exceeds the CIR, and green if it exceeds neither.

In color-aware mode, packets may be pre-colored yellow or red based on user-defined criteria prior to being processed by the meter. Packets that are pre-colored yellow or red are considered for discard as if the meter had colored them as a result of exceeding the meter. Pre-colored packets are not re-colored to green or yellow by the meter. Yellow packets may be colored red as a result of exceeding the meter.

Refer to RFC 2698 for further details.

Explicit Congestion Notification

Explicit Congestion Notification (ECN) is defined in RFC 3168. Conventional TCP networks signal congestion by dropping packets. A Random Early Discard scheme provides earlier notification than tail drop. ECN marks congested packets that would otherwise have been dropped and expects a ECN capable receiver to signal congestion back to the transmitter without the need to retransmit the packet that would have been dropped. For TCP, this means that the TCP receiver signals a reduced window size to the transmitter but does not request retransmission of the CE marked packet.

ECN uses the two least significant bits of Diffserv field (TOS octet in IPv4 / Traffic Class octet in IPv6) and codes them as follows:

00: Non ECN-Capable Transport – Non-ECT

10: ECN Capable Transport – ECT(0)

01: ECN Capable Transport – ECT(1)

11: Congestion Encountered – CE

ECN capable hosts communicate support for ECN via two flags in the TCP header:

- ECN-Echo (ECE)
- Congestion Window Reduced (CWR)

Dell Networking WRED considers packets for early discard only when the number of packets queued for transmission on a port exceeds the relevant minimum WRED threshold. Four thresholds are available for configuration. The green, yellow, and red thresholds operate on TCP packets. The fourth threshold operates on non-TCP packets.

When ECN is enabled and congestion is experienced, packets that are marked ECN-capable, are queued for transmission, and are randomly selected for discard by WRED are instead marked CE and are transmitted rather than dropped. This includes packets that exceed the WRED upper threshold. If the switch experiences severe congestion (no buffers available), then packets are discarded.

Dell Networking implements ECN capability as part of the WRED configuration process. Eligible packets are marked by hardware based upon the WRED configuration. The network operator can configure any CoS queue to operate in ECN marking mode and can configure different discard thresholds for each color.

Enabling ECN in Microsoft Windows

On many current Windows implementations, ECN capability is enabled via the **netsh** command as follows:

```
netsh int tcp set global ecncapability=enabled
```

The capability can be verified with the command **netsh int tcp show global**.

An example is shown below:

```
C:\Users\jmclendo>Netsh int tcp set global ecncapability=enabled
Ok.
C:\Users\jmclendo>netsh int tcp show global
Querying active state...

TCP Global Parameters
----------------------------------------------
Receive-Side Scaling State          : enabled
Chimney Offload State               : automatic
NetDMA State                        : enabled
Direct Cache Access (DCA)           : disabled
Receive Window Auto-Tuning Level    : normal
Add-On Congestion Control Provider  : none
ECN Capability                      : enabled
RFC 1323 Timestamps                 : disabled
```

In Windows Server 2012, DCTCP is self-activating based on the RTT of TCP packets. No user management is required. Use the PowerShell cmdlet Get-NetTcpConnection to verify DCTCP operation.

Example 1: SLA Configuration

The following example configures a simple meter and a trTCM meter is support of a network SLA. The SLA classes are segregated by CoS class as described in the comments.

1 Define a class-map so that all traffic will be in the set of traffic "cos-any".

```
console#config
console(config)#class-map match-all cos-any ipv4
console(config-classmap)#match any
console(config-classmap)#exit
```

2 Define a class-map such that all traffic with a CoS value of 1 will be in the set of traffic "cos1." We will use this as a conform color class map. Conform-color class maps must be one of cos, secondary cos, dscp, or ip precedence.

```
console(config)#class-map match-all cos1 ipv4
console(config-classmap)#match cos 1
console(config-classmap)#exit
```

3 Define a class-map such that all IPv4 traffic with a CoS value of 0 will be in the set of traffic "cos0." We will use this as a conform-color class map. Conform-color class maps must be one of cos, secondary cos, dscp, or ip precedence.

```
console(config)#class-map match-all cos0 ipv4
console(config-classmap)#match cos 0
console(config-classmap)#exit
```

4 Define a class-map such that all TCP will be in the set of traffic "TCP." We will use this as a base color class for metering traffic.

```
console(config)#class-map match-all tcp ipv4
console(config-classmap)#match protocol tcp
console(config-classmap)#exit
```

5 Define a policy-map to include packets matching class cos-any (IPv4). Ingress IPv4 traffic arriving at a port participating in this policy will be assigned red or green coloring based on the metering.

```
console(config)#policy-map simple-policy in
console(config-policy-map)#class cos-any
```

6 Create a simple policer in color blind mode. Packets below the committed information rate (CIR) or committed burst size (CBS) are assigned drop precedence green. Packets that exceed the CIR (in Kbps) or CBS (in Kbytes) are colored red. Both the conform and violate actions are set to transmit as WRED is used to drop packets when congested.

```
console(config-policy-classmap)#police-simple 10000000 64
conform-action transmit violate-action transmit
console(config-policy-classmap)#exit
console(config-policy-map)#exit
```

7 Define a policy-map in color aware mode matching class cos-any (IPv4). Ingress IPv4 traffic arriving at a port participating in this policy will be assigned green, yellow, or red coloring based on the meter.

```
console(config)#policy-map two-rate-policy in
console(config-policy-map)#class tcp
```

8 Create a two-rate policer per RFC 2698. The CIR value is 800 Kbps and the CBS is set to 96 Kbytes. The PIR is set to 950 Kbps and the PBS is set to 128 Kbytes.

Color-aware processing is enabled via the **conform-color** command; i.e., any packets not in CoS 0 or CoS 1 are pre-colored "red." Packets in CoS 0 are pre-colored yellow. Packets in CoS 1 are pre-colored green. Pre-coloring gives greater bandwidth to CoS 1 packets as they are initially subject to the CIR/CBS limits. Packets in CoS 0 are subject to the PIR limits. Based on the CIR/CBD, the PIR/PBS, and the conform, exceed, and violate actions specified below:

- TCP packets with rates less than or equal to the CIR/CBS in class CoS 1 are conforming to the rate (green).
- These packets will be dropped randomly at an increasing rate between 0 and 3% when the outgoing interface is congested between 80 and 100%.
- TCP packets with rates above the CIR/CBS and less than or equal to PIR/PBS in either class CoS 1 or class CoS 2 are policed as exceeding the CIR (yellow). These packets will be dropped randomly at an increasing rate between 0 and 5% when the outgoing interface is congested between 70 and 100%.

- TCP packets with rates higher than the PIR/PBS or which belong to neither class CoS 1 or class CoS 2 violate the rate (red). These packets will be dropped randomly at an increasing rate between 0 and 10% when the outgoing interface is congested between 50 and 100%.
- Non-TCP packets in CoS queue 0 or 1 will be dropped randomly at an increasing rate between 0 and 15% when the outgoing interface is congested between 50 and 100%.

```
console(config-policy-classmap)#police-two-rate 800 96 950 128
conform-action transmit exceed-action transmit violate-action
transmit
console(config-policy-classmap)#conform-color cos1 exceed-
color cos0
console(config-policy-classmap)#exit
console(config-policy-map)#exit
```

9 Enable WRED drop on traffic classes 0 and 1.

```
console(config)#cos-queue random-detect 0 1
```

10 Set the exponential-weighting-constant. The exponential weighting constant smooths the result of the average queue depth calculation by the following function:

*average depth = (previous queue depth * (1-1/2 ^n)) + (current queue depth * 1/2 ^n).*

The average depth is used in calculating the amount of congestion on a queue. Because the instantaneous queue depth fluctuates rapidly, larger values of the weighting constant cause the average queue depth value to respond to changes more slowly than smaller values.

```
console(config)#random-detect exponential-weighting-constant 4
```

11 Configure the queue parameters for traffic class 0 and 1. We set the minimum threshold and maximum thresholds to 80–100% for green traffic, 70–100% for yellow traffic, and 50–100% for red traffic. Non-TCP traffic drops in the 50–100% congestion range. Green traffic is dropped at a very low rate to slowly close the TCP window. Yellow and red traffic are dropped more aggressively.

```
console(config)#random-detect queue-parms 0 1 min-thresh 80 70
50 50 max-thresh 100 100 100 100 drop-prob-scale 3 5 10 15
```

12 Assign the color policies to ports. The metering policies are applied on ingress ports.

```
console(config)#interface Te1/0/22
console(config-if-Te1/0/22)#service-policy in simple-policy
console(config-if-Te1/0/22)#exit
console(config)#interface Te1/0/23
console(config-if-Te1/0/23)#service-policy in two-rate-policy
console(config-if-Te1/0/23)#exit
```

Example 2: Long-Lived Congestion

The following example enables WRED discard for non-color-aware traffic. Since a color-aware policer is not enabled, all traffic is treated as if it were colored "green." This means that only the "green" TCP and non-TCP WRED thresholds are active. Since the default CoS queue is 1, this example is suitable as a starting point for configuring WRED on a switch using the default settings. Packets will be randomly dropped on an egress port when the port becomes congested above the minimum threshold.

For the Dell Networking N2000/N3000 Series switches, a threshold of 100% corresponds to a buffer occupancy of 295428 bytes queued for transmission on an interface (about 300 jumbo frames). For the Dell Networking N4000 Series switch, a threshold of 100% corresponds to a buffer occupancy of 666757 bytes queued for transmission on an interface (about 600 jumbo frames).

If many (more than ¼) of the ports are becoming congested, then there is a chance that the switch buffer capacity will be exceeded and the switch will revert to tail-drop behavior. It is appropriate to lower the minimum threshold when many ports are becoming congested and to raise the minimum threshold when only one or two ports are becoming congested. Use the **show interfaces traffic** command to display interface congestion.

1 Configure the green thresholds for TCP traffic on CoS queue 1. Other thresholds are kept at their default values. The minimum threshold of 150% and maximum threshold of 200% with a drop probability of 2% are a good starting point for tuning the WRED parameters for an enterprise storage network that exhibits long term congestion on a few ports. Non-TCP traffic is configured for tail-drop at the 100% threshold. No color-aware processing is configured.

```
console(config)#random-detect queue-parms 1 min-thresh 150 30
20 100 max-thresh 200 90 80 100 drop-prob-scale 2 10 10 100
```

2 Enable WRED on cos-queue 1 (the default cos queue for packets marked user priority 0).

```
console(config)#cos-queue random-detect 1
```

Example 3: Data Center TCP (DCTCP) Configuration

This example globally configures a Dell Networking N2000/N3000 Series switch to utilize ECN marking of green packets queued for egress on CoS queues 0 and 1 using the DCTCP threshold as it appears in "DCTCP: Efficient Packet Transport for the Commoditized Data Center" Alizadeh, Greenberg, Maltz, Padhye, Patel, Prabhakar, Sengupta, and Sridharan, 2010.

NOTE: Data center TCP requires changes to the TCP stack on both ends of the connection. Reno TCP stacks do not always respond well to DCTCP settings.

In the first line of the configuration below, the first integer after the min-thresh keyword configures green-colored Congestion Enabled TCP packets in CoS queues 0 and 1 that exceed the WRED threshold (13% or ~38 Kbytes) to mark packets as Congestion Experienced. The first integer after the max-thresh parameter configures the upper threshold for green-colored TCP packets to the same value as the min-thresh threshold. This causes the switch to mark all ECN-capable queued packets as Congestion Experienced when the threshold is reached or exceeded. TCP packets without ECN capability bits set are dropped according to the normal WRED processing when the threshold is exceeded. Packets on other CoS queues are handled in the standard manner, i.e., tail-dropped when insufficient buffer is available. Yellow and red packet configuration (second and third threshold parameters) is kept at the defaults, as no metering to reclassify packets from green to yellow or red is present. The last threshold parameter configures non-TCP packets in CoS queues 0 and 1 to be processed with the WRED defaults. The **ecn** keyword enables ECN marking of ECN-capable packets on CoS queues 0 and 1. The weighting constant is set to 0 in the second line of the configuration, as described in the DCTCP paper cited above. Finally, CoS queues 0 and 1 are configured for WRED in the last line of the configuration.

```
console(config)#random-detect queue-parms 0 1 min-thresh 13 30 20
100 max-thresh 13 90 80 100 drop-prob-scale 100 10 10 10 ecn
console(config)#random-detect exponential-weighting-constant 0
console(config)#cos-queue random-detect 0 1
```

45

Auto VoIP

Dell Networking N1500, N2000, N3000, and N4000 Series Switches

Voice over Internet Protocol (VoIP) allows you to make telephone calls using a computer network over a data network like the Internet. With the increased prominence of delay-sensitive applications (voice, video, and other multimedia applications) deployed in networks today, proper QoS configuration will ensure high-quality application performance. The Auto VoIP feature is intended to provide an easy classification mechanism for voice packets so that they can be prioritized above data packets in order to provide better QoS. Because Auto VoIP is limited to 16 sessions, Voice VLAN is the preferred solution for enterprises wishing to deploy a large scale voice service. Voice VLAN and Auto VoIP are incompatible. Only one of Voice VLAN or Auto VoIP should be enabled in the switch.

The topics covered in this chapter include:

- Auto VoIP Overview
- Default Auto VoIP Values
- Configuring Auto VoIP (Web)
- Configuring Auto VoIP (CLI)

Auto VoIP Overview

The Auto VoIP feature explicitly matches VoIP streams in Ethernet switches and provides them with a better class of service than ordinary traffic. If you enable the Auto VoIP feature on an interface, the interface scans incoming traffic for the following call-control protocols:

- Session Initiation Protocol (SIP)
- H.323
- Skinny Client Control Protocol (SCCP)

When a call-control protocol is detected the switch assigns the traffic in that session to the highest CoS queue, which is generally used for time-sensitive traffic.

Auto VoIP is limited to 16 active sessions and makes use of the switch CPU to classify traffic. It is preferable to use the Voice VLAN feature in larger enterprise environments as it uses the switching silicon to classify voice traffic onto a VLAN.

Auto VoIP is incompatible with Voice VLAN and should not be enabled on switches on which Voice VLAN is enabled.

How Does Auto VoIP Use ACLs?

Auto VoIP utilizes ACL lists from the global system pool. ACL lists allocated by Auto VoIP reduce the total number of ACLs available for use by the network operator. Enabling Auto VoIP uses one ACL list (slice) to monitor for VoIP sessions. Each monitored VoIP session utilizes two rules from an additional ACL list. This means that the maximum number of ACL lists (slices) allocated by Auto VoIP is two.

Default Auto VoIP Values

Table 45-1 shows the global default value for Auto VoIP.

Table 45-1. Auto VoIP Global Defaults

Parameter	Default Value
Auto VoIP	Disabled

Configuring Auto VoIP (Web)

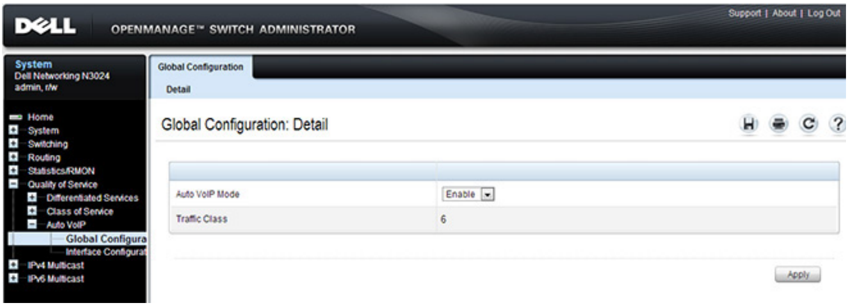
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring Auto VoIP features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click  at the top of the page.

Auto VoIP Global Configuration

Use the **Global Configuration** page to enable or disable Auto VoIP on all interfaces.

To display the Auto VoIP **Global Configuration** page, click **Quality of Service** → **Auto VoIP** → **Global Configuration** in the navigation menu.

Figure 45-1. Auto VoIP Global Configuration

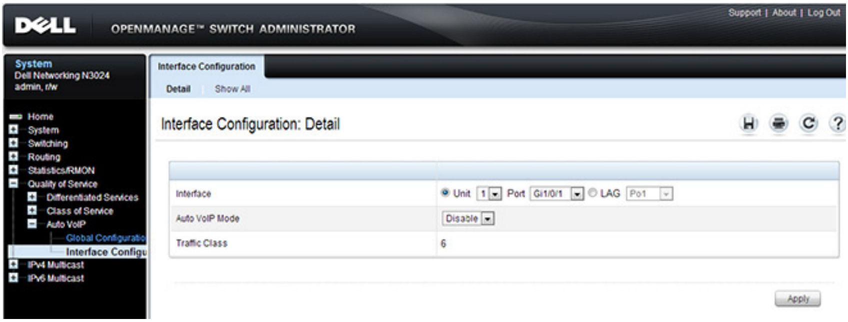


Auto VoIP Interface Configuration

Use the **Interface Configuration** page to enable or disable Auto VoIP on a particular interface.

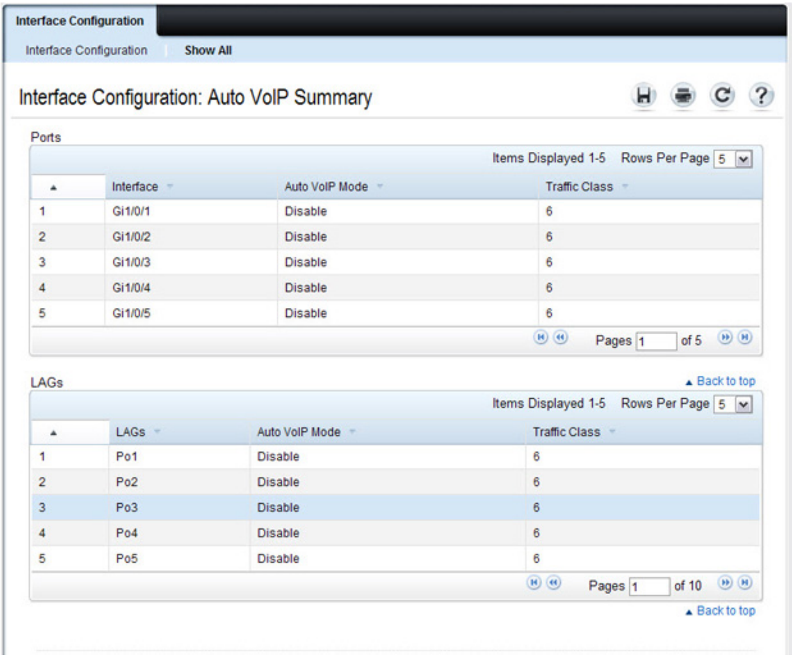
To display the **Interface Configuration** page, click **Quality of Service** → **Auto VoIP** → **Interface Configuration** in the navigation menu.

Figure 45-2. Auto VoIP Interface Configuration



To display summary Auto VoIP configuration information for all interfaces, click the **Show All** link at the top of the page.

Figure 45-3. Auto VoIP



Configuring Auto VoIP (CLI)

This section provides information about the commands used for configuring Auto VoIP settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Mapping Table Configuration

Beginning in Privileged Exec mode, use the following commands in to enable Auto VoIP and view its configuration.

CLI Command	Description
configure	Enter Global Configuration mode.
switchport voice detect auto	Enable the VoIP Profile on all the interfaces of the switch. The same command can be entered in Interface Configuration mode to enable it on a specific interface.
exit	Exit to Global Configuration Exec mode.
exit	Exit to Privilege Exec mode.
show switchport voice	Show the status of Auto VoIP on all interfaces or on an interface, if one is specified.

46

IPv4 and IPv6 Multicast

Dell Networking N3000 and N4000 Series Switches

NOTE: This feature is available only on Dell Networking N3000 and N4000 Series switches.

This chapter describes how to configure and monitor layer-3 (L3) multicast features for IPv4 and IPv6, including global IP and IPv6 multicast features as well as multicast protocols, including IGMP, DVMRP, and PIM for IPv4 and MLD and PIM for IPv6.

The topics covered in this chapter include:

- L3 Multicast Overview
- Default L3 Multicast Values
- Configuring General IPv4 Multicast Features (Web)
- Configuring IPv6 Multicast Features (Web)
- Configuring IGMP and IGMP Proxy (Web)
- Configuring MLD and MLD Proxy (Web)
- Configuring PIM for IPv4 and IPv6 (Web)
- Configuring DVMRP (Web)
- Configuring L3 Multicast Features (CLI)
- L3 Multicast Configuration Examples

L3 Multicast Overview

IP Multicasting enables a network host (or multiple hosts) to send an IP datagram to multiple destinations simultaneously. The initiating host sends each multicast datagram only once to a destination multicast group address, and multicast routers forward the datagram only to hosts who are members of the multicast group. Multicast enables efficient use of network bandwidth because each multicast datagram needs to be transmitted only once on each network link, regardless of the number of destination hosts. Multicasting contrasts with IP unicasting, which sends a separate datagram to each recipient host. The IP routing protocols can route multicast traffic, but the IP multicast protocols handle the multicast traffic more efficiently with better use of network bandwidth.

Applications that often send multicast traffic include video or audio conferencing, Whiteboard tools, stock distribution tickers, and IP-based television (IP/TV).

What Is IP Multicast Traffic?

IP multicast traffic is traffic that is destined to a host group. Host groups are identified by class D IP addresses, which range from 224.0.0.0 to 239.255.255.255. When a packet with a broadcast or multicast destination IP address is received, the switch will forward a copy into each of the remaining network segments in accordance with the IEEE MAC Bridge standard. Eventually, the packet is made accessible to all nodes connected to the network.

This approach works well for broadcast packets that are intended to be seen or processed by all connected nodes. In the case of multicast packets, however, this approach could lead to less efficient use of network bandwidth, particularly when the packet is intended for only a small number of nodes. Packets will be flooded into network segments where no node has any interest in receiving the packet. The L3 multicast features on the switch help to ensure that only the hosts in the multicast group receive the multicast traffic for that group.

Multicast applications send one copy of a packet, and address it to a group of receivers (Multicast Group Address) rather than to a single receiver (unicast address). Multicast depends on the network to forward the packets to only those networks and hosts that need to receive them.

Multicast Addressing

IPv4 multicast addresses in the range 224.0.0.X are link local addresses. These addresses are always flooded to all ports. Do not use these addresses for multicast data traffic. Because there is a 32:1 overlap in the IPv4 multicast address range, any address in the range 224–239.0.0.x and 224–239.128.0.0 should never be used for multicast data traffic.

239.0.0.0/8 is the locally scoped IPv4 multicast address range. Use addresses from this block for local/intra-domain multicast traffic. See RFC 2365 for further information

233.0.0.0/8 is the GLOP IPv4 public address range and is suitable for inter-domain multicast traffic. See RFC 2770 for further information.

232.0.0.0/8 is the PIM-SSM IPv4 public address space and is suitable for inter-domain traffic. See RFC 4608 for further information.

ffx5::/16 is the IPv6 site-local scope multicast address space and is suitable for intra-domain IPv6 multicast traffic.

ffx8::/16 is the IPv6 Organization-local scope multicast address space and is suitable for intra-domain traffic that may be sent over a VPN. These addresses are not valid in the public internet.

ffxe::/16 is the IPv6 Global scope multicast address space and is suitable for inter-domain traffic. These addresses are valid in the public internet.

What Multicast Protocols Does the Switch Support?

Multicast protocols are used to deliver multicast packets from one source to multiple receivers. Table 46-1 summarizes the multicast protocols that the switch supports.

Table 46-1. Multicast Protocol Support Summary

Protocol	IPv4 or IPv6	For Communication Between
IGMP	IPv4	Host-to-L3 switch/router
IGMP Proxy	IPv4	Host-to-L3 switch/router
MLD	IPv6	Host-to-L3 switch/router
MLD Proxy	IPv6	Host-to-L3 switch/router
PIM-SM	IPv4 and IPv6	L3-switch/router-to-L3 switch/router
PIM-DM	IPv4 and IPv6	L3-switch/router-to-L3 switch/router
DVMRP	IPv4	L3-switch/router-to-L3 switch/router

What Are the Multicast Protocol Roles?

Hosts must have a way to identify their interest in joining any particular multicast group, and routers must have a way to collect and maintain group memberships. These functions are handled by the IGMP protocol in IPv4. In IPv6, multicast routers use the Multicast Listener Discover (MLD) protocol to maintain group membership information.

Multicast routers must also be able to construct a multicast distribution tree that enables forwarding multicast datagrams only on the links that are required to reach a destination group member. Protocols such as DVMRP, and PIM handle this function.

IGMP and MLD are multicast group discovery protocols that are used between the clients and the local multicast router. PIM-SM, PIM-DM, and DVMRP are multicast routing protocols that are used across different subnets, usually between the local multicast router and remote multicast router.

When Is L3 Multicast Required on the Switch?

Use the IPv4/IPv6 multicast feature on Dell Networking N-Series switches to route multicast traffic between VLANs on the switch. If all hosts connected to the switch are on the same subnet, there is no need to configure the IP/IPv6 multicast feature. If the switch does not require L3 routing, IGMP snooping or MLD snooping can be used to manage port-based multicast group membership. For more information, see "What Is IGMP Snooping? " on page 869 and "What Is MLD Snooping? " on page 871. If the local network does not have a multicast router, the switch can be configured to act as the IGMP querier. For more information, see "IGMP Snooping Querier " on page 871.

If the switch is configured as a L3 switch and handles inter-VLAN routing through static routes, OSPF, or RIP, and multicast traffic is transmitted within the network, enabling and configuring L3 multicast routing on the switch is recommended.

By default, multicast packets locally routed into a VLAN by the router are flooded to all ports in the VLAN. Multicast packets ingressing a port that is a member of a routed VLAN are flooded to all ports in the VLAN other than the receiving port. Although IGMP/MLD snooping can be used to mitigate this behavior, it is strongly recommended that multicast routed VLANs only contain two ports, one on each connecting switch. A VLAN carrying multicast traffic should never traverse a multicast router, as ingress multicast traffic is layer-2-switched across the VLAN, defeating the purpose of the multicast router.

Determining Which Multicast Protocols to Enable

IGMP is required on any multicast router that serves IPv4 hosts. IGMP is not required on inter-router links. MLD is required on any router that serves IPv6 hosts. MLD is not required on inter-router links. PIM-DM, PIM-SM, and DVMRP are multicast routing protocols that help determine the best route for IP (PIM and DVMRP) and IPv6 (PIM) multicast traffic.

IGMP is automatically enabled whenever an IPv4 multicast routing protocol is enabled that requires it, i.e., PIM-SM, PIM-DM, and DVMRP via the CLI. Likewise, MLD is automatically enabled whenever an IPv6 multicast routing protocol is enabled that requires it (PIM-SM and PIM-DM) via the CLI. IGMP and MLD may not be separately enabled or disabled via the CLI. They may be separately enabled/disabled via the web.

For more information about when to use PIM-DM, see "Using PIM-DM as the Multicast Routing Protocol " on page 1525. For more information about when to use PIM-SM, see "Using PIM-SM as the Multicast Routing Protocol " on page 1516. For more information about when to configure DVMRP, see "Using DVMRP as the Multicast Routing Protocol " on page 1527.

What Is the Multicast Routing Table?

Multicast capable/enabled routers forward multicast packets based on the routes in the Multicast Routing Information Base (MRIB). These routes are created in the MRIB during the process of building multicast distribution trees by the Multicast Protocols running on the router. Different IP Multicast routing protocols use different techniques to construct these multicast distribution trees.

What Is IGMP?

The Internet Group Management Protocol (IGMP) is used by IPv4 systems (hosts, L3 switches, and routers) to report their IP multicast group memberships to any neighboring multicast routers. The Dell Networking N1500, N2000, N3000, and N4000 Series switches performs the multicast router role of the IGMP protocol, which means it collects the membership information needed by the active multicast routing protocol. IGMP is automatically enabled when PIM or DVMRP are enabled via the CLI.

The Dell Networking N1500, N2000, N3000, and N4000 Series switches also supports IGMP Version 3. Version 3 adds support for source filtering, which is the ability for a system to report interest in receiving packets only from specific source addresses, as required to support Source-Specific Multicast [SSM], or from all but specific source addresses, sent to a particular multicast address. Version 3 is designed to be interoperable with Versions 1 and 2.

Understanding IGMP Proxy

IGMP proxy enables a multicast router to learn multicast group membership information and forward multicast packets based upon the group membership information. The IGMP Proxy is capable of functioning only in certain topologies that do not require Multicast Routing Protocols (i.e., DVMRP, PIM-DM, and PIM-SM) and have a tree-like topology, as there is no support for features like reverse path forwarding (RPF) to correct packet route loops.

The proxy contains many downstream interfaces and a unique upstream interface explicitly configured. It performs the host side of the IGMP protocol on its upstream interface and the router side of the IGMP protocol on its downstream interfaces.

The IGMP proxy offers a mechanism for multicast forwarding based only on IGMP membership information. The router must decide about forwarding packets on each of its interfaces based on the IGMP membership information. The proxy creates the forwarding entries based on the membership information and adds it to the multicast forwarding cache (MFC) in order not to make the forwarding decision for subsequent multicast packets with same combination of source and group.

What Is MLD?

Multicast Listener Discovery (MLD) protocol enables IPv6 routers to discover the presence of multicast listeners, the hosts that wish to receive the multicast data packets, on its directly-attached interfaces. The protocol specifically discovers which multicast addresses are of interest to its neighboring nodes and provides this information to the active multicast routing protocol that makes decisions on the flow of multicast data packets. MLD is automatically enabled whenever IPv6 PIM is enabled on IPv6 interfaces via the CLI.

The Multicast router sends General Queries periodically to request multicast address listeners information from systems on an attached network. These queries are used to build and refresh the multicast address listener state on attached networks. Multicast listeners respond to these queries by reporting their multicast addresses listener state and their desired set of sources with Current-State Multicast address Records in the MLD2 Membership Reports. The Multicast router also processes unsolicited Filter-Mode-Change records and Source-List-Change Records from systems that want to indicate interest in receiving or not receiving traffic from particular sources.

The Dell Networking implementation of MLD v2 supports the multicast router portion of the protocol (i.e., not the listener portion). It is backward-compatible with MLD v1.

What Is PIM?

The Protocol Independent Multicast protocol is a simple, protocol-independent multicast routing protocol. PIM uses an existing unicast routing table and a Join/Prune/Graft mechanism to build a tree. Dell Networking N-Series switches support two types of PIM: sparse mode (PIM-SM) and dense mode (PIM-DM).

PIM-SM is most effective in networks with a sparse population of multicast receivers. In contrast, PIM-DM is most effective in networks with densely populated multicast receivers. In other words, PIM-DM can be used if the majority of network hosts request to receive a multicast stream, while PIM-SM might be a better choice in networks in which a small percentage of network hosts, located throughout the network, wish to receive the multicast stream.

Using PIM-SM as the Multicast Routing Protocol

PIM-SM is used to efficiently route multicast traffic to multicast groups that may span wide area networks and where bandwidth is constrained. PIM-SM uses shared trees by default and implements source-based trees for efficiency. PIM-SM assumes that no hosts want the multicast traffic unless they specifically ask for it. It initially creates a shared distribution tree centered on a defined "rendezvous point" (RP) through which source traffic is relayed to the ultimate receiver. Multicast traffic sources first send the multicast data to the RP, which in turn sends the data down the shared tree to the receivers. Shared trees centered on an RP do not necessarily provide the shortest or most optimal path. In such cases, a Dell Networking PIM-SM router adjacent to the host switches to the shortest path upon seeing the very first multicast data packet.

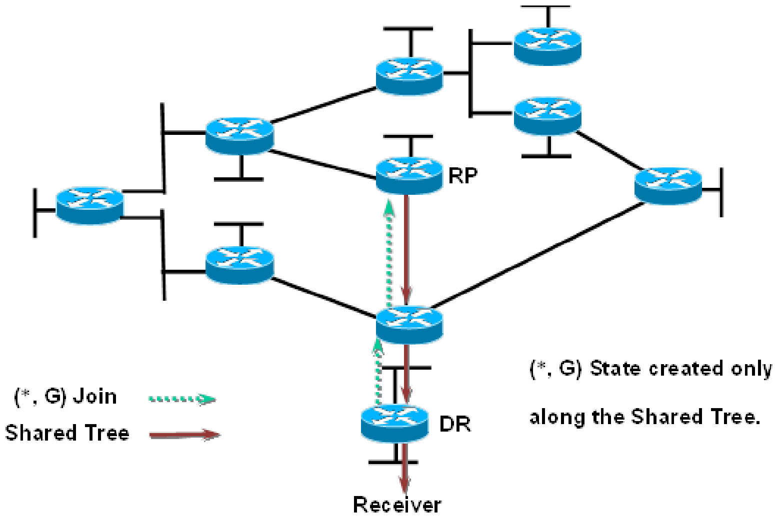
Many IP multicast applications, such as those that handle real-time dissemination of financial information, require high performance. Multicast group membership management (IGMP), unicast routing protocols (OSPF, RIP), and multicast routing protocols are all required to enable end-to-end multicast capabilities. The RP is a critical function for PIM-SM deployments. RP redundancy is always recommended. In a shared-tree model, multicast traffic from the multicast source is routed via the RP. If the RP goes down, the multicast receivers do not receive traffic until the RP comes up again. In general, more than one RP is configured (for a group range) to provide RP redundancy. The PIM-SM router acting as a BSR advertises the list of candidate RPs to all the PIM routers in the network. Each PIM router then runs the RP selection algorithm to determine an RP for the given group range. All the interested PIMSM routers then initiate re-reception of traffic through this new RP, and the multicast traffic is rerouted via the new RP. This is to provide high availability to the multicast applications and help ensure that the multicast traffic is recovered quickly in such scenarios.

PIM-SM Protocol Operation

This section describes the workings of PIM-SM protocol per RFC 4601. The protocol operates essentially in three phases, as explained in the following sections.

Phase-1: RP Tree

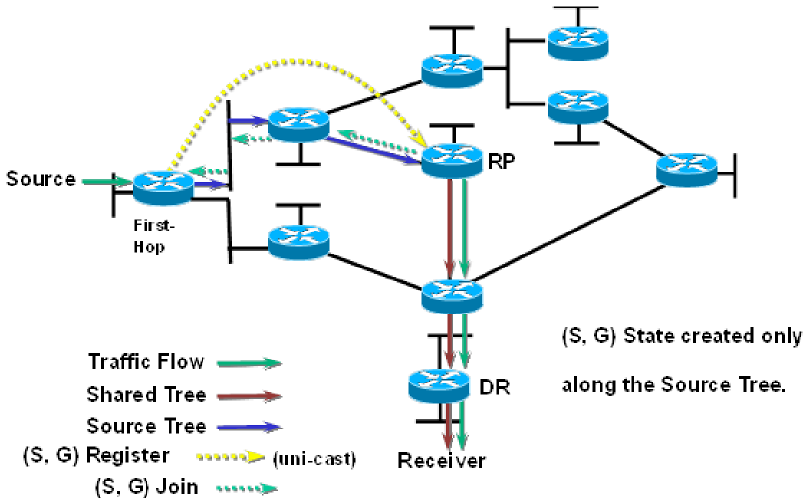
Figure 46-1. PIM-SM Shared Tree Join



- In this example, an active receiver (attached to leaf router at the bottom of the drawing) has joined multicast group "G".
- The leaf router (labeled DR above) knows the IP address of the Rendezvous Point (RP) for group G and sends a (*, G) Join for this group towards the RP.
- This (*, G) Join travels hop-by-hop to the RP, building a branch of the Shared Tree that extends from the RP to the last-hop router directly connected to the receiver.
- At this point, group "G" traffic can flow down the Shared Tree to the receiver.

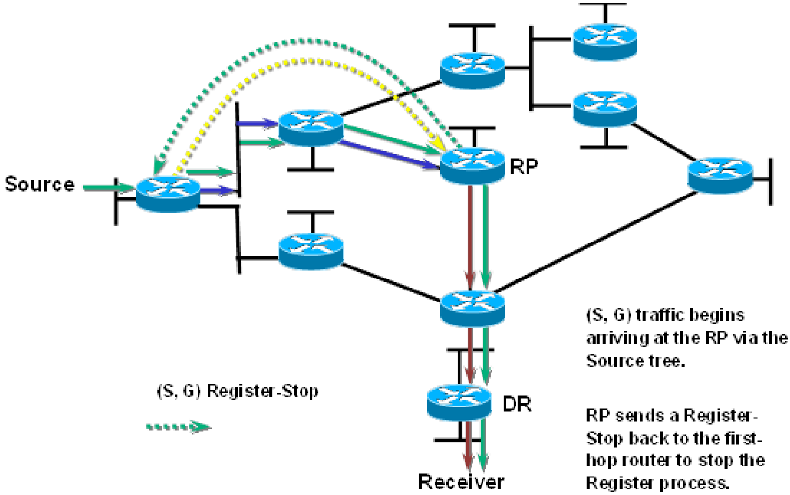
Phase-2: Register Stop

Figure 46-2. PIM-SM Sender Registration—Part1



- As soon as an active source for group G sends a packet, the designated router (DR) that is attached to this source is responsible for “Registering” this source with the RP and requesting the RP to build a tree back to that router.
- To do this, the source router encapsulates the multicast data from the source in a special PIM-SM message, called the Register message, and unicasts that data to the RP.
- When the RP receives the Register message, it does two things:
  - It de-encapsulates the multicast data packet inside of the Register message and forwards it down the Shared Tree.
  - The RP sends a source group (S, G) Join back towards the source to create a branch of an (S, G) Shortest-Path Tree (SPT). This results in the (S, G) state being created in the entire router path along the SPT, including the RP.

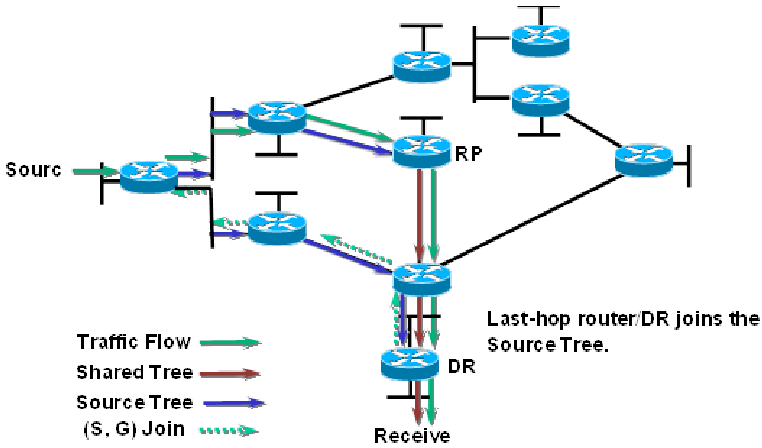
Figure 46-3. PIM-SM Sender Registration—Part 2



- As soon as the SPT is built from the Source router to the RP, multicast traffic begins to flow unencapsulated from source S to the RP.
- Once this is complete, the RP Router will send a "Register Stop" message to the first-hop router to tell it to stop sending the encapsulated data to the RP.

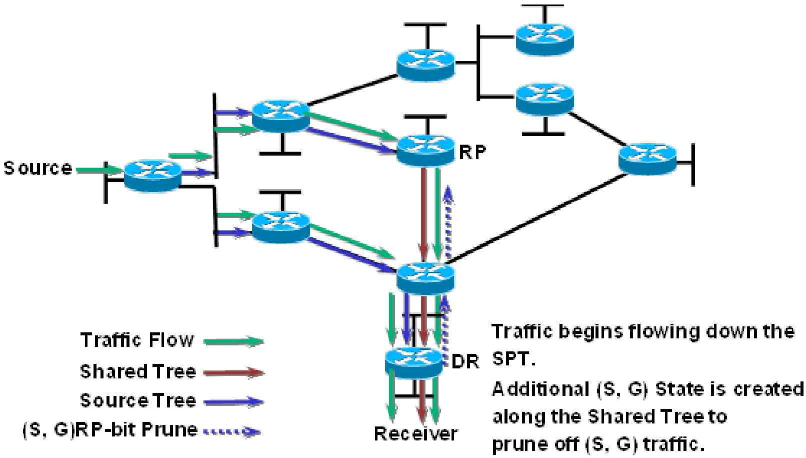
Phase 3: Shortest Path Tree

Figure 46-4. PIM-SM SPT—Part 1



- PIM-SM has the capability for last-hop routers (i.e., routers with directly connected group members) to switch to the Shortest-Path Tree and bypass the RP. This switchover is based upon an implementation-specific function called SwitchToSptDesired(S,G) in the standard and generally takes a number of seconds to switch to the SPT.
- In the above example, the last-hop router (at the bottom of the drawing) sends an (S, G) Join message toward the source to join the SPT and bypass the RP.
- This (S, G) Join messages travels hop-by-hop to the first-hop router (i.e., the router connected directly to the source), thereby creating another branch of the SPT. This also creates (S, G) state in all the routers along this branch of the SPT.

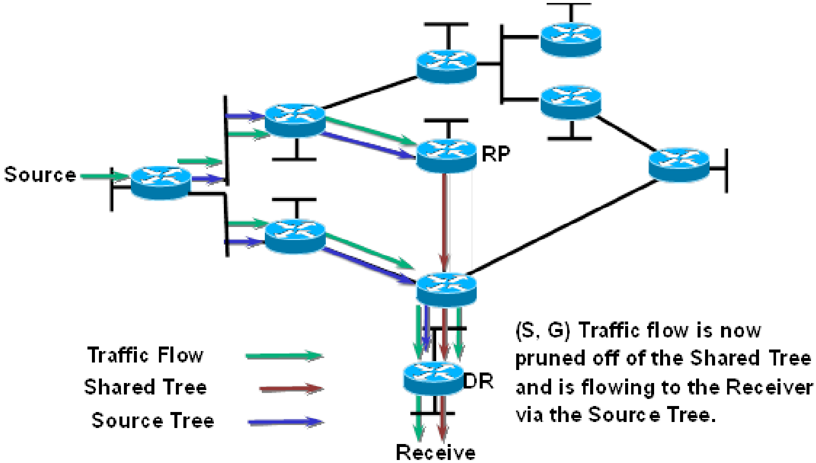
Figure 46-5. PIM-SM SPT—Part 2



- Finally, special (S, G) RP-bit Prune messages are sent up the Shared Tree to prune off this (S, G) traffic from the Shared Tree.

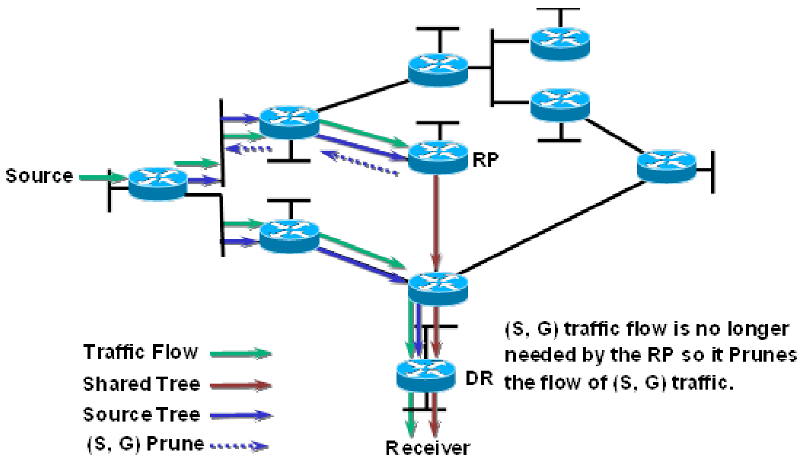
  If this were not done, (S, G) traffic would continue flowing down the Shared Tree resulting in duplicate (S, G) packets arriving at the receiver.

Figure 46-6. PIM-SM SPT—Part 3



- At this point, (S, G) traffic is now flowing directly from the first -hop router to the last-hop router and from there to the receiver.

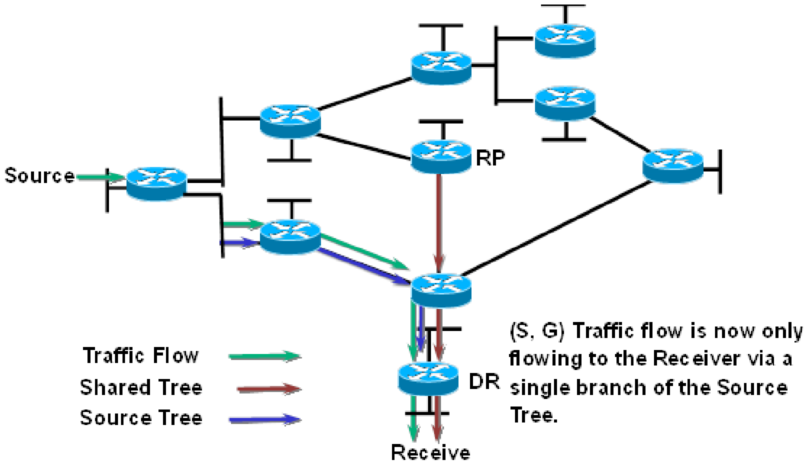
Figure 46-7. PIM-SM SPT—Part 4



- At this point, the RP no longer needs the flow of (S, G) traffic since all branches of the Shared Tree (in this case there is only one) have pruned off the flow of (S, G) traffic.
- As a result, the RP will send (S, G) Prunes back toward the source to shut off the flow of the now unnecessary (S, G) traffic to the RP.

NOTE: This will occur if the RP has received an (S, G) RP-bit Prune on all interfaces on the Shared Tree.

Figure 46-8. PIM-SM SPT—Part 5



- As a result of the SPT-Switchover, (S, G) traffic is now flowing only from the first-hop router to the last-hop router and from there to the receiver. Notice that traffic is no longer flowing to the RP.

The PIM standard requires support for multi-hop RP in that a router running PIM can act as an RP even if it is multiple router hops away from the multicast source. This requires that the first-hop router perform encapsulation of the multicast data and forward it as unicast toward the RP. In practice, this encapsulation is almost always performed in software due to the complexity of the operation. Likewise, the RP must perform de-encapsulation and forwarding of the multicast packets in software. This creates a performance problem in that it limits the number of packets that can be processed and places a high load on the CPUs in the first hop and RP routers, which can then adversely affect other router functions.

Dell Networking Optimizations to PIM-SM

Dell Networking N-Series switches perform the following optimizations to reduce the impact of multicast encapsulation/de-encapsulation and provide a higher level of multicast performance in the network.

- Limiting the number of packets sent to the RP by the first-hop router. When a multicast data source (S) starts sending data destined for a multicast group (G), the first-hop router receives these packets and traps them to its local CPU. A Dell Networking first-hop router immediately blocks further data packets in the stream and prevents them from reaching the CPU. The first-hop router then unicast-encapsulates the first received data packet in the form of a PIM Register message and software forwards it to the RP.

  When a Dell Networking first-hop router subsequently receives the PIM Join from the RP, the block is replaced with a regular multicast forwarding entry so that subsequent data packets are forwarded in the hardware.

  If the initial Register message(s) does not reach the RP, or the PIM Join sent in response does not reach the first-hop router, then the data stream would never get forwarded. To solve this, the negative entry is timed out and removed after 3 seconds so that the process can be repeated until it succeeds.

- In Phase 3—Shortest Path Tree, the last-hop router initiates a switchover to the SPT tree by sending a PIM (S,G) Join message towards the source as soon as it receives the first data packet via the (*,G) shared tree. Per the standard, this function is used to detect suboptimal routing of multicast traffic. Dell Networking multicast eliminates the SwitchToSptDesired(S,G) function and performs as if the SwitchToSptDesired(S,G) function always returns "true" as soon as it receives the first multicast packet instead of waiting for 30 seconds.
- Dell Networking RPs do not wait to receive the native multicast data but immediately respond to the PIM (S,G) Join by sending a 'Register Stop' message to the source's first-hop router to inform it that it can stop

sending the encapsulated Register messages. This removes the load from the CPU of the first-hop router and the RP, as they no longer need to encapsulate and de-encapsulate register messages with multicast data.

These optimizations significantly reduce the load on first-hop routers and RPs to encapsulate/de-encapsulate PIM register messages and their associated multicast data. In addition, the switchover to the SPT is initiated immediately upon the first multicast packet reaching the last-hop router. This leads to significantly faster response times for receiving the full multicast stream directly from the first-hop router (as opposed to the typical bandwidth-limited stream traversing the RP).

Using PIM-DM as the Multicast Routing Protocol

Unlike PIM-SM, PIM-DM creates source-based shortest-path distribution trees that make use of reverse-path forwarding (RPF). PIM-DM assumes that when a sender starts sending data, all downstream routers and hosts want to receive a multicast datagram. PIM-DM initially floods multicast traffic throughout the network. Routers that do not have any downstream neighbors prune back the unwanted traffic. In addition to PRUNE messages, PIM-DM makes use of graft and assert messages. Graft messages are used whenever a new host wants to join the group. Assert messages are used to shutoff duplicate flows on the same multi-access network.

There are two versions of PIM-DM. Version 2 does not use the IGMP message; instead, it uses a message that is encapsulated in an IP packet, with protocol number 103. In Version 2, a Hello message is introduced in place of a query message.

PIM-DM is appropriate for:

- Densely distributed receivers
- Few senders-to-many receivers (due to frequent flooding)
- High volume of multicast traffic
- Constant stream of traffic

To minimize the repeated flooding of datagrams and subsequent pruning associated with a particular source-group (S,G) pair, PIM-DM uses a State Refresh message. This message is sent by the router(s) directly connected to the source and is propagated throughout the network. When received by a router on its RPF interface, the State Refresh message causes an existing prune state to be refreshed. State Refresh messages are generated periodically by the router directly attached to the source.

What Is DVMRP?

DVMRP is an interior gateway protocol that is suitable for routing multicast traffic within an autonomous system (AS). DVMRP should not be used between different autonomous systems due to limitations with hop count and scalability.

NOTE: In addition to DVMRP, the switch supports the Protocol-Independent Multicast (PIM) sparse-mode (PIM-SM) and dense-mode (PIM-SM) routing protocol. Only one multicast routing protocol can be operational on the switch at any time. If you enable DVMRP, PIM must be disabled. Similarly, if PIM is enabled, DVMRP must be disabled.

DVMRP exchanges probe packets with all its DVMRP-enabled routers, it establishes two-way neighboring relationships, and it builds a neighbor table. DVMRP exchanges report packets and creates a unicast topology table, with which it builds the multicast routing table. This table is used to route the multicast packets. Since every DVMRP router uses the same unicast routing protocol, routing loops are avoided.

Understanding DVMRP Multicast Packet Routing

DVMRP is based on RIP; it forwards multicast datagrams to other routers in the AS and constructs a forwarding table based on information it learns in response. More specifically, it uses this sequence.

- A new multicast packet is forwarded to the entire multicast network, with respect to the time-to-live (TTL) of the packet.
- The TTL restricts the area to be flooded by the message.
- All routers that do not have members on directly-attached subnetworks send back *Prune messages* to the upstream router.
- The branches that transmit a prune message are deleted from the delivery tree.
- The delivery tree which is spanning to all the members in the multicast group, is constructed in the form of a DVMRP forwarding table.

Using DVMRP as the Multicast Routing Protocol

DVMRP is used to communicate multicast information between L3 switches or routers. If a Dell Networking N1500, N2000, N3000, and N4000 Series switches handles inter-VLAN routing for IP traffic, including IP multicast traffic, multicast routing might be required on the switch.

DVMRP is best suited for small networks where the majority of hosts request a given multicast traffic stream. DVMRP is similar to PIM-DM in that it floods multicast packets throughout the network and prunes branches where the multicast traffic is not desired. DVMRP was developed before PIM-DM, and it has several limitations that do not exist with PIM-DM.

You might use DVMRP as the multicast routing protocol if it has already been widely deployed within the network.

Microsoft Network Load Balancing

Dell Networking N-Series switches support Microsoft Network Load Balancing (NLB) in unicast mode only. When using Microsoft NLB, ensure that the Cluster Operation Mode is configured to the default value of Unicast.

Default L3 Multicast Values

IP and IPv6 multicast is disabled by default. Table 46-2 shows the default values for L3 multicast and the multicast protocols.

Table 46-2. L3 Multicast Defaults

Parameter	Default Value
IPv4 Multicast Defaults	
L3 Multicast Admin Mode	Disabled
Maximum Multicast Routing Table Entries	2048 (1536 IPv4/512 IPv6) Switch sizes are as follows: Dell Networking N2000/N3000 Series—1536 IPv4 / 512 IPv6 Dell Networking N4000 Series—512 IPv4 / 256 IPv6
Static Multicast Routes	None configured
Interface TTL Threshold	1
IGMP Defaults	
IGMP Admin Mode	Disabled globally and on all interfaces
IGMP Version	v3
IGMP Robustness	2
IGMP Query Interval	125 seconds
IGMP Query Max Response Time	100 seconds
IGMP Startup Query Interval	31 seconds
IGMP Startup Query Count	2
IGMP Last Member Query Interval	1 second
IGMP Last Member Query Count	2
IGMP Proxy Interface Mode	Disabled
IGMP Proxy Unsolicited Report Interval	1 second
MLD Defaults	
MLD Admin Mode	Disabled globally and on all interfaces

Table 46-2. L3 Multicast Defaults (Continued)

Parameter	Default Value
MLD Version	v2
MLD Query Interval	125 seconds
MLD Query Max Response Time	10,000 milliseconds
MLD Last Member Query Interval	1000 milliseconds
MLD Last Member Query Count	2
MLD Proxy Interface Mode	Disabled
MLD Proxy Unsolicited Report Interval	1 second
PIM Defaults	
PIM Protocol	Disabled globally and on all interfaces
PIM Hello Interval	30 seconds (when enabled on an interface)
PIM-SM Join/Prune Interval	60 seconds (when enabled on an interface)
PIM-SM BSR Border	Disabled
PIM-SM DR Priority	1 (when enabled on an interface)
PIM Candidate Rendezvous Points (RPs)	None configured
PIM Static RP	None configured
PIM Source-Specific Multicast (SSM) Range	None configured. Default SSM group address is 232.0.0.0/8 for IPv4 multicast and ff3x::/32 for IPv6 multicast.
PIM BSR Candidate Hash Mask Length	30 (IPv4) 126 (IPv6)
PIM BSR Candidate Priority	0
DVMRP Defaults	
DVMRP Admin Mode	Disabled globally and on all interfaces
DVMRP Version	3
DVMRP Interface Metric	1

Configuring General IPv4 Multicast Features (Web)

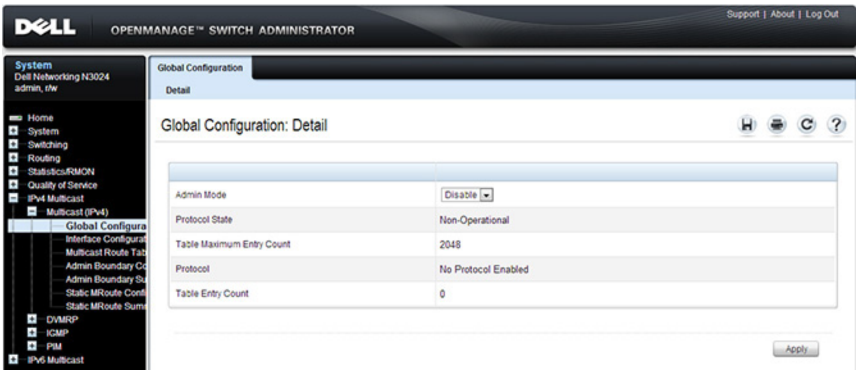
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring the L3 multicast features that are not protocol-specific on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

Multicast Global Configuration

Use the **Global Configuration** page to configure the administrative status of Multicast Forwarding in the router, and to display global multicast parameters.

To display the page, click **IPv4 Multicast** → **Multicast** → **Global Configuration** in the navigation panel.

Figure 46-9. Multicast Global Configuration

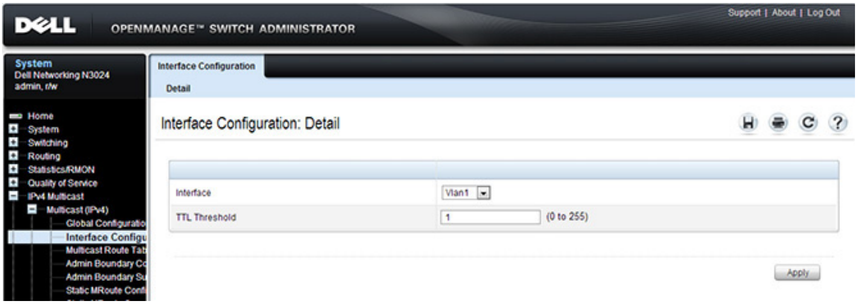


Multicast Interface Configuration

Use the **Interface Configuration** page to configure the TTL threshold of a multicast interface. At least one VLAN routing interface must be configured on the switch before fields display on this page.

To display the page, click **IPv4 Multicast** → **Multicast** → **Interface Configuration** in the navigation panel.

Figure 46-10. Multicast Interface Configuration

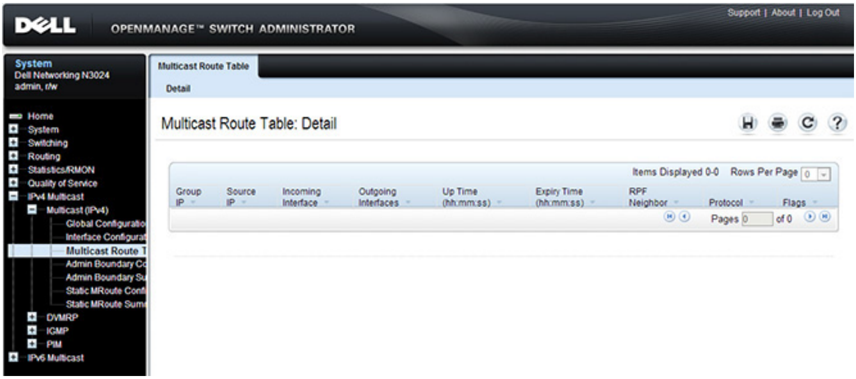


Multicast Route Table

Use the **Route Table** page to view information about the multicast routes in the IPv4 multicast routing table.

To display the page, click **IPv4 Multicast** → **Multicast** → **Multicast Route Table** Multicast Route Table

Figure 46-11. Multicast Route Table

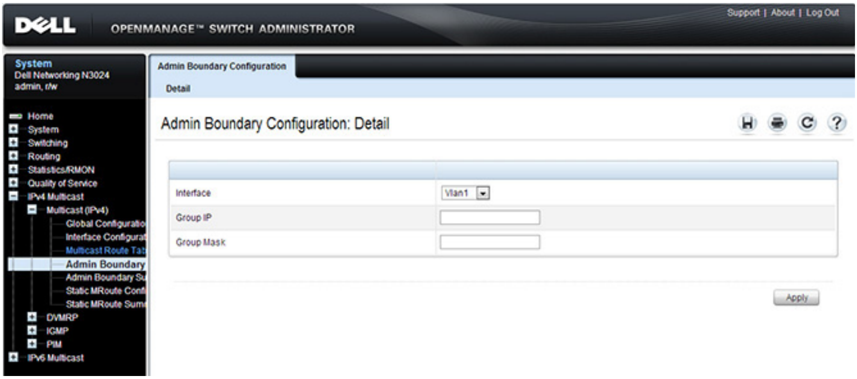


Multicast Admin Boundary Configuration

The definition of an administratively scoped boundary is a way to stop the ingress and egress of multicast traffic for a given range of multicast addresses on a given routing interface. Use the **Admin Boundary Configuration** page to configure a new or existing administratively scoped boundary. To see this page, you must have configured a valid routing interface and multicast.

To display the page, click **IPv4 Multicast** → **Multicast** → **Admin Boundary Configuration** in the navigation panel.

Figure 46-12. Multicast Admin Boundary Configuration

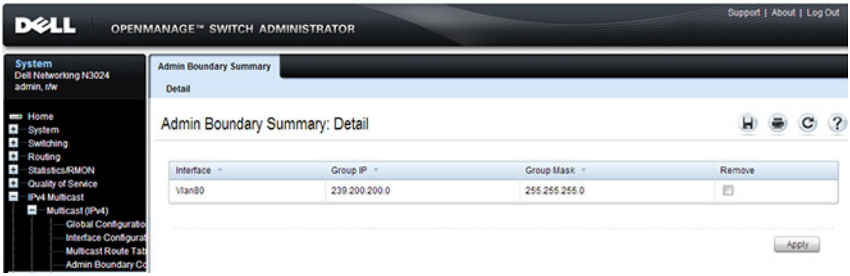


Multicast Admin Boundary Summary

Use the **Admin Boundary Summary** page to display existing administratively scoped boundaries.

To display the page, click **IPv4 Multicast** → **Multicast** → **Admin Boundary Summary** in the navigation panel.

Figure 46-13. Multicast Admin Boundary Summary

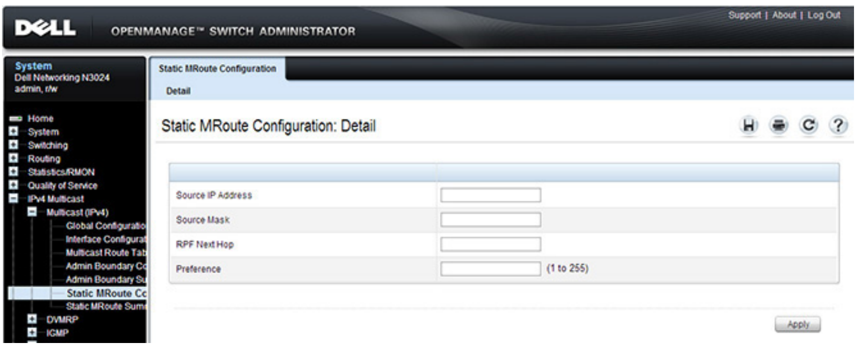


Multicast Static MRoute Configuration

Use the **Static MRoute Configuration** page to configure a new static entry in the Mroute table or to modify an existing entry.

To display the page, click **IPv4 Multicast** → **Multicast** → **Static MRoute Configuration** in the navigation panel.

Figure 46-14. Multicast Static MRoute Configuration

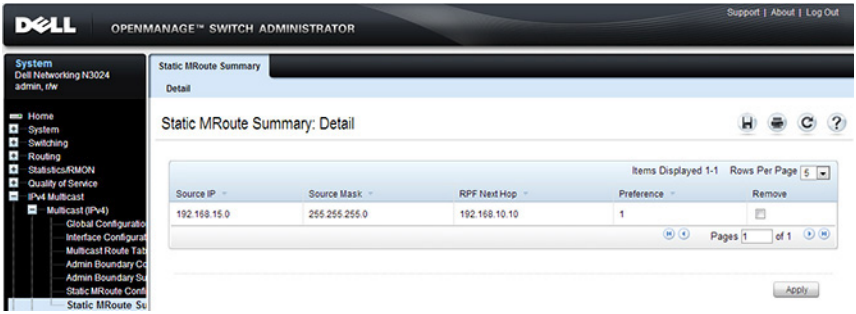


Multicast Static MRoute Summary

Use the **Static MRoute Summary** page to display static routes and their configurations.

To display the page, click **IPv4 Multicast** → **Multicast** → **Static MRoute Summary** in the navigation panel.

Figure 46-15. Multicast Static MRoute Summary



Configuring IPv6 Multicast Features (Web)

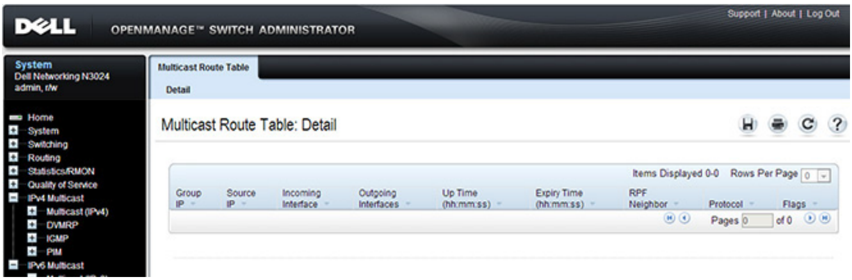
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring the IPv6 multicast features that are not protocol-specific on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

IPv6 Multicast Route Table

Use the **Multicast Route Table** page to view information about the multicast routes in the IPv6 multicast routing table.

To display the page, click **IPv6 Multicast** → **Multicast** → **Multicast Route Table**.

Figure 46-16. IPv6 Multicast Route Table



Configuring IGMP and IGMP Proxy (Web)

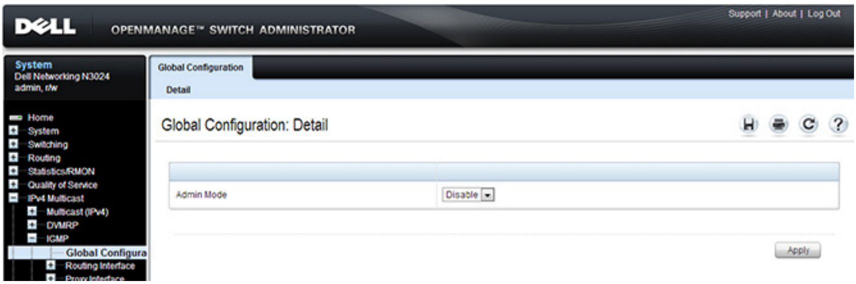
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring the IGMP and IGMP proxy features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click at the top of the page.

IGMP Global Configuration

Use the **Global Configuration** page to set IGMP on the system to active or inactive.

To display the page, click **IPv4 Multicast** → **IGMP** → **Global Configuration** in the navigation panel.

Figure 46-17. IGMP Global Configuration

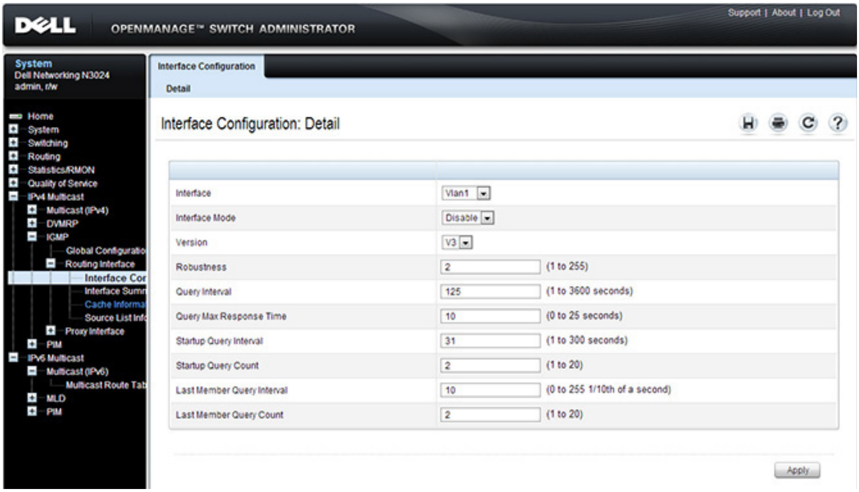


IGMP Interface Configuration

Use the **Interface Configuration** page to configure and/or display router interface parameters. At least one valid routing interface must be configured before this page can be accessed to configure IP Multicast IGMP.

To display the page, click **IPv4 Multicast** → **IGMP** → **Routing Interface** → **Interface Configuration** in the navigation panel.

Figure 46-18. IGMP Interface Configuration

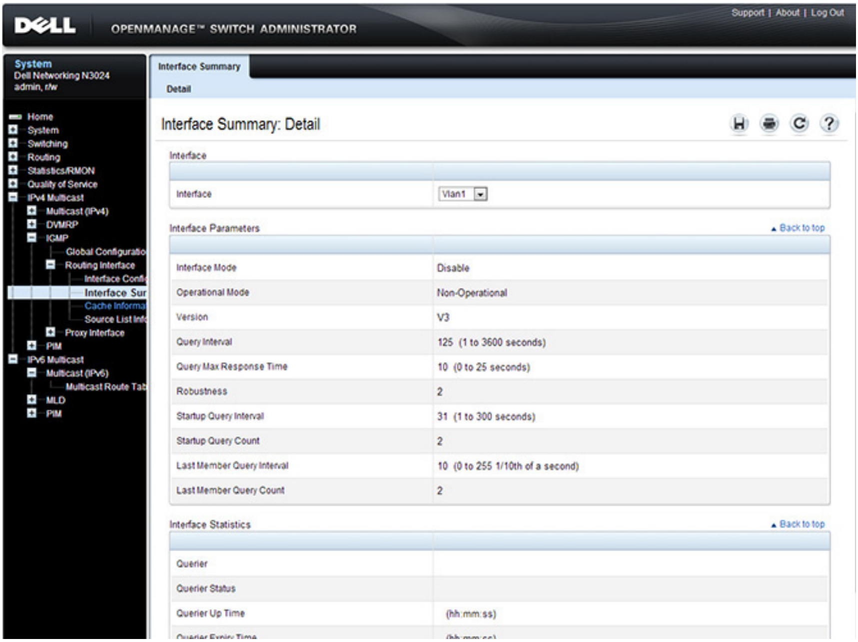


IGMP Interface Summary

Use the **Interface Summary** page to display IGMP routing parameters and data. You must configure at least one IGMP router interface to access this page.

To display the page, click **IPv4 Multicast** → **IGMP** → **Routing Interface** → **Interface Summary** in the navigation panel.

Figure 46-19. IGMP Interface Summary

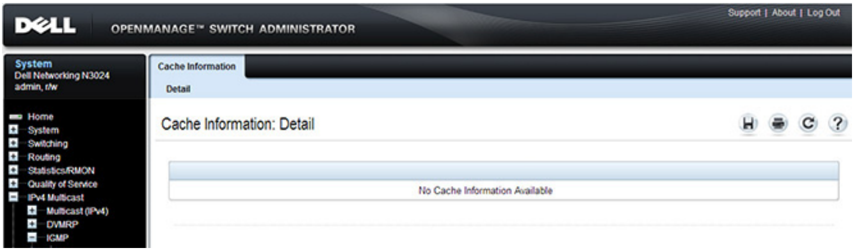


IGMP Cache Information

Use the **Cache Information** page to display cache parameters and data for an IP multicast group address. Group membership reports must have been received on the selected interface for data to display on the page.

To display the page, click **IPv4 Multicast** → **IGMP** → **Routing Interface** → **Cache Information** in the navigation panel.

Figure 46-20. IGMP Cache Information

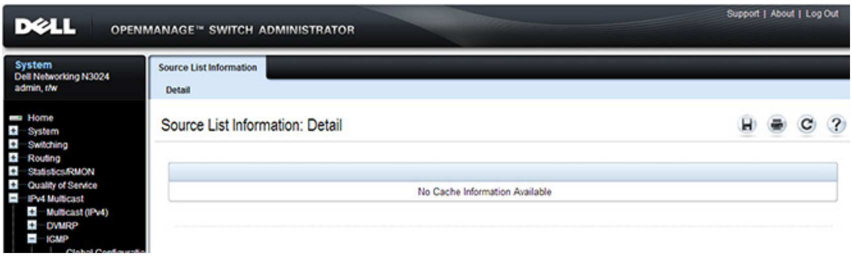


IGMP Interface Source List Information

Use the **Source List Information** page to display detailed membership information for an interface. Group membership reports must have been received on the selected interface for data to display information.

To display the page, click **IPv4 Multicast** → **IGMP** → **Routing Interface** → **Source List Information** in the navigation panel.

Figure 46-21. IGMP Interface Source List Information



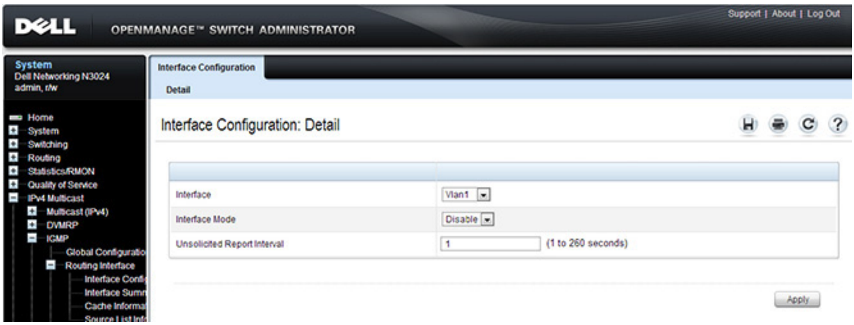
IGMP Proxy Interface Configuration

The IGMP Proxy is used by IGMP Router (IPv4 system) to enable the system to issue IGMP host messages on behalf of hosts that the system discovered through standard IGMP router interfaces. Thus, this feature acts as proxy to all hosts residing on its router interfaces.

Use the **Interface Configuration** page to configure IGMP proxy for a VLAN interface. You must have configured at least one VLAN routing interface before configuring or displaying data for an IGMP proxy interface, and it should not be an IGMP routing interface.

To display the page, click **IPv4 Multicast** → **IGMP** → **Proxy Interface** → **Interface Configuration** in the navigation panel.

Figure 46-22. IGMP Proxy Interface Configuration

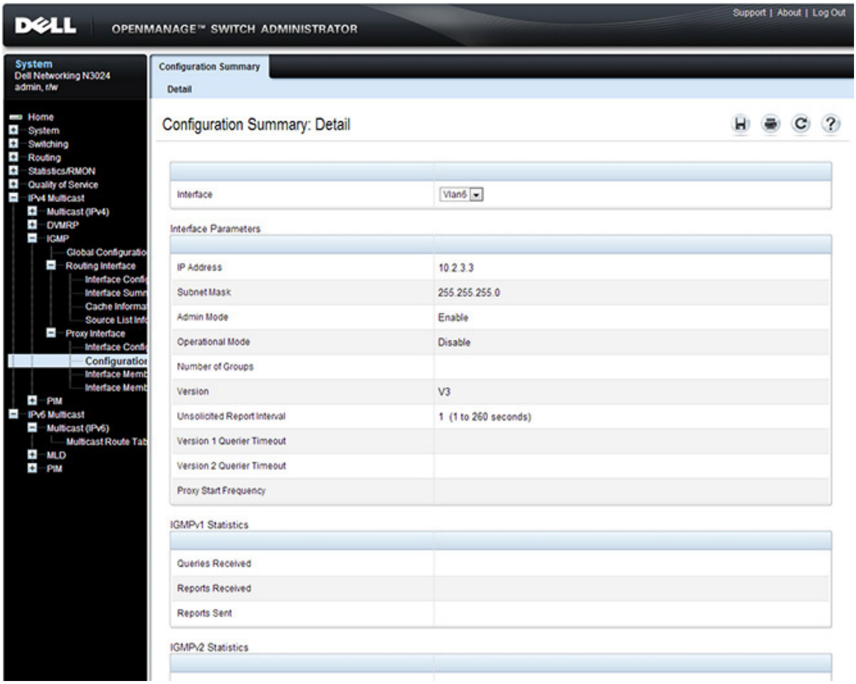


IGMP Proxy Configuration Summary

Use the **Configuration Summary** page to display proxy interface configurations by interface. You must have configured at least one VLAN routing interface configured before data displays on this page.

To display the page, click **IPv4 Multicast** → **IGMP** → **Proxy Interface** → **Configuration Summary** in the navigation panel.

Figure 46-23. IGMP Proxy Configuration Summary

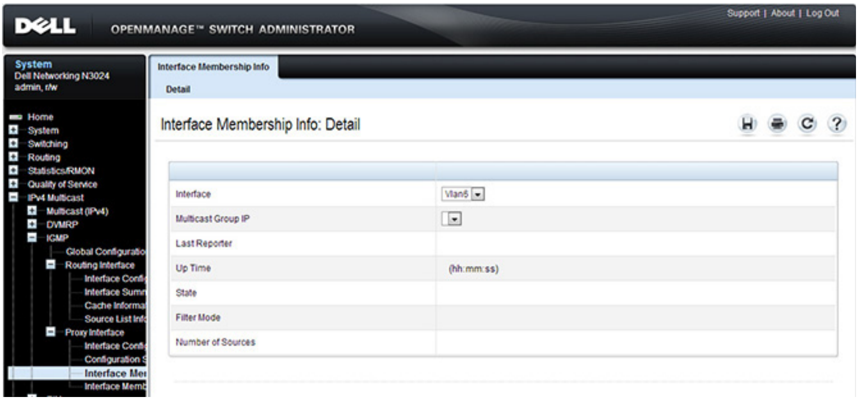


IGMP Proxy Interface Membership Info

Use the **Interface Membership Info** page to display interface membership data for a specific IP multicast group address. At least one VLAN routing interface must be configured for this page to display interface membership information, and it should not be an IGMP routing interface. Also, if no group membership reports have been received on the selected interface, no data displays on this page.

To display the page, click **IPv4 Multicast** → **IGMP** → **Proxy Interface** → **Interface Membership Info** in the navigation panel.

Figure 46-24. IGMP Proxy Interface Membership Info

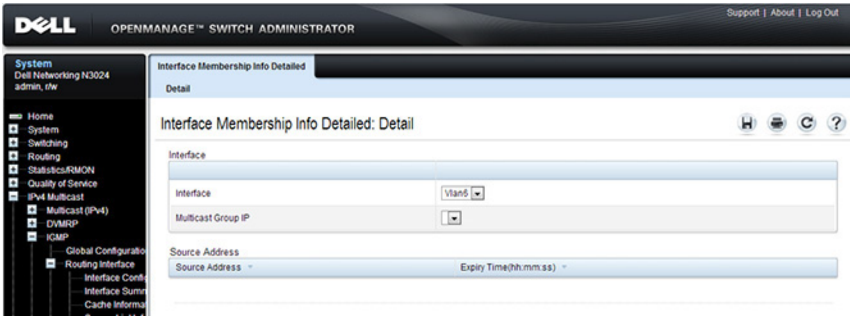


Detailed IGMP Proxy Interface Membership Information

Use the **Interface Membership Info Detailed** page to display detailed interface membership data. At least one VLAN routing interface must be configured before detailed interface membership information can be displayed, and it should not be an IGMP routing interface. Also, if no group membership reports have been received on the selected interface, then no data can be displayed.

To display the page, click **IPv4 Multicast** → **IGMP** → **Proxy Interface** → **Interface Membership Info Detailed** in the navigation panel.

Figure 46-25. IGMP Proxy Interface Membership Info Detailed



Configuring MLD and MLD Proxy (Web)

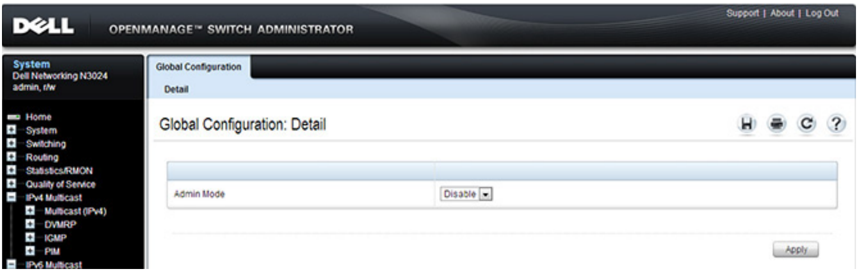
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring the MLD and MLD proxy features on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

MLD Global Configuration

Use the **Global Configuration** page to administratively enable and disable the MLD service.

To display the page, click **IPv6 Multicast** → **MLD** → **Global Configuration** in the navigation panel.

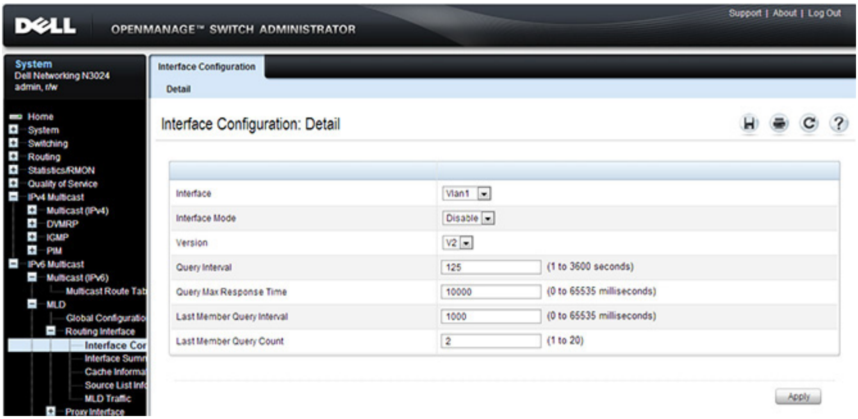
Figure 46-26. MLD Global Configuration



MLD Routing Interface Configuration

Use the **Interface Configuration** page to enable selected IPv6 router interfaces to discover the presence of multicast listeners, the nodes who wish to receive the multicast data packets, on its directly attached interfaces. To access this page, click **IPv6 Multicast** → **MLD** → **Routing Interface Interface Configuration** in the navigation panel.

Figure 46-27. MLD Routing Interface Configuration

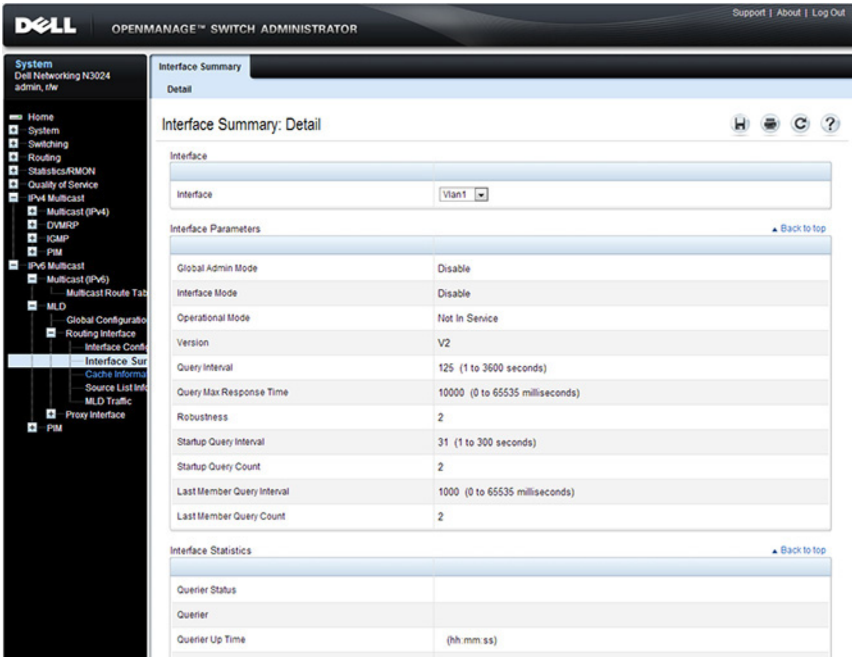


MLD Routing Interface Summary

Use the **Interface Summary** page to display information and statistics on a selected MLD-enabled interface. You must configure at least one IGMP VLAN routing interface to access this page.

To access this page, click **IPv6 Multicast** → **MLD** → **Routing Interface** → **Interface Summary** in the navigation panel.

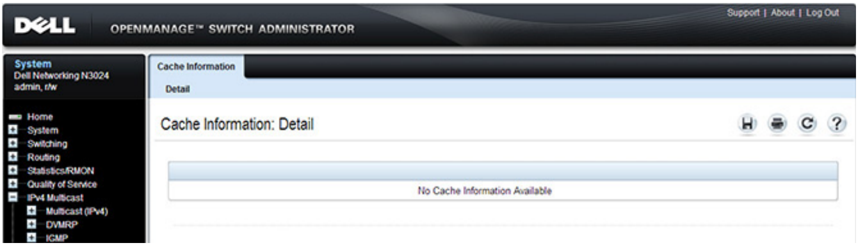
Figure 46-28. MLD Routing Interface Summary



MLD Routing Interface Cache Information

The **Interface Cache Information** page displays cache parameters and data for an IP multicast group address that has been reported to operational MLD routing interfaces. You must configure at least one MLD VLAN routing interface to access this page. Also, group membership reports must have been received on the selected interface in order for data to be displayed here. To access this page, click **IPv6 Multicast** → **MLD** → **Routing Interface** → **Cache Information** in the navigation panel.

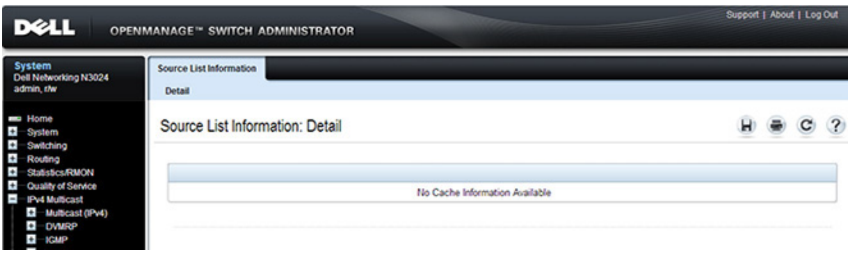
Figure 46-29. MLD Routing Interface Cache Information



MLD Routing Interface Source List Information

The **Interface Source List Information** page displays detailed membership information for an interface. You must configure at least one MLD VLAN routing interface to access this page. Also, group membership reports must have been received on the selected interface in order for data to be displayed here. To access this page, click **IPv6 Multicast** → **MLD** → **Routing Interface** → **Source List Information** in the navigation panel.

Figure 46-30. MLD Routing Interface Source List Information

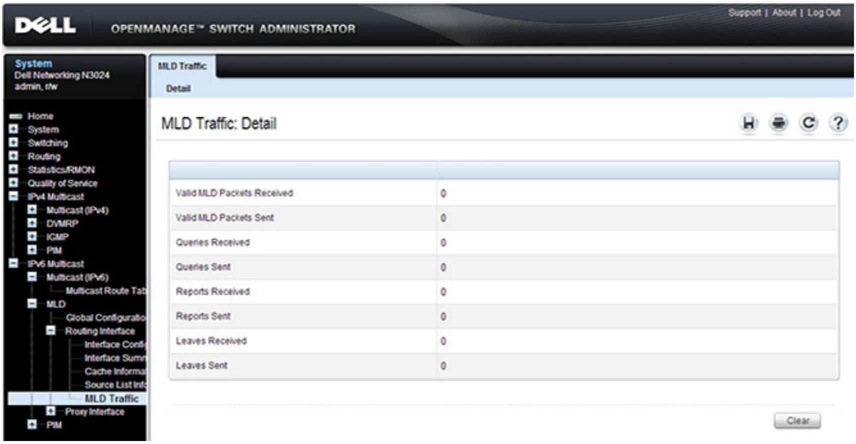


MLD Traffic

The **MLD Traffic** page displays summary statistics on the MLD messages sent to and from the router.

To access this page, click **IPv6 Multicast** → **MLD** → **Routing Interface** → **MLD Traffic** in the navigation panel.

Figure 46-31. MLD Traffic

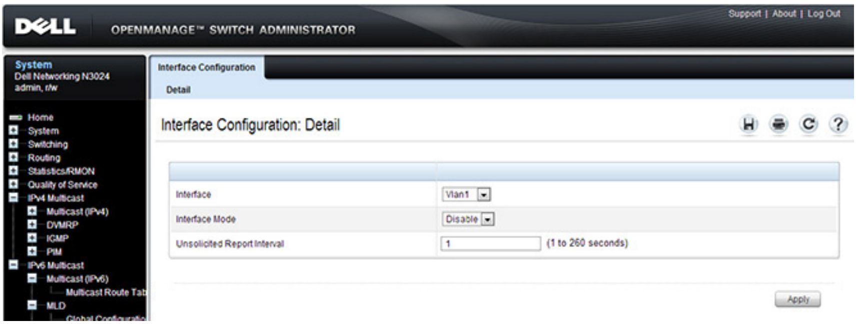


MLD Proxy Configuration

When you configure an interface in MLD proxy mode, it acts as a proxy multicast host that sends MLD membership reports on one VLAN interface for MLD Membership reports received on all other MLD-enabled VLAN routing interfaces.

Use the **Interface Configuration** page to enable and disable ports as MLD proxy interfaces. To display this page, click **IPv6 Multicast** → **MLD** → **Proxy Interface** → **Interface Configuration** in the navigation panel.

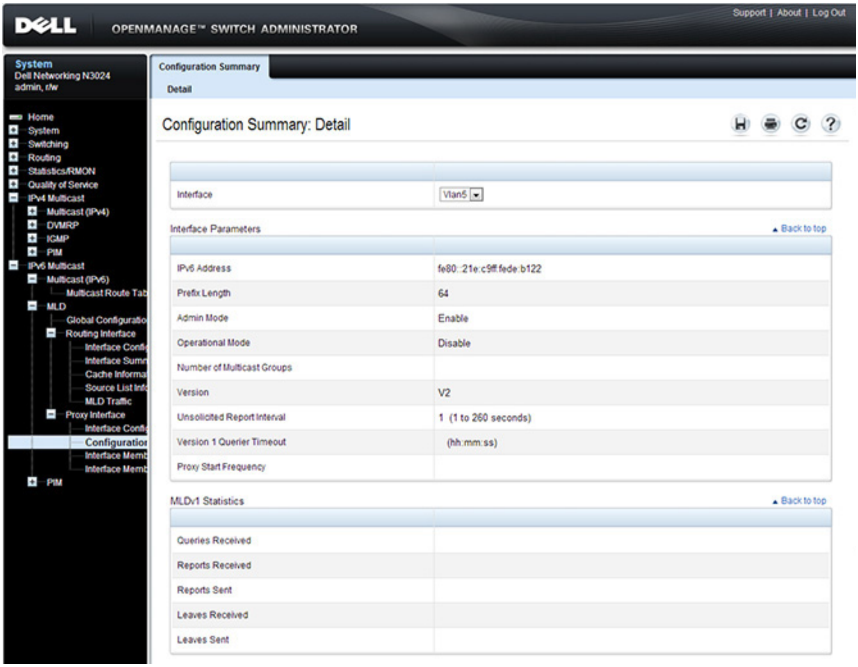
Figure 46-32. MLD Proxy Interface Configuration



MLD Proxy Configuration Summary

Use the **Configuration Summary** page to view configuration and statistics on MLD proxy-enabled interfaces. To display this page, click **IPv6 Multicast** → **MLD** → **Proxy Interface** → **Configuration Summary** in the navigation panel.

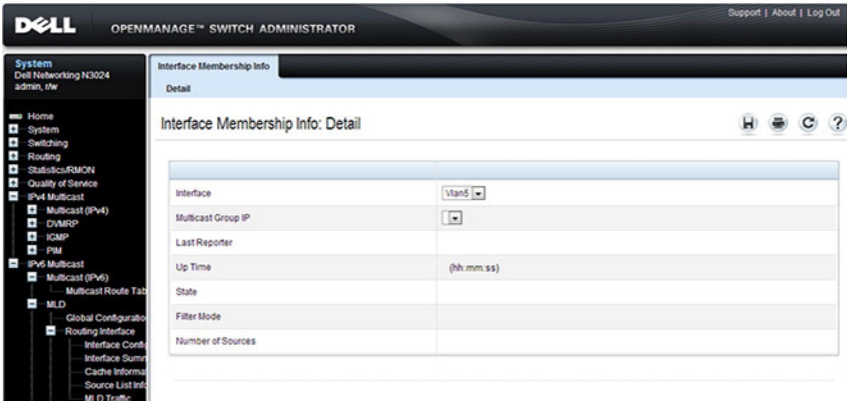
Figure 46-33. MLD Proxy Configuration Summary



MLD Proxy Interface Membership Information

The **Interface Membership Information** page lists each IP multicast group for which the MLD proxy interface has received membership reports. To display this page, click **IPv6 Multicast** → **MLD** → **Proxy interface** → **Interface Membership Info** in the navigation panel.

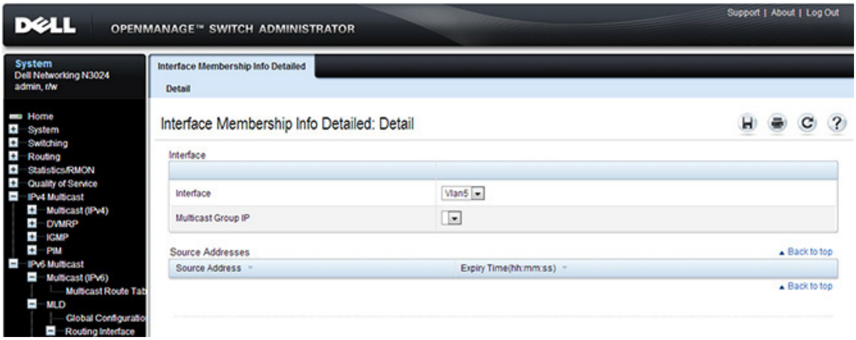
Figure 46-34. Interface Membership Information



Detailed MLD Proxy Interface Membership Information

The **Interface Membership Information Detailed** page provides additional information about the IP multicast groups for which the MLD proxy interface has received membership reports. To display this page, click **IPv6 Multicast** → **MLD** → **Proxy Interface** → **Interface Membership Info Detailed** in the navigation panel.

Figure 46-35. Interface Membership Information—Detailed



Configuring PIM for IPv4 and IPv6 (Web)

This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring PIM-SM and PIM-DM for IPv4 and IPv6 multicast routing on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click ? at the top of the page.

NOTE: The OpenManage Switch Administrator pages to configure IPv4 multicast routing and IPv6 multicast routing is very similar. The figures in this section show the IPv4 multicast configuration pages. To configure IPv6 multicast with PIM, use the pages available from the IPv6 Multicast → PIM menu.

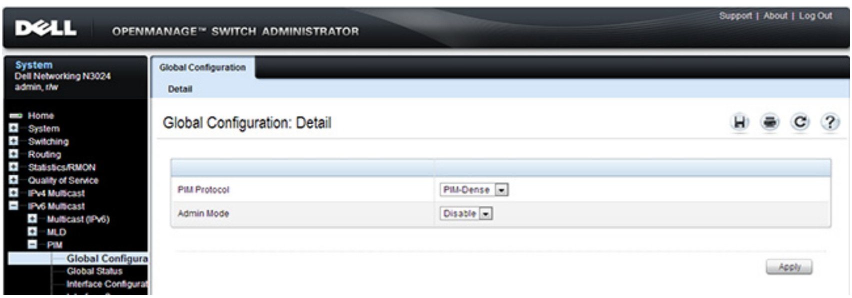
PIM Global Configuration

Use the **Global Configuration** page to configure the administrative status of PIM-DM or PIM-SM on the switch. It is strongly recommended that IGMP be enabled on any switch on which IPv4 PIM is enabled and MLD be enabled on any switch for which IPv6 PIM is enabled. This ensures that the multicast router behaves as expected.

The CLI behavior is different than the web interface. Enabling PIM on an IPv4 interface via the CLI automatically enables IGMP on the interface. Likewise, enabling PIM on an IPv6 interface via the CLI automatically enables MLD on the interface.

To display the page, click **IPv4 Multicast → PIM → Global Configuration** or **IPv6 Multicast → PIM → Global Configuration** in the navigation panel.

Figure 46-36. PIM-DM Global Configuration

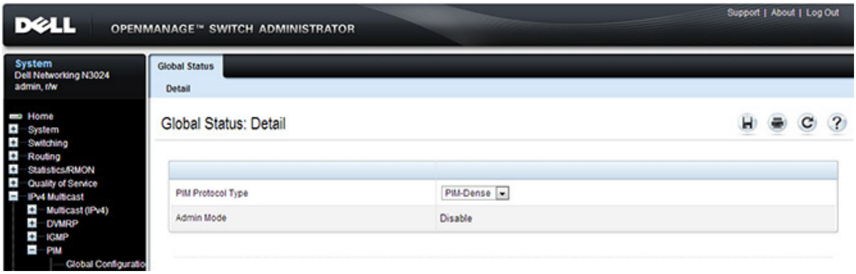


PIM Global Status

Use the **Global Status** page to view the administrative status of PIM-DM or PIM-SM on the switch.

To display the page, click **IPv4 Multicast** → **PIM** → **Global Status** or **IPv6 Multicast** → **PIM** → **Global Status** in the navigation panel.

Figure 46-37. PIM Global Status

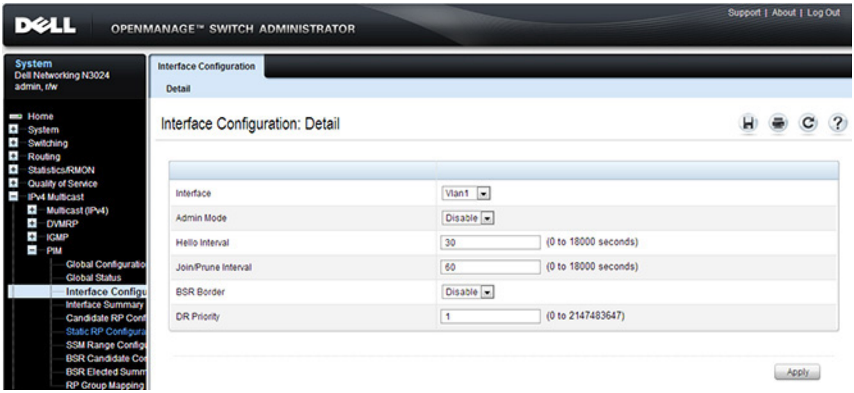


PIM Interface Configuration

Use the **Interface Configuration** page to configure specific VLAN routing interfaces with PIM.

To display the page, click **IPv4 Multicast** → **PIM** → **Interface Configuration** or **IPv6 Multicast** → **PIM** → **Interface Configuration** in the navigation panel.

Figure 46-38. PIM Interface Configuration

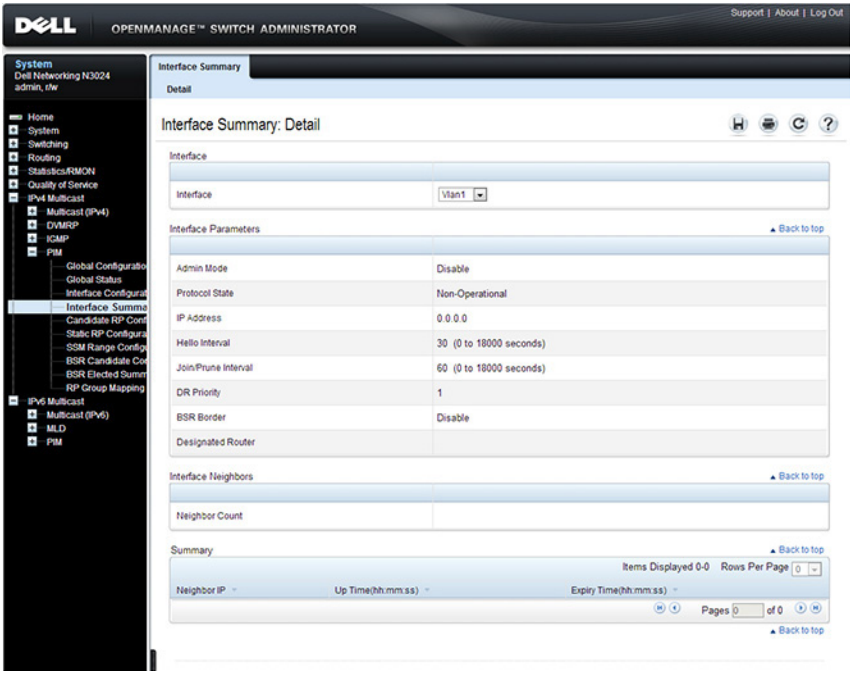


PIM Interface Summary

Use the **Interface Summary** page to display a PIM-enabled VLAN routing interface and its settings.

To display the page, click **IPv4 Multicast** → **PIM** → **Interface Summary** or **IPv6 Multicast** → **PIM** → **Interface Summary** in the navigation panel.

Figure 46-39. PIM Interface Summary

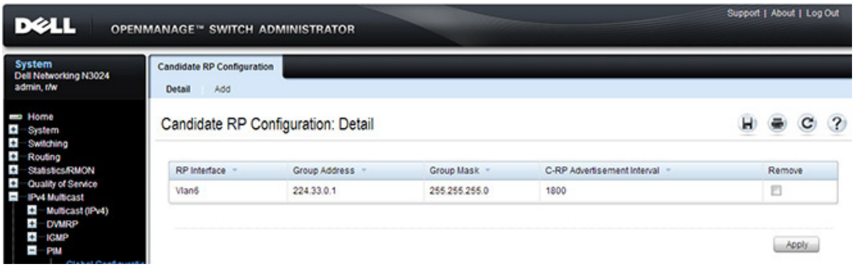


Candidate RP Configuration

The Candidate RP is configured on the **Add Candidate RP** page. Use the **Candidate RP Configuration** page to display and delete the configured rendezvous points (RPs) for each port using PIM.

To access the page, click **IPv4 Multicast** → **PIM** → **Candidate RP Configuration** or **IPv6 Multicast** → **PIM** → **Candidate RP Configuration**.

Figure 46-40. Candidate RP Configuration



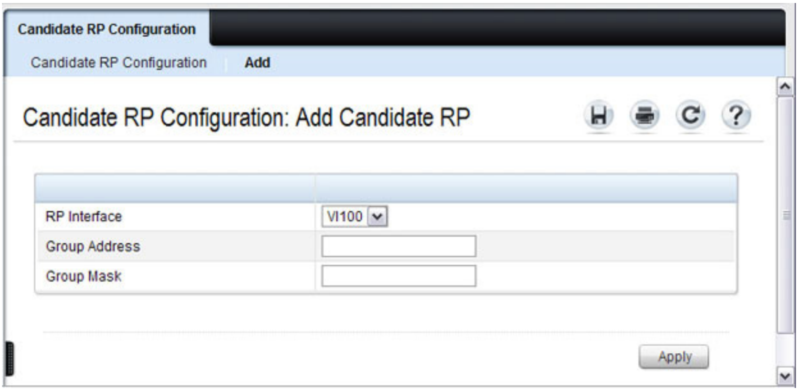
Adding a Candidate RP

To add PIM Candidate rendezvous points (RPs) for each IP multicast group:

1. Open the **Candidate RP Configuration** page.
2. Click **Add**.

   The **Add Candidate RP** page displays.

Figure 46-41. Add Candidate RP



3 Select the VLAN interface for which the Candidate RP is to be configured.

4 Enter the group address transmitted in Candidate-RP-Advertisements.

5 Enter the prefix length transmitted in Candidate-RP-Advertisements to fully identify the scope of the group which the router supports if elected as a Rendezvous Point.

6 Click **Apply Changes**.

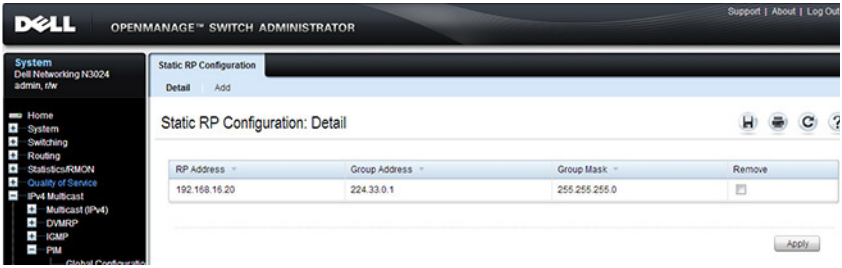
The new Candidate RP is added, and the device is updated.

Static RP Configuration

Use the **Static RP Configuration** page to display or remove the configured RP. The page also allows adding new static RPs by clicking the **Add** button. Only one RP address can be used at a time within a PIM domain. If the PIM domain uses the BSR to dynamically learn the RP, configuring a static RP is not required. However, the static RP can be configured to override any dynamically learned RP from the BSR.

To access the page, click **IPv4 Multicast** → **PIM** → **Static RP Configuration** or **IPv6 Multicast** → **PIM** → **Static RP Configuration**.

Figure 46-42. Static RP Configuration



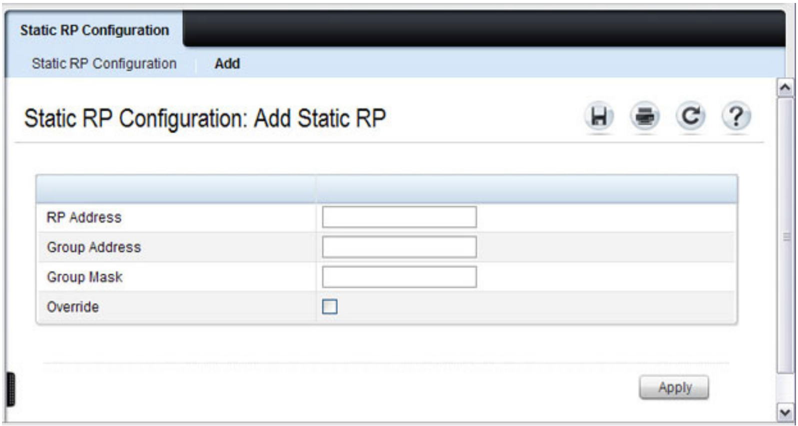
Adding a Static RP

To add a static RP for the PIM router.

1. Open the **Static RP Configuration** page.
2. Click **Add**.

   The **Add Static RP** page displays.

Figure 46-43. Add Static RP



3 Enter the IP address of the RP for the group range.

4 Enter the group address of the RP.

5 Enter the group mask of the RP.

6 Check the **Override** option to configure the static RP to override the dynamic (candidate) RPs learned for same group ranges.

7 Click **Apply**.

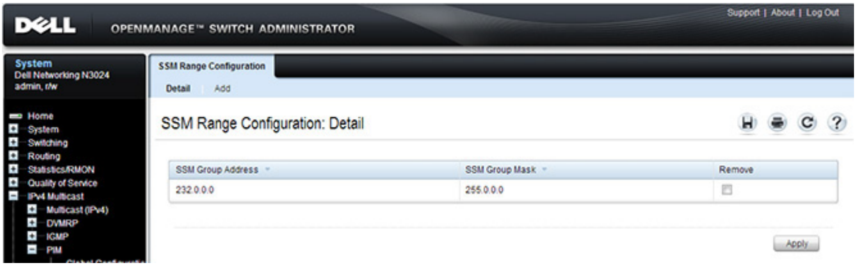
The new Static RP is added, and the device is updated.

SSM Range Configuration

Use this page to display or remove the Source Specific Multicast (SSM) group IP address and group mask for the PIM router.

To display the page, click **IPv4 Multicast** → **PIM** → **SSM Range Configuration** or **IPv6 Multicast** → **PIM** → **SSM Range Configuration**.

Figure 46-44. SSM Range Configuration



Adding an SSM Range

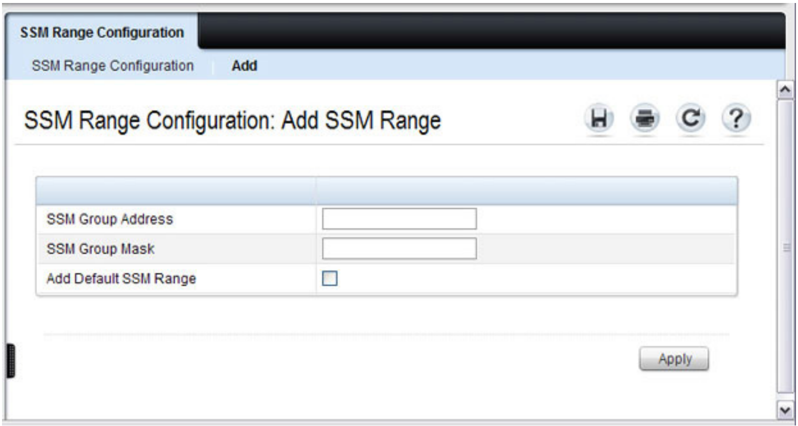
To add the Source-Specific Multicast (SSM) Group IP Address and Group Mask (IPv4) or Prefix Length (IPv6) for the PIM router:

1 Open the **SSM Range Configuration** page.

2 Click **Add**.

The **Add SSM Range** page displays.

Figure 46-45. Add SSM Range



3 Click the Add Default SSM Range check box to add the default SSM Range. The default SSM Range is 232.0.0.0/8 for IPv4 multicast and ff3x::/32 for IPv6 multicast.

4 Enter the SSM Group IP Address.

5 Enter the SSM Group Mask (IPv4) or SSM Prefix Length (IPv6).

6 Click **Apply**.

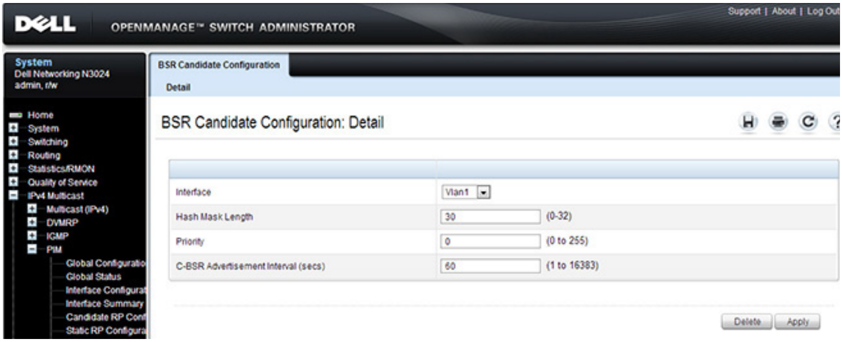
The new SSM Range is added, and the device is updated.

BSR Candidate Configuration

Use this page to configure information to be used if the interface is selected as a bootstrap router.

To display the page, click **IPv4 Multicast** → **PIM** → **BSR Candidate Configuration** or **IPv6 Multicast** → **PIM** → **BSR Candidate Configuration**.

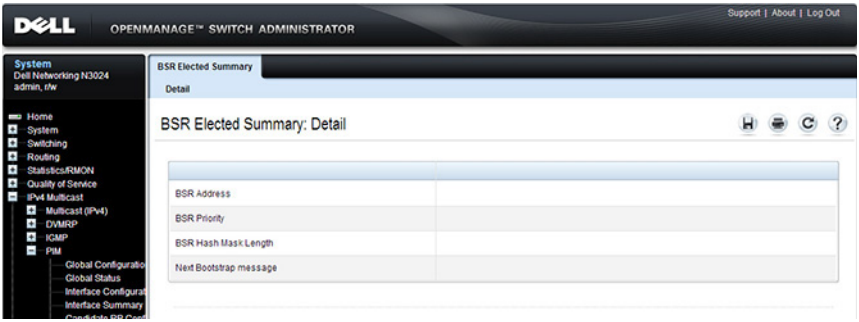
Figure 46-46. BSR Candidate Configuration



BSR Candidate Summary

Use this page to display information about the configured BSR candidates. To display this page, click **IPv4 Multicast** → **PIM** → **BSR Candidate Summary** or **IPv6 Multicast** → **PIM** → **BSR Elected Summary**.

Figure 46-47. BSR Elected Summary



Configuring DVMRP (Web)

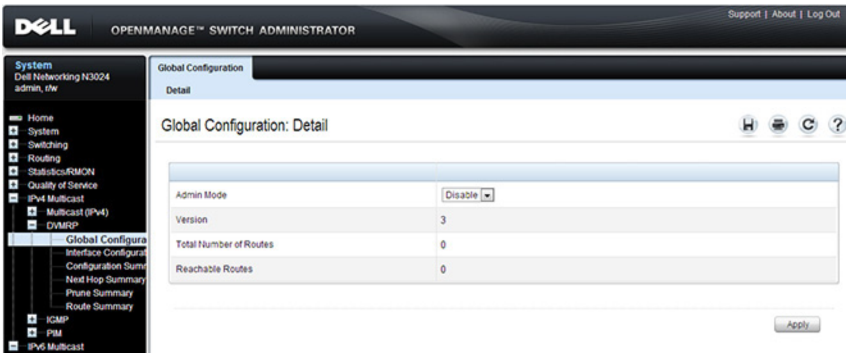
This section provides information about the OpenManage Switch Administrator pages for configuring and monitoring DVMRP on a Dell Networking N1500, N2000, N3000, and N4000 Series switches. For details about the fields on a page, click at the top of the page.

DVMRP Global Configuration

Use the **Global Configuration** page to configure global DVMRP settings. It is strongly recommended that IGMP be enabled on any switch on which DVMRP is enabled. The use cases for enabling DVMRP without IGMP are few, and enabling IGMP ensures that the multicast router behaves as expected.

To display the page, click **IPv4 Multicast → DVMRP → Global Configuration** in the navigation panel.

Figure 46-48. DVMRP Global Configuration

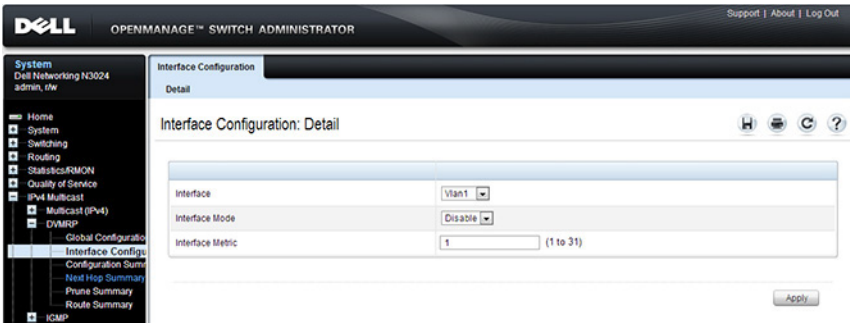


DVMRP Interface Configuration

Use the **Interface Configuration** page to configure a DVMRP VLAN routing interface. You must configure at least one router interface before you configure a DVMRP interface. Otherwise you see a message telling you that no router interfaces are available, and the configuration screen is not displayed. It is strongly recommended that IGMP be enabled on any interface on which DVMRP is enabled. This ensures that the multicast router behaves as expected.

To display the page, click **IPv4 Multicast** → **DVMRP** → **Interface Configuration** in the navigation panel.

Figure 46-49. DVMRP Interface Configuration

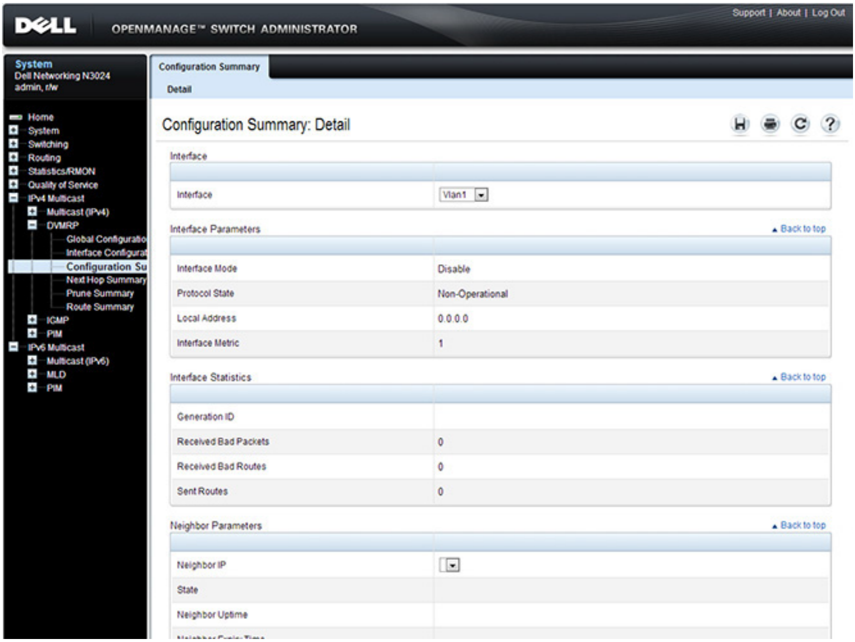


DVMRP Configuration Summary

Use the **Configuration Summary** page to display the DVMRP configuration and data for a selected interface. At least one VLAN routing interface must be configured before data can be displayed for a DVMRP interface. Otherwise, a message displays that no VLAN router interfaces are available, and the configuration summary screen is not displayed.

To display the page, click **IPv4 Multicast → DVMRP → Configuration Summary** in the navigation panel.

Figure 46-50. DVMRP Configuration Summary

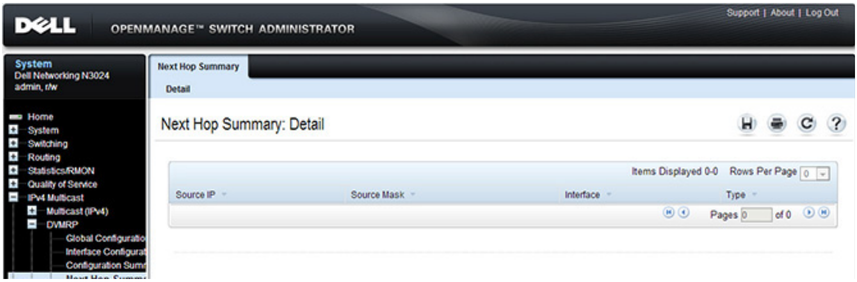


DVMRP Next Hop Summary

Use the **Next Hop Summary** page to display the next hop summary by Source IP.

To display the page, click **IPv4 Multicast** → **DVMRP** → **Next Hop Summary** in the navigation panel.

Figure 46-51. DVMRP Next Hop Summary

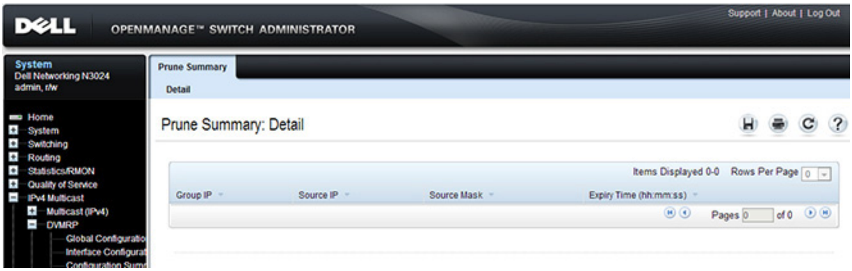


DVMRP Prune Summary

Use the **Prune Summary** page to display the prune summary by Group IP.

To display the page, click **IPv4 Multicast** → **DVMRP** → **Prune Summary** in the navigation panel.

Figure 46-52. DVMRP Prune Summary

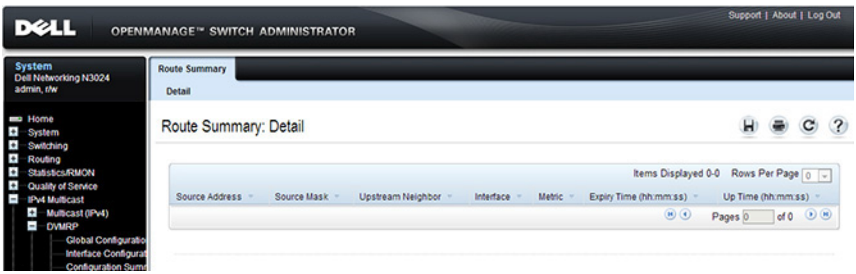


DVMRP Route Summary

Use the **Route Summary** page to display the DVMRP route summary.

To display the page, click **IPv4 Multicast** → **DVMRP** → **Route Summary** in the navigation panel.

Figure 46-53. DVMRP Route Summary



Configuring L3 Multicast Features (CLI)

This section provides information about the commands used for configuring general IPv4 multicast settings on the switch. For more information about the commands, see the *Dell Networking N1500, N2000, N3000, and N4000 Series Switches CLI Reference Guide* at www.dell.com/support.

Configuring and Viewing IPv4 Multicast Information

Beginning in Privileged EXEC mode, use the following commands to enable IPv4 multicast on the switch and to view and configure other general multicast settings.

Command	Purpose
configure	Enter global configuration mode.
ip multicast	Enable IPv4/IPv6 multicast on the switch.
ip mroute *source-address mask rpf-address preference*	Create a static multicast route for a source range. • *source-address* — The IP address of the multicast data source. • *mask* — The IP subnet mask of the multicast data source. • *rpf-address* — The IP address of the next hop towards the source. • *preference* — The cost of the route (Range: 1–255).
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ip mcast boundary *groupipaddr mask*	Add an administrative scope multicast boundary specified by the multicast group IP address (*groupipaddr*) and group IP subnet mask (*mask*) for which this multicast administrative boundary is applicable. The group IP address valid range is 239.0.0.0 to 239.255.255.255.
ip multicast ttl-threshold *ttlvalue*	Apply a Time to Live (TTL) value to the interface. The *ttlvalue* is the TTL threshold which is applied to the multicast data packets forwarded through the interface.

Command	Purpose
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ip multicast	View system-wide multicast information.
show ip mcast boundary {vlan *vlan-id* | **all}**	View all the configured administrative scoped multicast boundaries.
show ip mcast mroute {detail | summary}	View a summary or all the details of the multicast table.
show mac address-table multicast [count]	View information about the entries in the multicast address table.
show ip mcast mroute group *groupipaddr* **{detail | summary}**	View the multicast configuration settings such as flags, timer settings, incoming and outgoing interfaces, RPF neighboring routers, and expiration times of all the entries in the multicast mroute table containing the *groupipaddr* value.
show ip mcast mroute source *sourceipaddr* **{summary |** *groupipaddr***}**	View the multicast configuration settings such as flags, timer settings, incoming and outgoing interfaces, RPF neighboring routers, and expiration times of all the entries in the multicast mroute table containing the *sourceipaddr* or *sourceipaddr* | *groupipaddr* pair value(s).
show ip mcast mroute static [*sourceipaddr*]	View all the static routes configured in the static mcast table if it is specified or display the static route associated with the particular *sourceipaddr*.

Configuring and Viewing IPv6 Multicast Route Information

Beginning in Privileged EXEC mode, use the following commands to configure static IPv6 multicast routes on the switch and to view IPv6 multicast table information.

Command	Purpose
configure	Enter global configuration mode.
ip multicast	Enable IPv4/IPv6 multicast routing.
ipv6 mroute *source-address/prefix-length rpf-address* [**interface vlan** *vlan-id*] *preference*	Create a static multicast route for a source range. • *source-address/prefix-length* — The IPv6 address of the multicast data source. • *rpf-address* — The IPv6 address of the next hop towards the source. • *vlan-id*— If the *rpf-address* is a link-local address then the VLAN interface must also be specified. If the *rpf-address* is a global address, then specifying the VLAN interface is not required. • *preference* — The cost of the route (Range: 1–255).
exit	Exit to Privileged EXEC mode.
show ipv6 mroute {**detail** | **summary**}	View a summary or all the details of the multicast table.
show ipv6 mroute group *groupipaddr* {**detail** | **summary**}	View the multicast configuration settings such as flags, timer settings, incoming and outgoing interfaces, RPF neighboring routers, and expiration times of all the entries in the multicast mroute table containing the *groupipaddr* value.
show ipv6 mroute source *sourceipaddr {***summary** | *groupipaddr*}	View the multicast configuration settings such as flags, timer settings, incoming and outgoing interfaces, RPF neighboring routers, and expiration times of all the entries in the multicast mroute table containing the *sourceipaddr* or *sourceipaddr* | *groupipaddr* pair value(s).
show ipv6 mroute static [*sourceipaddr]*	View all the static routes configured in the static mcast table if it is specified or display the static route associated with the particular *sourceipaddr*.

Configuring and Viewing IGMP

Beginning in Privileged EXEC mode, use the following commands to configure IGMP on the switch and on VLAN routing interfaces and to view IGMP information.

Command	Purpose
configure	Enter global configuration mode.
ip multicast	Enable IPv4/IPv6 multicast routing.
ip igmp	Enable IGMP on the switch.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ip igmp	Enable IGMP on the interface.
ip igmp version *version*	Set the version of IGMP for an interface. The *version* variable can be 1, 2, or 3.
ip igmp robustness *robustness*	Configure the robustness that allows tuning of the interface, that is, tuning for the expected packet loss on a subnet. If a subnet is expected to have significant loss, the robustness variable may be increased for the interface. The range for *robustness* is 1–255.
ip igmp query-interval *seconds*	Configure the query interval for the specified interface. The query interval determines how fast IGMP Host-Query packets are transmitted on this interface. The range for *seconds* is 0–3600 seconds.
ip igmp query-max-response-time *seconds*	Configure the maximum response time interval for the specified interface. It is the maximum query response time advertised in IGMPv2 queries on this interface. The range for *seconds* is 0–25 seconds.
ip igmp startup-query-interval *seconds*	Set the interval between general queries sent at startup on the interface. The range for *seconds* is 0–300 seconds.

Command	Purpose
ip igmp startup-query-count *count*	Set the number of queries sent out on startup —at intervals equal to the startup query interval for the interface. The range for *count* is 1–20.
ip igmp last-member-query-interval *tenthsofseconds*	Configure the Maximum Response Time inserted in Group-Specific Queries which are sent in response to Leave Group messages. The range is 0–255 tenths of a second.
ip igmp last-member-query-count *count*	Set the number of Group-Specific Queries sent before the router assumes that there are no local members on the interface. The range for *count* is 1–20.
CTRL + Z	Exit to Privileged EXEC mode.
show ip igmp	View system-wide IGMP information.
show ip igmp interface [**vlan** *vlan-id*]	View IGMP information for all interfaces or for the specified interface.
show ip igmp interface stats [**vlan** *vlan-id*]	View IGMP statistics for all interfaces or for the specified interface.
show ip igmp groups [**interface vlan** *vlan-id*]	View the registered multicast groups on the interface.
show ip igmp membership	View the list of interfaces that have registered in any multicast group.

Configuring and Viewing IGMP Proxy

Beginning in Privileged EXEC mode, use the following commands to configure the upstream VLAN routing interface as an IGMP proxy. The IGMP proxy issues host messages on behalf of the hosts that have been discovered on IGMP-enabled interfaces. The upstream interface is the interface closest to the root multicast router, which should be running IGMP.

NOTE: Configure only the upstream interface as the IGMP proxy. IGMP should be enabled on all downstream interfaces. IP routing and IP multicast must be enabled on the switch for the IGMP proxy feature to operate.

Command	Purpose
configure	Enter global configuration mode.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ip igmp-proxy	Configure the interface as an IGMP proxy interface.
ip igmp-proxy reset-status	(Optional) Reset the host interface status parameters of the IGMP Proxy.
ip igmp-proxy unsolicit-rprt-interval *seconds*	Configure the unsolicited report interval for the IGMP proxy interface. The range for *seconds* is 0–260 seconds.
CTRL + Z	Exit to Privileged EXEC mode.
show ip igmp-proxy	View a summary of the host interface status parameters.
show ip igmp-proxy interface	View a detailed list of the host interface status parameters. This command displays information only when IGMP Proxy is operational.
show ip igmp-proxy groups	View a table of information about multicast groups that IGMP Proxy reported. This command displays information only when IGMP Proxy is operational.

Configuring and Viewing MLD

Beginning in Privileged EXEC mode, use the following commands to configure MLD on the switch and on VLAN routing interfaces and to view IGMP information.

Command	Purpose
configure	Enter global configuration mode.
ip multicast	Enable IPv4/IPv6 multicast routing.
ipv6 mld router	Enable MLD on the switch.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ipv6 mld router	Enable MLD on the interface.
ipv6 mld version *version*	Set the version of MLD for an interface. The *version* variable can be 1 or 2.
ipv6 mld query-interval *seconds*	Configure the query interval for the specified interface. The query interval determines how fast MLD Host-Query packets are transmitted on this interface. The range for *seconds* is 0–3600 seconds.
ipv6 mld query-max-response-time *seconds*	Configure the maximum response time interval for the specified interface. It is the maximum query response time advertised in MLD queries on this interface. The range for *seconds* is 0–25 seconds.
ipv6 mld last-member-query-interval *tenthsofseconds*	Set the last member query interval for the MLD interface, which is the value of the maximum response time parameter in the group-specific queries sent out of this interface. The range is 0–65535 milliseconds.
ipv6 mld last-member-query-count *count*	Set the number of listener-specific queries sent before the router assumes that there are no local members on the interface. The range for *count* is 1–20.
CTRL + Z	Exit to Privileged EXEC mode.

Command	Purpose
show ipv6 mld interface [**vlan** *vlan-id*]	View MLD information for all interfaces or for the specified interface.
show ipv6 mld interface stats [**vlan** *vlan-id*]	View MLD statistics for all interfaces or for the specified interface.
show ipv6 mld groups [**interface vlan** *vlan-id*]	View the registered multicast groups on the interface.
show ipv6 mld membership	View the list of interfaces that have registered in any multicast group.

Configuring and Viewing MLD Proxy

Beginning in Privileged EXEC mode, use the following commands to configure the upstream VLAN routing interface as an MLD proxy. The MLD proxy issues host messages on behalf of the hosts that have been discovered on the downstream MLD-enabled interfaces. The upstream interface is the interface closest to the root multicast router, which should be running IGMP.

NOTE: Configure only the upstream interface as the MLD proxy. MLD should be enabled on all downstream interfaces. IPv6 routing must be enabled on the switch for the MLD proxy feature to operate.

Command	Purpose
configure	Enter global configuration mode.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ipv6 mld-proxy	Configure the interface as an MLD proxy interface.
ipv6 mld-proxy reset-status	(Optional) Reset the host interface status parameters of the MLD Proxy.
ipv6 igmp-proxy unsolicit-rprt-interval *seconds*	Configure the unsolicited report interval for the MLD proxy interface. The range for *seconds* is 0–260 seconds.
CTRL + Z	Exit to Privileged EXEC mode.

Command	Purpose
show ipv6 mld-proxy	View a summary of the host interface status parameters.
show ipv6 mld-proxy interface	View a detailed list of the host interface status parameters. This command displays information only when MLD Proxy is operational.
show ipv6 mld-proxy groups	View a table of information about multicast groups that MLD Proxy reported. This command displays information only when MLD Proxy is operational.

Configuring and Viewing PIM-DM for IPv4 Multicast Routing

Beginning in Privileged EXEC mode, use the following commands to configure PIM-DM for IPv4 multicast routing on the switch and on VLAN routing interfaces and to view PIM-DM information.

Command	Purpose
configure	Enter global configuration mode.
ip routing	Enable ip routing. Routing is required for PIM to calculate where to prune the multicast trees.
ip pim dense	Enable PIM-DM on the switch.
ip igmp	Enable IGMP. IGMP is required for PIM to operate properly.
ip multicast	Enable IPv4/IPv6 multicast routing.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ip pim	Enable PIM-DM on the interface.
ip igmp	Enable IGMP on the interface. IGMP is required for proper operation of PIMDM
ip pim hello-interval *seconds*	Specify the number of seconds (range: 0–65535) to wait between sending PIM hello messages on the interface.
exit	Exit to Privileged EXEC mode.
show ip pim	View system-wide PIM information.

Command	Purpose
show ip pim interface vlan *vlan-id*	View the PIM-DM information for the specified interface.
show ip pim neighbor [**interface vlan** *vlan-id* | **all**]	View a summary or all the details of the multicast table.

Configuring and Viewing PIM-DM for IPv6 Multicast Routing

Beginning in Privileged EXEC mode, use the following commands to configure PIM-DM for IPv6 multicast routing on the switch and on VLAN routing interfaces and to view PIM-DM information.

Command	Purpose
configure	Enter global configuration mode.
ip routing	Enable IP routing. Routing is required for PIM operation.
ipv6 unicast-routing	Enable IPv6 routing. IPv6 routing is required for the operation of PIM.
ipv6 pim dense	Enable PIM-DM on the switch.
ip multicast	Enable IPv6/IPv6 multicast routing.
ip igmp	Enable IGMP. IGMP is required for PIM to operate properly.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ipv6 pim	Enable PIM on the interface.
ipv6 enable	Enable IPv6 on the VLAN.
ipv6 mld router	Enable MLD on the VLAN. MLD is required for PIM.
ipv6 pim hello-interval *seconds*	Specify the number of seconds (range: 0–65535) to wait between sending PIM hello messages on the interface.
exit	Exit to Privileged EXEC mode.
show ipv6 pim	View system-wide PIM information.

Command	Purpose
show ipv6 pim interface vlan *vlan-id*	View the PIM information for the specified interface.
show ipv6 pim neighbor [**interface vlan** *vlan-id* | **all**]	View a summary or all the details of the multicast table.

Configuring and Viewing PIM-SM for IPv4 Multicast Routing

Beginning in Privileged EXEC mode, use the following commands to configure PIM-SM for IPv4 multicast routing on the switch and on VLAN routing interfaces and to view PIM-SM information.

Command	Purpose
configure	Enter global configuration mode.
ip routing	Enable ip routing. Routing is required for PIM operation.
ip pim sparse	Enable PIM-SM as the multicast routing protocol on the switch.
ip igmp	Enable IGMP.
ip multicast	Enable IPv4/IPv6 multicast routing.
ip pim bsr-candidate vlan *vlan-id hash-mask-length* [*priority*] [**interval** *interval*]	Configure the switch to announce its candidacy as a bootstrap router (BSR). • *vlan-id* — A valid VLAN ID. • *hash-mask-length* — The length of a mask that is to be ANDed with the group address before the hash function is called. All groups with the same seed hash correspond to the same RP. For example, if this value is 24, only the first 24 bits of the group addresses matter. This allows you to get one RP for multiple groups. (Range 0–32 bits). • *priority* — The priority of the candidate BSR. The BSR with the higher priority is preferred. If the priority values are the same, the router with the higher IP address is the BSR. (Range 0–255). • *interval* — (Optional) Indicates the BSR candidate advertisement interval. The range is from 1 to 16383 seconds. The default value is 60 seconds.

Command	Purpose
ip pim rp-candidate vlan *vlan-id group-address group-mask* [**interval** *interval*]	Configure the router to advertise itself to the BSR router as a PIM candidate Rendezvous Point (RP) for a specific multicast group range. • *vlan-id* — A valid VLAN ID. • *group-address* — Group IP address supported by RP. • *group-mask* — Group subnet mask for group address. • *interval* — (Optional) Indicates the RP candidate advertisement interval. The range is from 1 to 16383 seconds. The default value is 60 seconds.
ip pim rp-address *rp-address group-address group-mask* [**override**]	(Optional) Statically configure the RP address for one or more multicast groups. Only one RP address can be used at a time within a PIM domain The optional keyword **override** indicates that if there is a conflict, the RP configured with this command prevails over the RP learned by BSR.
ip pim ssm {**default** | *group-address group-mask*}	Define the Source Specific Multicast (SSM) range of IP multicast addresses. • **default** — Defines the SSM range access list to 232.0.0.0/8. • *group-address group-mask* — defines the SSM range.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ip pim hello-interval *seconds*	Specify the number of seconds (range: 0–65535) to wait between sending PIM hello messages on the interface.
ip pim bsr-border	Prevent bootstrap router (BSR) messages from being sent or received through the interface.
ip pim dr-priority *priority*	Set the priority value for which a router is elected as the designated router (DR). The election priority range is 0–2147483647.
ip pim join-prune-interval *interval*	Configure the interface join/prune interval for the PIM-SM router. The interval range is 0–18000 seconds.

Command	Purpose
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ip pim	View system-wide PIM information.
show ip pim interface vlan *vlan-id*	View the PIM information for the specified interface.
show ip pim neighbor [**interface vlan** *vlan-id* | **all**]	View a summary or all the details of the multicast table.
show ip pim rp-hash *groupaddr*	View the RP router being selected for the specified multicast group address from the set of active RP routers. The RP router for the group is selected by using a hash algorithm.
show ip pim bsr-router [**candidate** | **elected**]	View the bootstrap router (BSR) information.
show ip pim rp mapping	View group-to-RP mappings of which the router is aware (either configured or learned from the BSR)

Configuring and Viewing PIM-SM for IPv6 Multicast Routing

Beginning in Privileged EXEC mode, use the following commands to configure PIM-SM for IPv6 multicast routing on the switch and on VLAN routing interfaces and to view PIM-SM information.

Command	Purpose
configure	Enter global configuration mode.
ip routing	Enable IP routing. Routing is required for PIM operation.
ipv6 unicast-routing	Enable IPv6 routing. IPv6 routing is required for IPv6 PIM.
ipv6 pim sparse	Enable PIM-SM as the multicast routing protocol on the switch.
ip mld router	Enable MLD. MLD is required for the proper operation of IPv6 PIM.
ip multicast	Enable IPv4/IPv6 multicast.

Command	Purpose
ipv6 pim bsr-candidate vlan *vlan-id hash-mask-length* [*priority*] [**interval** *interval*]	Configure the switch to announce its candidacy as a bootstrap router (BSR) • *vlan-id* — A valid VLAN ID. • *hash-mask-length* — The length of a mask that is to be ANDed with the group address before the hash function is called. All groups with the same seed hash correspond to the same RP. For example, if this value is 24, only the first 24 bits of the group addresses matter. This allows you to get one RP for multiple groups. (Range 0–32 bits). • *priority* — The priority of the candidate BSR. The BSR with the higher priority is preferred. If the priority values are the same, the router with the higher IPv6 address is the BSR. (Range 0–255). • *interval* — (Optional) Indicates the BSR candidate advertisement interval. The range is from 1 to 16383 seconds. The default value is 60 seconds.
ipv6 pim rp-candidate vlan *vlan-id group-address/prefix-length* [**interval** *interval*]	Configure the router to advertise itself to the BSR router as a PIM candidate Rendezvous Point (RP) for a specific multicast group range. • *vlan-id* — A valid VLAN ID. • *group-address/prefix-length*— Group IPv6 address and prefix length supported by RP. • *interval* — (Optional) Indicates the RP candidate advertisement interval. The range is from 1 to 16383 seconds. The default value is 60 seconds.
ipv6 pim rp-address *rp-address group-address/prefix-length* [**override**]	(Optional) Statically configure the RP address for one or more multicast groups. Only one RP address can be used at a time within a PIM domain The optional keyword **override** indicates that if there is a conflict, the RP configured with this command prevails over the RP learned by BSR.

Command	Purpose
ipv6 pim ssm {**default** | *group-address/prefix-length*}	Define the Source Specific Multicast (SSM) range of IPv6 multicast addresses. • **default** — Defines the SSM range access list to FF3x::/32. • *group-address/prefix-length* — defines the SSM range.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN.
ipv6 pim	Enable PIM on the VLAN.
ipv6 enable	Enable IPv6 on the VLAN.
ipv6 mld router	Enable MLD on the VLAN. MLD is required for IPv6 PIM.
ipv6 pim hello-interval *seconds*	Specify the number of seconds (range: 0–65535) to wait between sending PIM hello messages on the interface.
ipv6 pim bsr-border	Prevent bootstrap router (BSR) messages from being sent or received through the interface.
ipv6 pim dr-priority *priority*	Set the priority value for which a router is elected as the designated router (DR). The election priority range is 0–2147483647.
ipv6 pim join-prune-interval *interval*	Configure the interface join/prune interval for the PIM-SM router. The interval range is 0–18000 seconds.
exit	Exit to Global Config mode.
exit	Exit to Privileged EXEC mode.
show ipv6 pim	View system-wide PIM information.
show ipv6 pim interface vlan *vlan-id*	View the PIM information for the specified interface.
show ipv6 pim neighbor [**interface vlan** *vlan-id* | **all**]	View a summary or all the details of the multicast table.

Command	Purpose
show ipv6 pim rp-hash *groupaddr*	View the RP router being selected for the specified multicast group address from the set of active RP routers. The RP router for the group is selected by using a hash algorithm.
show ipv6 pim bsr-router	View the bootstrap router (BSR) information.
show ipv6 pim rp mapping	View group-to-RP mappings of which the router is aware (either configured or learned from the BSR)

Configuring and Viewing DVMRP Information

Beginning in Privileged EXEC mode, use the following commands to configure DVMRP on the switch and on VLAN routing interfaces and to view DVMRP information.

Command	Purpose
configure	Enter global configuration mode.
ip dvmrp	Enable DVMRP on the switch.
ip multicast	Enable IP multicast.
interface vlan *vlan-id*	Enter Interface Configuration mode for the specified VLAN routing interface.
ip dvmrp	Enable DVMRP on the interface.
ip dvmrp metric *metric*	Configure the metric (range: 1–31) for an interface. This value is used in the DVMRP messages as the cost to reach this network.
exit	Exit to Privileged EXEC mode.
show ip dvmrp interface vlan *vlan-id*]	View the multicast information for the specified interface.
show ip dvmrp neighbor	View neighbor information for DVMRP.
show ip dvmrp nexthop	View the next hop information on outgoing interfaces for routing multicast datagrams.
show ip dvmrp prune	View the table that lists the router's upstream prune information
show ip dvmrp route	View the multicast routing information for DVMRP.

L3 Multicast Configuration Examples

This section contains the following configuration examples:

- Configuring Multicast VLAN Routing With IGMP and PIM-SM
- Configuring DVMRP

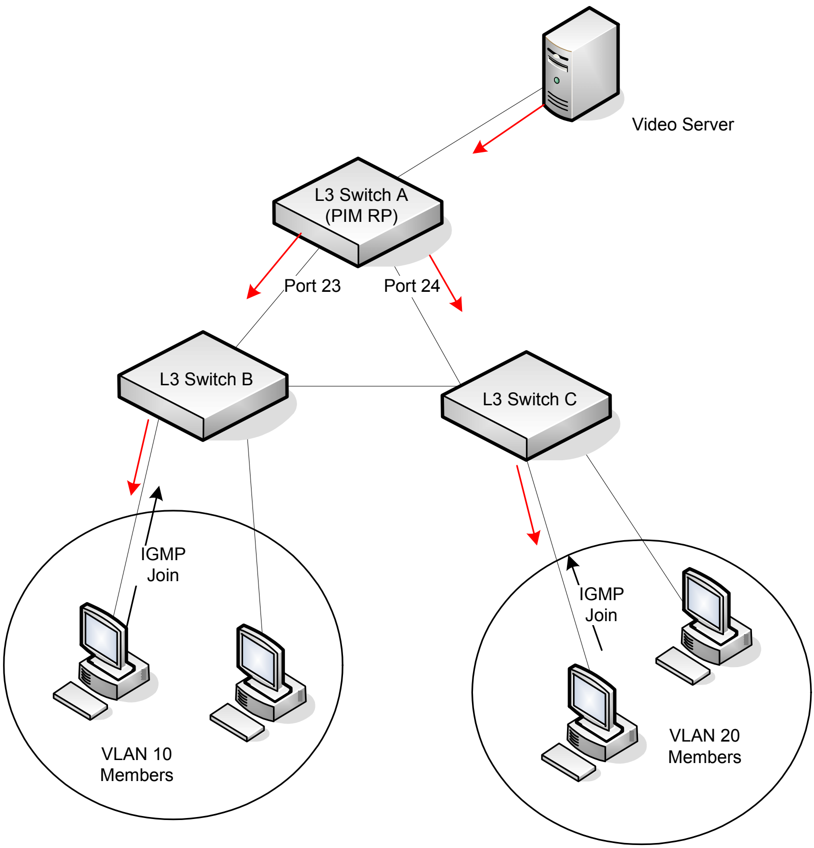
Configuring Multicast VLAN Routing With IGMP and PIM-SM

This example describes how to configure a Dell Networking N-Series switch with two VLAN routing interfaces that route IP multicast traffic between the VLANs. PIM and IGMP are enabled on the switch and interfaces to manage the multicast routing. VLAN 10 is statically configured as the RP for the multicast group.

> **NOTE:** PIM does not require OSPF specifically; static routing or RIP could also be configured for unicast routing.

The configuration in this example takes place on L3 switch A shown in Figure 46-54. The red arrows indicate the path that multicast traffic takes. L3 Switch A is configured as the RP for the PIM domain, so it is in charge of sending the multicast stream to L3 Switch B and L3 Switch C, and these switches forward the multicast data to the hosts that have requested to receive the data.

Figure 46-54. IPv4 Multicast VLAN Routing



In addition to multicast configuration, this example includes commands to configure STP and OSPF on L3 Switch A. STP is configured on the ports that connects the switch to other switches. OSPF is configured to route unicast traffic between the VLANs and PIM is enabled to rout multicast traffic between the two VLANs. Since IGMP snooping is enabled by default on all VLANs, no commands to enable it appear in the example below.

To configure Switch A:

1 Create the two VLANs. IGMP/MLD Snooping is disabled globally.

```
console#configure
console(config)#no ip igmp snooping
console(config)#no ipv6 mld snooping
console(config)#vlan 10,20
console(config-vlan10,20)#exit
```

2 Configure port 23 and 24 as trunk ports.

```
console(config)#interface te1/0/23
console(config-if-Te1/0/23)#switchport mode trunk
console(config-if-Te1/0/23)#switchport trunk allowed vlan remove 10
console(config-if-Te1/0/23)#exit

console(config)#interface te1/0/24
console(config-if-Te1/0/24)#switchport mode trunk
console(config-if-Te1/0/24)#switchport trunk allowed vlan remove 20
console(config-if-Te1/0/24)#exit
```

3 Enable routing on the switch and configure the OSPF router ID.

```
console(config)#ip routing
console(config)#router ospf
console(config-router)#router-id 3.3.1.1
console(config-router)#exit
```

4 Configure VLAN 10 as a VLAN routing interface and specify the OSPF area. When you assign an IP address to the VLAN, routing is automatically enabled.

```
console(config)#interface vlan 10
console(config-if-vlan10)#ip address 192.168.10.4 255.255.255.0
console(config-if-vlan10)#ip ospf area 0
```

5 Enable IGMPv2 and PIM-SM on the VLAN routing interface.

```
console(config-if-vlan10)#ip igmp
console(config-if-vlan10)#ip igmp version 2
console(config-if-vlan10)#ip pim
console(config-if-vlan10)#exit
```

6 Configure VLAN 20 as a VLAN routing interface and specify the OSPF area.

```
console(config)#interface vlan 20
console(config-if-vlan20)#ip address 192.168.20.4 255.255.255.0
console(config-if-vlan20)#ip ospf area 0
```

7 Enable IGMPv2 and PIM-SM on the VLAN routing interface.

```
console(config-if-vlan20)#ip igmp
console(config-if-vlan10)#ip igmp version 2
console(config-if-vlan20)#ip pim
```

```
console(config-if-vlan20)#exit
```

8 Globally enable IP multicast, IGMP, and PIM-SM on the switch.

```
console(config)#ip multicast
console(config)#ip igmp
console(config)#ip pim sparse
```

9 Configure VLAN 10 as the RP and specify the range of multicast groups for PIM-SM to control. The 239.9.x.x address is chosen as it is a locally administered address that maps to MAC addresses that do not conflict with control plane protocols.

```
console(config)#ip pim rp-address 192.168.10.4 239.9.0.0 255.255.0.0
```

Configuring DVMRP

The following example configures two DVMRP interfaces on the switch to enable inter-VLAN multicast routing.

To configure the switch:

1 Globally enable IP routing and IP multicast.

```
console#configure
console(config)#ip routing
console(config)#ip multicast
```

2 Globally enable IGMP so that this L3 switch can manage group membership information for its directly-connected hosts. Enabling IGMP is not required if there are no directly-connected hosts; however, it is recommended that it be enabled to ensure correct operation of multicast routing. Disable IGMP/MLD snooping.

```
console(config)#ip igmp
console(config)#no ip igmp snooping
console(config)#no ipv6 mld snooping
```

3 Globally enable DVMRP.

```
console(config)#ip dvmrp
```

4 Enable DVMRP and IGMP on VLAN routing interfaces 10 and 20.

```
console(config)#interface vlan 10
console(config-if-vlan10)#ip address 192.168.10.1 255.255.255.0
console(config-if-vlan10)#ip dvmrp
console(config-if-vlan10)#ip igmp
console(config-if-vlan10)#exit

console(config)#interface vlan 20
console(config-if-vlan20)#ip address 192.168.20.1 255.255.255.0
console(config-if-vlan20)#ip dvmrp
console(config-if-vlan20)#ip igmp
console(config-if-vlan20)#exit
```

47

Audio Video Bridging

Dell Networking N4000 Series Switches

NOTE: This feature is available on Dell Networking N4000 Series switches only.

Overview

Audio Video Bridging (AVB) is a suite of protocols for reserving resources in the network to facilitate an end-to-end time-sensitive traffic flow. AVB uses the following protocols:

- IEEE 802.1AS—Measures wire propagation time for precise synchronization.
- Multiple VLAN Registration Protocol (MVRP)—Replaces the role of GVRP in dynamic VLAN creation. MVRP propagates dynamic VLAN information to participating bridges. The participating bridges register (or withdraw) VLAN ID registrations for propagation of AVB streams. If a VLAN ID is dynamically registered on a bridge port, the bridge forwards frames for that VLAN ID on the port.
- Multiple MAC Registration Protocol—Replaces the role of GMRP in dynamic (M)FDB entry creation. MMRP propagates the association of a MAC address (associated with an AVB stream) to a VLAN. This helps to determine to what part of a network a given MAC address needs to be transmitted. If a MAC address is registered on a bridge port by MMRP, the bridge forwards frames addressed to that MAC address on the port.
- Multiple Stream Reservation Protocol (MSRP)—Controls bandwidth reservation of the audio-video stream. MSRP calculates the reservation size, protects port queues by remapping conflicting traffic, and shapes traffic according to the reservation criteria. MSRP defines the payloads and rules for propagating registrations for talkers (stream sources) and listeners (stream destinations).

AVB data is usually multicast traffic, not necessarily in standard IPv4 multicast format. For example, the IEEE 1722 Audio Video Transport Protocol uses MAC addresses in the following ranges:

Address Range	Function
91:E0:F0:00:00:00–91:E0:F0:00:FD:FF	Dynamic Allocation Pool
91:E0:F0:00:FE:00–91:E0:F0:00:FE:FF	Locally administered pool
91:E0:F0:00:FF:00–91:E0:F0:00:FF:FF	Reserved pool

MMRP, MVRP and MSRP share a common framework that provides services to the individual protocols. The common framework is the Multiple Registration Protocol (MRP). MRP allows participants in an MRP application to register attributes with other participants in a Bridged LAN. Each MRP participant maintains:

- Registrar and Applicant state machine for each attribute of interest
- LeaveAll and PeriodicTransmission state machine support for the participant

MRP propagates the attribute registrations throughout the AVB network. AVB network participants are aware of all other participants and their attribute registrations. VLAN bridges that do not support MRP forward received MRPDUs on all ports that are in forwarding state.

MRP implements as many MRP Attribute Protocol (MAP) contexts as there are MSTP instances. Within each MAP context, one participant is created for each bridge port and for each MRP application (MMRP, MSRP or MVRP). The AVB protocol family implements the following attributes:

Protocol	Attribute
MVRP	A VLAN identifier
MMRP	A VLAN MAC address association
MSRP	A domain attribute is a logical traffic class (A or B).
	A talker attribute identifying a stream from a transmitting station.
	A listener attribute identify a steam request from receiving station.

The Dell Networking N4000 AVB feature supports:

- IEEE 802.1ak (D8.0) Multiple Registration Protocol (MRP)
  - MVRP—Multiple VLAN Registration Protocol
  - MMRP—Multiple Multicast Registration Protocol
- IEEE 802.1as (D7.6)
  - Single unit only. No support on stack members.
  - Clock Master
  - Timing propagation
- IEEE 802.1Qat (D6.1)
  - Stream Reservation Protocol (MSRP)
- IEEE 802.1Qav (D7.0)
  - Forwarding and Queuing Enhancements for Time-Sensitive Streams
- IEEE 802.1ba (D2.5)
  - AVB Network Operation

MSRP

MSRP provides a mechanism for the reservation of resources for specific traffic streams traversing a bridged network.

MSRP categorizes AVB devices into talkers (stream sources) and listeners (stream destinations). An AVB device may be both a talker and a listener.

MSRP operates via several types of announcements (MRP declarations). The announcements are propagated throughout the AVB network. Announcements may occur in any order except when noted otherwise.

Listeners announce that they wish to receive a stream by sending an asking declaration. The AV bridges and talkers register this declaration.

Talkers announce that they are able to supply a stream with certain characteristics by sending an offering declaration. This declaration includes the ranking of the reservation's importance, its bandwidth, Ethernet priority level, and destination MAC address. Optionally, a higher-layer stream ID is also included. If a bridge has sufficient resources to propagate the stream, it forwards the declaration. If a bridge has insufficient resources to supply the stream, it converts the declaration into a failed declaration. In addition, the bridge converts the asking declaration into an asking failed declaration and forwards that declaration to the talker.

Listeners announce that they have configured their queues for handling a stream and that transmission can begin by sending a ready declaration. Only if a bridge has registered both the offering and the ready declarations will it configure its output queues on the port from which the ready declaration was received and propagate the ready declaration.

When a ready declaration is registered by the talker, it will start sending the data stream.

Regardless of the order in which the registrations occur, it is the presence of both the offering registration on an input port and the ready registration on an output port that triggers the bridge to configure the output port and allow propagation of the AVB stream from the talker.

MVRP

MVRP provides a mechanism for the declaration of dynamic registration of VLANs and propagation of VLAN information over a bridged network. The propagation of VLAN information via MRP allows MVRP-aware devices to dynamically establish and update the set of VLANs that are active on network devices and the ports through which those devices can be reached.

With MVRP both end stations and bridges may issue and revoke VLAN membership declarations. The effect of issuing such declaration is that each MVRP Participant that receives the declaration will create or update a dynamic VLAN Registration entry in the Filtering Database to indicate whether that VLAN is registered on the reception Port.

The MVRP protocol serves end stations that want to exchange data across the network with the specific VID. MVRP guarantees that the required VID will be present on all MVRP-aware devices on the path from one station to another without any manual configuration required. An MVRP request from a device (end station or bridge) means that the device wants to receive traffic on the requested VID. If the data flow is bidirectional, then each end station desiring the flow must issue the same MVRP request.

Receiving the MVRP request for a specific VID on a port of the bridge implies:

- The requested VLAN is dynamically added to the bridge's Dynamic VLAN Registration Entries table.
- The port where the request is received is dynamically added to the set of ports that participate in the requested VLAN.
- For a bridge, the MVRP request is propagated to all other ports that are in the forwarding state in at least one instance of a Multiple Spanning Tree context.

The port of a bridge that receives an MVRP request converts the Join Request into a Join Indication, and an MVRP attribute is registered on these ports. On receipt of a Join Indication, MVRP creates the requested VLAN and adds the ingress port as a member of the newly created VLAN. Also, the Join Indication calls the MAP function to propagate the attribute to all other MVRP-enabled ports in the same MAP context.

Declarations are "alive" while at least one registration exists. Registrations can be purged by LeaveTimer if no MVRPDUs with confirmation are received within the LeaveTimer value after LeaveAll timer expiration, or by receiving an MSRPDU with the Leave event. The LeaveAll timer is running constantly. The purging time is variable and depends on when the LeaveAll timer expires after traffic has been stopped. The possible range is [LeaveTimerValue, LeaveTimerValue + LeaveAllTimerValue * 1.5].

MMRP

MMRP allows hosts and bridges to dynamically register and de-register multicast group membership or individual MAC addresses with bridges attached to the network. MMRP propagates that information across all the bridges that support Extended Filtering Services in the network. The MAC address attributes registered, deregistered, and disseminated via MMRP can apply to a group MAC address or individual MAC addresses. The exchange of multicast group membership information can result in the creation or updating of the MAC Address Registration Entries in the Filtering Database to indicate the ports and VLAN IDs on which the multicast groups have been registered.

Operationally, MMRP defines a sub-tree of the active spanning tree as a result of the creation of MAC Address Registration Entries in the filtering databases of the bridges. End stations may also make use of the group membership information registered via MMRP to keep track of the groups for which active members currently exist and the service requirements of upstream devices.

This allows end stations that are sources of frames destined for a Group to suppress the transmission of such frames if their registered Group membership and Group service requirement information indicates that there are no valid recipients of those frames reachable via the networks to which they are attached.

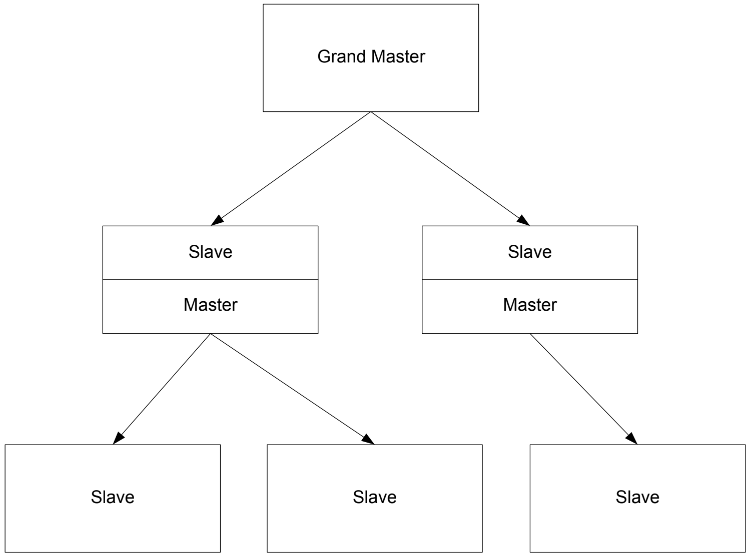
This end system behavior (known as source pruning) allows MAC service users transmitting MAC frames destined for a number of groups to avoid unnecessary flooding of traffic in the local network when there are no group members registered to receive traffic.

IEEE 802.1AS

IEEE 802.1AS is a protocol designed to synchronize clocks in the nodes of a distributed system that communicate in a bridged network. 802.1AS also provides a mechanism to measure link delays, which may be used to calculate end-to-end propagation delay.

The IEEE 802.1AS standard specifies the protocol and procedures for ensuring that QoS requirements are met for time-sensitive applications such as audio and video. The IEEE 1588 Precision Time Protocol (PTP) forms the basis of the IEEE 802.1AS standard. PTP specifies a precise clock synchronization protocol that relies on timestamped packets. The PTP protocol is applicable to distributed systems that consist of one or more nodes communicating over some set of communication media. The distribution of synchronous time information is performed in hierarchical manner with a grandmaster clock at the root of the hierarchy. The grandmaster provides a common and precise time reference for one or more directly-attached slave devices by periodically exchanging timing information. In other words, all slave devices synchronize their clocks with grandmaster clock. The slave devices can, in-turn, act as master devices for further hierarchical layers of slave devices. These relationships are shown in Figure 47-1.

Figure 47-1. IEEE 802.1S Master/Slave Device Relationships



The 802.1AS implementation described in this document is based on the IEEE P802.1AS/D7.6 draft standard [1].

IEEE 802.1AS time synchronization provides a common time base for sampling data streams at a source device and presenting those streams at a destination device with the same relative timing. End-to-end synchronization of clocks is critical for traffic that is highly time-sensitive and has stringent latency and jitter requirements.

The advantages of using 802.1AS include:

- Network clock synchronization in the sub-microsecond range
- Synchronization of clocks with different precision, resolution, and stability
- Simple installation and maintenance
- Fast convergence when topology changes occur
- Low-cost implementation in multicast messaging networks such as Ethernet

A device that can issue or receive IEEE 802.1AS communications is termed a "time-aware system". A time-aware system can either be an end station device attached to a network or a bridge that interconnects end stations. Typically, an end station device has single port and a bridge has multiple ports. The segment of an 802.1AS network that enables direct communication between two time-aware systems is defined as an 802.1AS communication path. The port on time-aware end station can be a master or slave. A time-aware bridge can have at most one port in slave state and all other ports in master state. If the time-aware system is the grandmaster, all ports on that system are in the master state and there are no slave ports.

The main functions of 802.1AS component are:

- Best Master selection (BMCA)
- Time Synchronization
- Link delay measurement

Best Master Selection

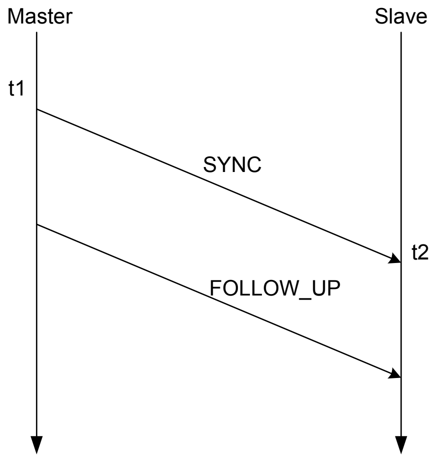
An 802.1AS domain consists of one of more time-aware systems that communicate with each other as defined by the 802.1AS protocol. An 802.1AS network consists of single 802.1AS domain and a single grandmaster clock.

The best master clock algorithm (BMCA) determines the best master for an 802.1AS domain and establishes a tree structured master-slave hierarchy with the best master as the root. Best master clock selection information is periodically exchanged between time-aware systems via ANNOUNCE messages. Every time-aware system invokes the BMCA to compare data associated with the announcing clocks to determine the better clock. If a master port receives an ANNOUNCE message from a better clock, then the port ceases to be a master and becomes a slave port. Likewise, if a clock with a port acting as slave determines that it is a better master than the current master clock, it becomes a master port and starts sending ANNOUNCE messages. In a steady state, once the best master is elected, the ANNOUNCE messages are sent only on master port(s), and the slave port receives ANNOUNCE messages sent by its master port. The mean default time interval between successive ANNOUNCE message transmissions is 1 second. The default ANNOUNCE receipt timeout is 3 times the ANNOUNCE message transmission interval. If ANNOUNCE receipt timeout occurs, the time-aware station invokes the BMCA to determine the new best master clock. If the best master clock is grandmaster-capable, then the clock becomes the grandmaster clock for the 802.1AS domain, generating time synchronization information periodically. The ANNOUNCE message also includes a path trace TLV that tracks the path to best master clock. Each time-aware system updates the received ANNOUNCE message by appending its clock identity to the path trace TLV. The path trace TLV is mandatory as long as the link MTU is not exceeded, in which case the ANNOUNCE message is transmitted without path trace TLV information.

Time Synchronization

A grandmaster clock periodically synchronizes the attached slave clock devices through the use of SYNC and FOLLOW_UP messages. The master timestamps the transmission of each SYNC message and the slave timestamps its reception. The transmit timestamp is conveyed to the slave via the FOLLOW_UP message. This method is illustrated below:

Figure 47-2. SYNC and FOLLOW_UP Messages



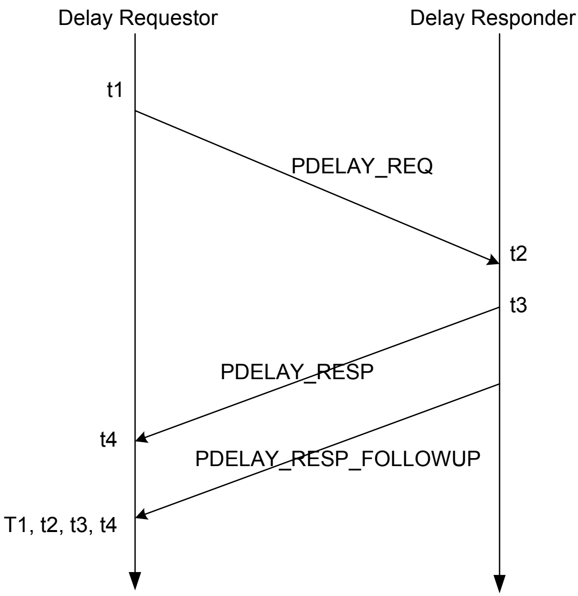
The receipt of FOLLOW_UP message causes the local clock to be synchronized. SYNC and FOLLOW_UP messages must be sent by the current master clock, otherwise the slave port discards the messages. To match a FOLLOW_UP message to SYNC message, there is a sequence ID field in the SYNC and FOLLOW_UP messages. The master sends the FOLLOW_UP message with the same sequence ID as the SYNC message. The value (t2 – t1) gives the (offset + link delay) between the master and slave. The link delay is calculated as described below. Assuming that the link delay is symmetric, the offset value can be derived from (t2 – t1). This sequence of SYNC and FOLLOW_UP messages is repeated at every SYNC transmission interval.

A slave device can, in turn, act as a master for the downstream nodes by forwarding incoming synchronization information to the downstream nodes. The received SYNC/FOLLOW_UP messages are not forwarded as-is; a new set of SYNC/FOLLOW_UP messages is generated based on the downstream master port properties (port identity, port sequence ID, etc.). Also, the correction field in the FOLLOW_UP packet is updated with the residence time before transmitting the message on master port(s). The rate at which SYNC messages are transmitted from the master port is limited. For example, if SYNC messages are received on a bridge's slave port at a very high rate, the SYNC messages are transmitted out of the bridge's master port at a rate greater than or equal to one half of its configured SYNC interval. Ideally a SYNC message is transmitted from a master port whenever a SYNC message is received on its slave port. If a bridge acts as a grandmaster, SYNC messages are generated based on the CPU clock speed; however, the rate at which SYNC messages are transmitted out of the master interface is always greater than or equal to half of the configured SYNC interval.

Link Delay Measurement

The peer delay mechanism measures the port-to-port propagation time (link delay) between two ports. A delay measurement is made on all ports and can be initiated independently of the 802.1AS port state. The transmission of PDELAY_REQ message is the first step in the measurement process. The delay requestor captures the time stamp of the transmission time of PDELAY_REQ packet (t1). When the PDELAY_REQ message is received by the delay responder, the packet's RX time stamp is captured (t2). The delay responder issues two packets in response to the PDELAY_REQ: a delay response PDELAY_RESP and a delay response follow up PDELAY_RESP_FOLLOWUP. The receive timestamp of PDELAY_REQ (t2) and transmit timestamp of PDELAY_RESP (t3) is captured and is conveyed using the follow up message. The delay requestor captures RX timestamp of PDELAY_RESP_FOLLOWUP message (t4). This sequence is shown in the below diagram:

Figure 47-3. Link Delay Measurement Sequence



After the completion of delay request/response exchange, the delay requestor has all four time stamps (t1, t2, t3, t4). The link delay is computed using the following equation:

Tmean_delay = ((t2 - t1) + (t4 - t3))/ 2

The mean time interval between successive PDELAY_REQ messages sent over a link is defined by the 802.1AS standard. If the link delay between two time-aware systems is greater than a “reasonable” cable delay (a few microseconds), then the ports on those links are not enabled for 802.1AS. The accuracy of this algorithm depends on symmetric link propagation delay between the peers, meaning that the link delay from requestor to responder is the same as the delay from responder to the requestor. The peer delay mechanism also requires that there are no transparent devices (bridges) that can add extra delay between the peers.

As part of the PDELAY exchange, the requestor computes the ratio of the frequency of the responder's local clock at the other end of the link and the frequency of the requestor's local clock. To account for the frequency offset between the clocks at each end, the peer delay is adjusted based on the computed ratio.

Caveats and Limitations

IEEE 802.1AS PTP requires the presence of a grand master clock. Dell Networking N-Series AVB switches are not capable of acting as grand master clocks.

IEEE 802.1AS PTP does not operate across a stack. Both the master and all slave ports must be located on the same stack unit.

CoS queues 5–7 should not be used for SRP traffic, as they are also used for control plane traffic. Use CoS queues 2–4 instead.

Dell Networking N-Series MSRP supports MMRP/MSRP only for identification of multicast sources. GMRP and IGMP are not supported.

The MRP protocols are not tied to the spanning-tree protocols. A spanning-tree must exist, but MRP operates independently of the spanning-tree protocols and does not interact with them.

The MRP "New" message can be used to discard learned information without using the service registration (pruning) state machines of MVRP.

The IEEE 1588 Precision Time Protocol supports the logical grouping of PTP clocks into multiple sub-domains, with each sub-domain having a best master clock. PTP clocks are synchronized to the grandmaster clock in that sub-domain. There is no such grouping of clocks in an 802.1AS domain.

The time synchronization mechanism is applicable only after a grandmaster is elected and port states are set to Master/Slave. In other words, SYNC and FOLLOWUP messages are not sent when the Best Master Clock Algorithm selection is in progress.

In compliance with sections 11.2.3 and 11.2.4 of IEEE 802.1AS, canonical flow control (PAUSE) and Priority Flow Control (PFC) must be disabled on bridges that are enabled for PTP. This configuration is not enforced by bridge management.

In-situ measurements have shown residence times of up to 10 ms and PDELAY turnaround times of up to 1 ms.

AVB Configuration Example

The following example configures an AVB switch.

1 Create VLAN 2. This VLAN is used to carry the MSRP traffic.

```
console#config
console(config)#vlan 2
console(config-vlan2)#exit
```

2 Configure two CoS queues for minimum latency, with CoS queue 3 receiving a guarantee of 10% of the scheduler and CoS queue 4 receiving a guarantee of 15% of the scheduler.

```
console(config)#cos-queue min-bandwidth 0 0 0 10 15 0 0
```

3 Configure CoS queues 3 and 4 as strict priority queues. This means that CoS queue 4 packets will be scheduled for transmission first, with up to 15% of the scheduler slots. CoS queue 3 packets will be scheduled next, with up to 10% of the scheduler slots, and then all other packets will be scheduled fairly with the remaining scheduler slots.

```
console(config)#cos-queue strict 3 4
```

4 Configure interfaces Te1/0/1-4 as trunk ports and enable them for 802.1AS, MVRP, MSRP, and MMRP. MSRP classes A and B are allowed to reserve 15% and 10% of the switch bandwidth respectively.

```
console(config)#interface range Te1/0/1-4
console(config-if)#switchport mode trunk
console(config-if)#dot1as
console(config-if)#mvrp
console(config-if)#mmrp
console(config-if)#msrp
console(config-if)#msrp delta-bw class A 15
console(config-if)#msrp delta-bw class B 10
console(config-if)#exit
```

5 Globally enable IEEE 802.1AS and set the local clock type to 2 with a priority of 128.

```
console(config)#dot1as
console(config)#dot1as priority 2 128
```

6 Globally enable MVRP, MMRP, and MSRP and enable the periodic state machines to purge registrations periodically. Also enable MSRP talker pruning.

```
console(config)#mvrp global
console(config)#mvrp periodic state machine
console(config)#mmrp global
console(config)#mmrp periodic state machine
console(config)#msrp talker-pruning
console(config)#msrp global
```

7 Configure the MSRP fan-in to be 4 to 1. This is used in calculating the available bandwidth, i.e. four ingress ports are capable of receiving streams to be transmitted out of a single egress port. Class A traffic is mapped onto priority class 4 and class B traffic is mapped onto priority class 3.

```
console(config)#msrp max-fan-in-ports 4
console(config)#msrp srclassqav class A pcp 4
console(config)#msrp srclassqav class B pcp 3
console(config)#exit
```

8 Use commands such as the following, among others, to verify the configuration and operation of the various protocols.

```
console#show msrp interface te1/0/1

MSRP Interface Admin Mode...................... Enabled
srclass-pvid................................... 2
MSRP class A Boundary port status.............. True
MSRP class B Boundary port status.............. True
MSRP QAV class A delta bandwidth............... 15
MSRP QAV class B delta bandwidth............... 10
MSRP class A bandwidth (allocated/total)....... 0 / 18745200
MSRP class B bandwidth (allocated/total)....... 0 / 31242000
MSRP total bandwidth (allocated/total)......... 0 / 93726000
QAV class A priority........................... 4
QAV class A remap priority..................... 1
QAV class B priority........................... 3
QAV class B remap priority..................... 1

console#show dot1as summary
```

```
802.1AS Global Admin Mode...................... Enabled
Grandmaster Capable............................ Yes
Best Clock Identity............................
F8:B1:56:FF:FE:0F:2B:49
Best Clock Priority1........................... 246
Best Clock Priority2........................... 128
Steps to Best Clock............................ 0
Local Clock Identity...........................
F8:B1:56:FF:FE:0F:2B:49
Local Clock Priority1.......................... 246
Local Clock Priority2.......................... 128
Grandmaster Change Count....................... 0
Last Grandmaster Change Timestamp.............. 0
```

console#**show mrp interface summary**

```
Intf      JoinTimer LeaveTimer LeaveAllTimer
--------- --------- ---------- -------------
Te1/0/1   20        300        2000
Te1/0/2   20        300        2000
Te1/0/3   20        300        2000
Te1/0/4   20        300        2000
Te1/0/5   20        300        2000
Te1/0/6   20        300        2000
```

48

OpenFlow

Dell Networking N2000, N3000, and N4000 Series Switches

Dell Networking OpenFlow Hybrid Overview

The following acronyms are used in this chapter.

Table 48-1. OpenFlow Acronyms

Acronym	Definition
ICAP	Ingress Content Aware Processor. This is a hardware flow matching table. The term ICAP is used synonymously with IFP.
IFP	Ingress Field Processor. The IFP is a hardware flow matching table.
OVS	Open vSwitch
VCAP	VLAN Content Aware Processor. This is a hardware flow matching table. The term VCAP is used synonymously with VFP.
VFP	VLAN Field Processor. The VFP is a hardware flow matching table.

The Dell Networking OpenFlow Hybrid feature implements a true OpenFlow hybrid model as opposed to a 'ships-in-the-night' model. Packets may be forwarded normally via bridging or routing for interfaces that do not have flows installed. On interfaces with flows installed, packet forwarding proceeds normally (except as noted herein) for non-matching flows. Dell Networking OpenFlow Hybrid enables the switch to be managed by a centralized OpenFlow Controller using the OpenFlow protocol. Dell Networking OpenFlow Hybrid is not supported in stacking deployments. Remove all stack members prior to enabling Dell Networking OpenFlow Hybrid.

Dell Networking partially supports the OpenFlow 1.0 and OpenFlow 1.3 standards. The Dell Networking OpenFlow Hybrid switch contains OpenFlow agent version 2.3.0 from the Open vSwitch (OVS) project. The Open vSwitch code is licensed under the Apache 2 license. The OpenFlow agent has been validated with the Helium release of OpenDaylight (ODL).

The OpenFlow 1.0 standard supports a single-table data forwarding path. Dell Networking switches support Open Vswitch proprietary extensions to enable the OpenFlow controller access to multiple forwarding tables.

The OpenFlow 1.3 standard enables a multi-table data forwarding path. Dell Networking switches, however, support a single-table OpenFlow 1.3 data forwarding path.

The Dell Networking OpenFlow feature has the following major functions:

1. Enabling Dell Networking OpenFlow Hybrid.
2. Deploying OpenFlow Configuration.
3. Interacting with the OpenFlow Controllers.
4. Deploying OpenFlow Controller Flows.
5. Collecting Port and Queue Status and Statistics.
6. Supporting OpenFlow Controller Group tables

Enable Dell Networking OpenFlow Hybrid

The OpenFlow feature can be enabled and disabled by the network administrator. Although Dell Networking OpenFlow Hybrid may be administratively enabled, it is not operational until the switch has an IP address.

The OpenFlow feature can be administratively disabled at any time. After administratively disabling the feature, the network administrator must wait until the OpenFlow Feature is operationally disabled before re-enabling the feature.

The administrator can allow the switch to automatically assign an IP address to OpenFlow or to specifically select which address should be used. The administrator can also direct the OpenFlow feature to always use the out-of-band interface.

If the address is assigned automatically and the interface with the assigned address goes down, the switch selects another active interface if one is available. Dell Networking OpenFlow Hybrid becomes operationally disabled and re-enabled when a new IP address is selected. If the address is assigned statically, the OpenFlow feature comes up only when a switch interface with the matching IP address becomes active.

Automatic IP address selection is done in the following order of preference.

1. Loopback interfaces.
2. Routing interfaces.
3. Out-of-band interface.

Dell Networking switches support IPv4 addresses for connecting to the OpenFlow controller. IPv6 addresses are not supported.

If IP routing is enabled, the out-of-band interface cannot be used as the OpenFlow interface.

Once the OpenFlow IP address is selected, it is used until the interface goes down or the OpenFlow feature is disabled or, in case of automatic address selection, a more preferred interface becomes available.

If the out-of-band interface is manually selected as the OpenFlow IP address, the Open Flow feature becomes enabled immediately, even if there is no IP address assigned to the interface.

The selected IP address is used as the endpoint of the IP connections to the OpenFlow controllers.

When the OpenFlow feature is operationally disabled, the switch drops connections with the OpenFlow Controllers. The switch also purges all flows programmed by the controllers.

If the administrator changes the OpenFlow variant while the OpenFlow feature is enabled, the switch automatically disables and re-enables OpenFlow. This causes all flows to be deleted and connections to the controllers to be dropped.

If the administrator changes the default hardware table for OpenFlow 1.0 and the switch is currently operating in OpenFlow 1.0 variant, the OpenFlow feature is automatically disabled and re-enabled.

Interaction with the OpenFlow Controllers

Dell Networking OpenFlow Hybrid implements a subset of the OpenFlow 1.0 protocol and a subset of the OpenFlow 1.3 protocol. Dell Networking OpenFlow Hybrid also implements certain enhancements to the OpenFlow protocol to optimize it for the Data Center environment and to make it compatible with Open vSwitch. Dell Networking OpenFlow Hybrid interacts with any OpenFlow controller that supports OpenFlow 1.0 or the OpenFlow 1.3 standards.

This section covers the following topics:

- Dell Networking OpenFlow Hybrid Supported OpenFlow messages and options.

Dell Networking OpenFlow Hybrid Principles of Operation

The Dell Networking OpenFlow Hybrid OpenFlow implementation is targeted for the data center market as opposed to the education market. As a consequence of this design decision, some aspects of the OpenFlow 1.0/1.3 specifications are not supported, while extra features are added to enhance the data center networking environment.

Key limitations are:

- A single bridge instance.
- A limited subset of supported flow actions.
- IPv4 and IPv6 flows cannot both be supported with iSCSI enabled on the N4000. Disable iSCSI to support IPv4 and IPv6 flows simultaneously on the N4000.
- The IPv6 destination address field is not supported on the N4000. However, the IPv6 flow label is supported.

The Dell Networking OpenFlow Hybrid implements the following behaviors:

1. The switch behaves as an OpenFlow-Enabled Hybrid switch. This means that the switch can forward OpenFlow and normal layer-2 and layer-3 traffic on the same ports and the same VLANs at the same time. When the controller adds flows, the ports mentioned in the match criteria or egress actions are automatically assumed to be OpenFlow ports, so the switch disables ingress and egress filtering on those ports and allows the ports to receive and transmit traffic for any VLAN. This change in the ingress and egress filtering behavior may affect how the switch handles the non-OpenFlow traffic on those ports.

2 The switch supports only one bridge instance.

3 In OpenFlow 1.0 mode, the switch supports several backup OpenFlow controllers. The backup controllers can exchange hello messages with the switch, but cannot add flows or monitor switch status. A vendor message is defined to allow a backup controller become a primary controller. In the OpenFlow 1.3 mode several OpenFlow controllers can manage the switch at the same time.

4 In the OpenFlow 1.0 mode, the switch supports multiple hardware tables to which flows are added. The switch advertises to the controller as having multiple tables. The multi-table support in OpenFlow 1.0 is weak because it does not allow the OpenFlow controller to specify the table to add the flow to. Dell Networking OpenFlow Hybrid extends the OpenFlow 1.0 protocol to specify the table number into which the flow is inserted by using the most significant byte of the command field in the OFPT_FLOW_MOD message. "OpenFlow 1.0 Supported Flow Match Criteria, Actions and Status " on page 1616 defines which flows are added to which hardware tables.

5 In OpenFlow 1.3 mode, the switch supports only one hardware table.

6 When operating in the OpenFlow 1.3 mode, the switch supports the group table. See "Group Table " on page 1638 for more information.

7 The switch does not support the OpenFlow 1.0 emergency flow table.

8 The switch does not support forwarding packets in software. If a flow cannot be added to the hardware, the switch generates an error message.

9 The switch does not support adding flow match criteria and forwarding actions for ports that are not currently present in the system. However, if ports are removed after the flow is installed, then the flow is updated with the correct port forwarding rules. If the match port is not present on the switch, the switch holds the flow in a software table and applies the flow to the hardware when the port becomes available. If the port for a forwarding action is not present on the switch, the switch adds the flow without the missing port and modifies the flow when the port becomes available. This behavior can cause a flow to be added with no egress ports, which causes packets matching the flow to be dropped.

10 When the switch loses connection to the OpenFlow controller it continues to forward traffic using the flows previously programmed by the controller. When the switch reconnects to the controller, it keeps using the previously programmed flows until the OpenFlow controller tells it otherwise.

11 At boot time, when the switch does not have any flows, it forwards traffic normally using the layer-2/layer-3 forwarding rules.

12 The switch supports sending data packets to the controller. However, the controller must explicitly install a flow to forward packets to the controller. Packets that do not match any flow entries are forwarded normally using the layer-2 or layer-3 logic.

13 The switch supports the ability for the controller to inject packets into the network via the switch. This means that the controller can inject packets into the network.

14 The switch supports only a limited set of flow match criteria and actions. See "OpenFlow 1.0 Supported Flow Match Criteria, Actions and Status " on page 1616.

15 The switch supports flows for physical ports and LAGs. These ports can be used as destinations and match criteria. Status and statistics are reported for these ports.

16 The switch supports eight CoS queues per physical port. Redirection to queues is supported only for the OpenFlow 1.3 protocol. Queue status reporting is supported.

17 The switch supports flow aging. The switch checks the flow install time and idle time every 30 seconds. If either of the timers exceeds the configured values for the flow, the switch deletes the flow. For hardware tables that do not support flow statistics, the switch does not support the idle timeout.

OpenFlow 1.0 Supported Flow Match Criteria, Actions and Status

The Dell Networking OpenFlow Hybrid switch supports a limited set of match criteria and actions. This section defines which match criteria and flow actions are supported in each hardware table.

Dell Networking OpenFlow Hybrid adds flows into one of the following hardware tables: the VLAN Field Processor or the Ingress Field Processor. The Ingress Field Processor is subdivided into two different hardware tables: the

"MAC Forwarding Table" and the "OpenFlow 1.0 Rule Table". The hardware table to which the flow is added depends on the flow table identifier specified in the OFPT_FLOW_MOD message.

The flows are added, modified, and removed using the OFPT_FLOW_MOD message. The OFPT_FLOW_MOD message is handled by the Open vSwitch layer and the resulting flow modification commands are passed to Dell Networking OpenFlow Hybrid using the ofproto_class interface.

Dell Networking OpenFlow Hybrid enables the OpenFlow 1.0 Controller to add flows to different tables by making use of the most significant byte in the command field in the OFPT_FLOW_MOD message. If this byte is 0, the flow is added to the default table configured by the administrator. If the byte is not zero, the flow is added to the flow table specified in Table 48-2.

The following table identifiers are mapped to the listed hardware tables. The table identifiers are not contiguous because some identifiers are reserved for future enhancements. The supported hardware table IDs, sizes, and descriptions are accessible through the switch user interface.

Table 48-2. Flow Table Identifiers

ID	Usage	Description
0	User-Configured table.	This table ID in the OFPT_FLOW_MOD messages indicates that the rule should be added to the default table configured by the administrator. The standard OpenFlow 1.0 controllers always send 0 to the switch. Table 0 is not reported in the OFPST_TABLE message.
1–3	Reserved	Unused.
4	Source MAC VLAN Assignment	This table is in the VLAN Field Processor.
5–23	Reserved	Unused.
24	OpenFlow 1.0 Rule Table	IFP table containing OpenFlow 1.0 rules.
25	MAC Forwarding Table	IFP table containing multicast and unicast DA-MAC-based forwarding rules.
26–31	Reserved	Unused

Table 48-2. Flow Table Identifiers (Continued)

ID	Usage	Description
32–255	Unsupported	The enhanced OpenFlow 1.0 protocol only supports table IDs 0 to 31.

When using multiple hardware tables, it is possible to set up the hardware so that, for example, the MAC Forwarding Table and OpenFlow 1.0 Rule Table match the same packet. If the packet matches multiple slices in the IFP, the hardware performs all non-conflicting actions on the packet. For example, the OpenFlow 1.0 Rule Table may set the packet priority and the MAC Forwarding Table may direct the packet to a specific output port.

If the packet actions conflict, the egress action is not predictable. The controller-based applications should take care not to insert flow with conflicting actions.

If the packet matches an IFP rule installed by a different component, such as QoS. any conflicting actions are generally resolved in favor of the other component. The IFP slices allocated to the OpenFlow component have the lowest priority except for the system rules. The OpenFlow actions override actions installed by the system rules.

Although the OpenFlow IFP slices are lower priority than IFP slices used by other Dell Networking OpenFlow Hybrid components, the IFP itself is positioned in the ingress pipeline after the forwarding database and the routing tables. This means that IFP rules inserted by the OpenFlow feature can affect switching and routing decisions.

VFP-based flows also may affect switching decisions and alter switching protocols behavior by changing MAC addresses or/and VLAN IDs.

To avoid interfering with non-OpenFlow traffic, the rules should be qualified with a VLAN ID reserved for the OpenFlow traffic. The Dell Networking OpenFlow Hybrid switch does not enforce any specific VLAN IDs and also accepts wildcard VLAN IDs, so it is up to the OpenFlow Controller to configure the switch correctly.

Refer to "Limitations, Restrictions, and Assumptions " on page 1655 for the list of known interferences.

This section includes the following topics:

OpenFlow 1.0 Rule Table

The OpenFlow 1.0 rule table implements many of the OpenFlow match criteria and actions defined in the OpenFlow 1.0 standard.

The table is implemented in the Ingress Field Processor using slices configured in the intra-slice double-wide mode. This means that the number of rules in each IFP slice is divided in half to provide the necessary rule width.

The following sections describe the match criteria and actions supported by the OpenFlow 1.0 table.

- OpenFlow 1.0 Match Criteria

  Table 48-3 defines the OpenFlow 1.0 match criteria supported by Dell Networking OpenFlow Hybrid. The fields in the table correspond to the fields defined in Table 3 in the OpenFlow 1.0 Switch Specification.

  In summary, the Dell Networking OpenFlow Hybrid switch supports matching on all fields specified in the OpenFlow 1.0 standard except the IP address fields in ARP frames and the ARP op-code.

  If the switch is configured to operate as a router, then for IPv4 packets, the hardware matches the packet fields only if the packet can be forwarded by the hardware. Packets with IP header errors and packets with options in the IP header are sent to the switch protocol stack and cannot be intercepted by the OpenFlow controller.

  If the switch is not a router and is only performing layer-2 switching, then it ignores IPv4 header errors and applies the match rules to all IPv4 packets.

  All fields in this table can be wild-carded.

Table 48-3. Supported OpenFlow Match Criteria

Match Field	Description
Ingress Port	Physical port or LAG.
Ethernet Source Address	The 6-byte source MAC.
Ethernet Destination Address	The 6-byte destination MAC.

Table 48-3. Supported OpenFlow Match Criteria (Continued)

Match Field	Description
Ethernet Type	The Ethertype in Ethernet V2 tagged and untagged packets.
VLAN ID	The VLAN Identifier field in the VLAN header. The valid range for the VLAN ID is 1 to 4094. Note that all packets are tagged in the system when they are processed by the OpenFlow 1.0 classifier. The packets that entered the switch without a tag are assigned a tag either by the ingress port PVID or by the Source MAC VLAN Assignment Table. Thus 0xFFFF, a special VLAN designator indicating that the entry should match untagged traffic, cannot be used as a match criteria for this field.
VLAN Priority	The VLAN Priority field in the VLAN header. The valid range for the VLAN Priority is 0 to 7. Note that all packets are tagged in the system when they are processed by the OpenFlow 1.0 classifier. The packets that entered the switch without a tag are assigned a default port priority configured for the port.
IP Source Address	The 4-byte IP source address in IPv4 packets. Only packets with Ethertype 0x0800 can match to the IP Source Address field. The OpenFlow controller is not required to explicitly set up the Ethernet Type match field. The Ethernet Type field may be wildcarded and the switch can still match IPv4 packets. The switch supports subnet masking for the IP Source Address. The Source IP Address matching within ARP packets is not supported.

Table 48-3. Supported OpenFlow Match Criteria (Continued)

Match Field	Description
IP Destination Address	The 4-byte IP destination address in IPv4 packets. Only packets with Ethertype 0x0800 can match to the IP Destination Address field. The OpenFlow controller is not required to explicitly set up the Ethernet Type match field. The Ethernet Type field may be wildcarded and the switch can still match IPv4 packets. The switch supports subnet masking for the IP Destination Address. The Destination IP Address matching within ARP packets is not supported.
IP Protocol	1-byte IP Protocol field in the IPv4 packets. The hardware matches this field only against IPv4 packets. The protocol field in ARP packets is not supported.
IP ToS	The most significant 6-bits of the Type of Service byte. The value is actually interpreted as the DiffServ Codepoint. Only IPv4 frames can match this classifier.
Transport Source Port / ICMP Type	Source IP port for TCP and UDP IPv4 packets or ICMP Type. To correctly match on the ICMP type, the controller must set the IP Protocol value to ICMP (1). The IP Protocol can be wildcarded for matching on the IP Port number
Transport Destination Port / ICMP Code	The destination IP port for TCP and UDP IPv4 packets or ICMP code. To correctly match on the ICMP code, the Controller must set the IP Protocol to ICMP (1). The IP Protocol can be wildcarded for matching on the IP port number.

- OpenFlow 1.0 Actions

  The switch supports single-port and multi-port forwarding actions as well as some optional packet modifications actions.

  Table 48-4 defines the supported and unsupported forwarding actions.

Table 48-4. Supported/Unsupported OpenFlow Forwarding Actions

Forwarding Action	Description
Forward—Physical Port	The switch can redirect traffic to one or more ports. A valid port can be a physical port or a LAG. When redirecting traffic to multiple ports, a combination of physical ports and LAGs can be specified in the actions.
Forward—CONTROLLER	Send packet to the OpenFlow controller. The "CONTROLLER" reserved port can also be used with switch port numbers. The flow must not include any packet modification actions.
Forward—ALL	Not Supported. This is a "Required" action for OpenFlow, however it is not practical on a VLAN-enabled switch.
Forward—LOCAL	Not Supported. This is a "Required" action, but does not make sense on an OpenFlow-Enabled switch. The packets are sent to the switch CPU only when directed by its bridging our routing protocol stack.
Forward—TABLE	Not Applicable. The action is defined only for packet-out messages.
Forward—IN_PORT	Not Supported. This is a "Required" action, but is not supported because this action is too dangerous on switches in production network.

Table 48-4. Supported/Unsupported OpenFlow Forwarding Actions (Continued)

Forwarding Action	Description
Forward—NORMAL	This is a supported forwarding action. "NORMAL" reserved port can be either the only action in the list, or can be specified along with the "CONTROLLER" port. No packet modifications are allowed when this action is specified. The packet is forwarded according to normal layer-2 or layer-3 tables. There are two use cases identified for this action: • An access list, where traffic matching the rule is allowed while traffic not matching the rule is dropped. • A statistics monitor, enabling the OpenFlow controller to collect byte and packet counters for the matching traffic.
Forward—FLOOD	Not Supported. This is an "Optional" action and is not supported.
Enqueue	Not Supported. This is an "Optional" action and is not supported. The egress queue is selected based on the packet 802.1p priority and the switch configuration.
Drop	Packets matching the flow with this action are dropped. When this action is specified, it must be the only action in the action list.

Table 48-4. Supported/Unsupported OpenFlow Forwarding Actions (Continued)

<table>
<tr><th>Forwarding Action</th><th>Description</th></tr>
<tr><td>Modify Field</td><td>The switch supports modifying certain fields in the packet. The feature can be used to give higher priority to certain packets by modifying the 802.1p and DSCP fields. The feature can also be used to implement policy based routing.

The packet modifications can be made to the single-port and multi-port flows. If multiple egress ports are specified in the flow then all of the packet modification actions must precede the port forwarding actions. All ports in a multi-port flow perform the same packet modifications. Dell Networking OpenFlow Hybrid does not support modifying packets differently for different ports. This action is only supported for table 4 and 24.

The field modification is supported for the following fields:
• Set VLAN ID
• Set VLAN Priority
• Modify Source MAC Address
• Modify Destination MAC Address
• Modify IPv4 ToS bits

If the flow has a modify VLAN action and does not specify a tagged matched criterion, the flow is rejected.

The packet actions may appear in any order, but must precede any forwarding actions. Also, each type of packet modification action must appear only one time in the flow.

The remaining OpenFlow 1.0 packet modification actions are not supported. The unsupported actions are:
• Strip VLAN Header.
• Modify IPv4 Source Address
• Modify IPv4 Destination address.
• Modify Transport Source Port.
• Modify Transport Destination Port.</td></tr>
</table>

Source MAC VLAN Assignment Table

The Source MAC VLAN Assignment table matches on SA MAC, VLAN, and Input Port. Dell Networking OpenFlow Hybrid checks the 'wildcards' field in the ofp_match structure and returns an error if any of the bits other than OFPFW_IN_PORT, OFPFW_DL_VLAN, or OFPFW_DL_SRC are set to 0. If the OpenFlow Controller specifies an unsupported action, the switch rejects the flow with an error.

Table 48-5. Source MAC VLAN Assignment Table Match Criteria

Name	Description	Match Criteria/Actions
Phase-1-Untagged-MAC	Assign a VLAN to the station. The flow is added to the VFP. Only one untagged VLAN may be used per port.	dl_vlan – 0xFFFF — Special VLAN designator indicating that entry should match untagged traffic.
		in_port — Valid physical or LAG port number on the switch.
		dl_src — Source MAC.
		Action type — OFPAT_SET_VLAN_VID
		VLAN — Valid VLAN ID.
Phase-1-MAC	Assign a VLAN to the station. The flow is added to the VFP.	dl_vlan — Valid VLAN ID.
		in_port — Valid physical or LAG port number on the switch.
		dl_src — Source MAC.
		Action type — OFPAT_SET_VLAN_VID
		VLAN — Valid VLAN ID.
Phase-1-Drop	Drop packets that don't match more specific VFP rules.	dl_vlan — Wildcard.
		in_port — Valid physical or LAG port number on the switch.
		dl_src — Wildcard
		No Actions (Packet is Dropped)

MAC Forwarding Table

The MAC Forwarding table matches on DA MAC, SA MAC, VLAN, and Input Port. Dell Networking OpenFlow Hybrid checks the 'wildcards' field in the ofp_match structure and returns an error if any of the bits other than OFPFW_IN_PORT, OFPFW_DL_VLAN, OFPFW_DL_SRC, or OFPFW_DL_DST are set to 0. 0xFFFF, a special VLAN designator indicating that entry should match untagged traffic, cannot be used as a match criteria for VLAN ID field dl_vlan.

Table 48-6. MAC Forwarding Table Match Criteria

Name	Description	Match Criteria/Actions
Local — MAC	Entry used for sending traffic between local ports.	dl_vlan — Valid VLAN ID dl_dst — Non-multicast destination MAC address. in_port — Wildcard dl_src — Wildcard Action Type — OFPAT_OUTPUT port — Valid physical port or LAG. max_len — Ignored
Local — Broadcast	Match on Broadcast packets sent by the local ports.	dl_vlan — Valid VLAN ID dl_dst — ff:ff:ff:ff:ff:ff in_port — Valid physical port or LAG. dl_src — Wildcard Action Type — OFPAT_OUTPUT (Can be repeated) • port — Valid physical port or LAG. • max_len — Ignored

Table 48-6. MAC Forwarding Table Match Criteria (Continued)

Name	Description	Match Criteria/Actions
Local — Multicast	Match on any MAC address with the multicast bit enabled. All other bits in the destination MAC are implicitly masked.	dl_vlan — Valid VLAN ID dl_dst – 01:00:00:00:00:00 — Special MAC address in_port — Valid Physical Port or LAG. dl_src — Wildcard Action Type — OFPAT_OUTPUT (Can be repeated) • port — Valid physical port or LAG. • max_len — Ignored
Local — Default	Match traffic arriving on local port and a specific VLAN.	dl_vlan — Valid VLAN ID in_port — Valid Physical Port or LAG. dl_dst — Wildcard dl_src — Wildcard Action Type — OFPAT_OUTPUT (Can be repeated) • port — Valid physical port or LAG. • max_len — Ignored
Layer-2-Match	This flow matches all layer-2 fields required for a learning bridge.	dl_dst — Non-Multicast destination MAC address. dl_src — Source MAC Address. dl_vlan — VLAN ID in_port — Ingress physical port or LAG. Action Type OFPAT_OUTPUT • port — Valid physical port or LAG. • max_len — Ignored

Table 48-6. MAC Forwarding Table Match Criteria (Continued)

Name	Description	Match Criteria/Actions
Controller — VLAN	Match traffic for a specific VLAN and send the packet to the OpenFlow Controller.	dl_vlan — Valid VLAN ID dl_dst — Wildcard in_port — Wildcard dl_src — Wildcard Action Type — OFPAT_OUTPUT (Can be specified only one time) • port — OFPP_CONTROLLER (0xfffd) • max_len — An integer from 0 to 9216.

Flow Addition and Modification Error Messages

If the switch detects a problem with a newly added flow, or is unable to add or modify a flow due to lack of hardware resources, the switch generates an error message in response to the ofproto_class Flow Put function and generates a syslog message with a text string representing the error type.

Table 48-7 lists the syslog messages that can be generated by the switch in response to the flow modification requests. The syslog notification level for all these message is 3-Warning.

Table 48-7. Syslog Messages in Response to Flow Modification Requests

ASCII Text	Description
Unexpected 'wildcards' value <hex-value>.	The wildcards field contains bits that are set to 0 for match criteria unsupported by the switch.
Unsupported Match Criteria	The match criteria do not correspond to any supported pattern defined in "OpenFlow 1.0 Rule Table " on page 1620, "Source MAC VLAN Assignment Table " on page 1626, and "MAC Forwarding Table " on page 1627.
Unsupported Match Port <hex-value>.	The match criteria port is not wild-carded and not in the range from 0-0xff00 or 0xffc0.
Invalid Match VLAN <hex-value>.	The VLAN is not in the range from 1 to 4094 or special untagged VLAN 0xFFFF.
Unable to add the flow to the hardware, xid - <hex-value>, table = <integer>.	Hardware does not have enough room to add this flow.
Unsupported Flow Actions, xid - <hex-value>	One or more actions for the flow type corresponding to the match criteria are not supported. The supported actions are defined in "OpenFlow 1.0 Supported Flow Match Criteria, Actions and Status " on page 1616.
Invalid Output Port <hex-value>.	The output port number is not in the range 0–0xff00.

Flow Status and Statistics

The OpenFlow Controller uses the OFPT_STATS_REQUEST message with the type OFPST_FLOW to request flow status and statistics. The switch supports all flow match criteria in the OFPT_STATS_REQUEST defined by the OpenFlow 1.0 standard.

The switch supports packet and byte counters for the OpenFlow 1.0 Rule Table and the MAC Forwarding Table.

The OFPT_STATS_REPLY message includes the flow match criteria and actions.

OpenFlow 1.3 Flow Match Criteria and Actions

The Policy ACL Flow Table supports wide, multi-field matching. Most fields can be wildcard matched, and relative priority must be specified in all flow entries. The Policy ACL Flow Table has actions to redirect packets to different destination groups. It can be used to output copies of packets (for example, ARP packets or BPDU frames) to the Controller.

The Policy ACL Flow Table is organized into mutually exclusive logical sub-tables. Flow entries in the IPv6 logical tables match only packets that require matching on IPv6 header fields. The non-IPv6 logical table matches any packet that does not require matching on IPv6 header fields. Following the OpenFlow single-entry match semantics, since the Policy ACL Flow Table is considered a single table, a packet can match at most one rule in the entire table.

Flow entries must conform to match field prerequisite requirements defined in the OpenFlow specification or in this document. In other words, if a prerequisite field is identified for a particular match field, it must be explicitly provided. For example, to match a TCP source port, the IP protocol must be 4 (TCP) and the Ethertype must be 0x0800 (IPv4) or 0x86dd (IPv6).

The default on table miss is to do nothing. The packet will be forwarded using the output or group in the action set, if any. If the action set does not have a group or output action the packet is dropped.

Flow Match Fields

The available match fields for Policy ACL Flow Table flow entry types are as described in the following tables.

Table 48-8. Policy ACL Flow Table Layer 2 Match Fields

Field	Bits	Maskable	Optional	Description or Prerequisite
IN_PORT	32	No	Yes	Physical or logical ingress port.
ETH_SRC	48	Yes	Yes	Ethernet source MAC
ETH_DST	48	Yes	Yes	Ethernet destination MAC
ETH_TYPE	16	No	Yes	Any value except 0x86dd. Explicit prerequisite must be 0x800 if IP fields are to be matched.
VLAN_VID	16	Yes	Yes	VLAN ID. Cannot be masked for a VLAN bridging rule that redirects to a different L2 output group. Only applicable to VLAN flow entry types.
VLAN_PCP	3	No	Yes	802.1p priority field from VLAN tag. Always has a value, will be zero if packet did not have a VLAN tag.

Table 48-9. Policy ACL Flow Table IPv4 Match Fields

Field	Bits	Maskable	Optional	Description or Prerequisite
IN_PORT	32	No	Yes	Physical or logical ingress port.
ETH_SRC	48	Yes	Yes	Ethernet source MAC
ETH_DST	48	Yes	Yes	Ethernet destination MAC
ETH_TYPE	16	No	Yes	Any value except 0x86dd. Explicit prerequisite must be 0x800 if IP fields are to be matched.
VLAN_VID	16	Yes	Yes	VLAN ID. Cannot be masked for a VLAN bridging rule that redirects to a different L2 output group. Only applicable to VLAN flow entry types.

Table 48-9. Policy ACL Flow Table IPv4 Match Fields (Continued)

Field	Bits	Maskable	Optional	Description or Prerequisite
VLAN_PCP	3	No	Yes	802.1p priority field from VLAN tag. Always has a value, will be zero if packet did not have a VLAN tag.
IPV4_SRC	32	Yes	Yes	Matches SIP if Ethertype = 0x0800
IPV4_DST	32	Yes	Yes	Matches DIP if Ethertype = 0x0800
IP_PROTO	8	No	Yes	IP protocol field from IP header if Ethertype = 0x0800
IP_DSCP	6	No	Yes	Bits 0 through 5 of the IP ToS Field as defined in RFC 2474 if Ethertype = 0x0800
IP_ECN	2	No	Yes	Bits 6 through 7 of the IP ToS Field as defined in RFC 3168 if Ethertype = 0x0800
TCP_SRC	16	No	Yes	If Ethertype = 0x0800 and IP_PROTO = 6
UDP_SRC	16	No	Yes	If Ethertype = 0x0800 and IP_PROTO = 17
SCTP_SRC	16	No	Yes	If Ethertype = 0x0800 and IP_PROTO = 132
ICMPV4_TYPE	8	No	Yes	If Ethertype = 0x0800 and IP_PROTO = 1
TCP_DST	16	No	Yes	If Ethertype = 0x0800 and IP_PROTO = 6
UDP_DST	16	No	Yes	If Ethertype = 0x0800 and IP_PROTO = 17
SCTP_DST	16	No	Yes	If Ethertype = 0x0800 and IP_PROTO = 132
ICMPv4_COD E	8	No	Yes	If Ethertype = 0x0800 and IP_PROTO = 1

Table 48-10. Policy ACL Flow Table IPv6 Match Fields

Field	Bits	Maskable	Optional	Description
IN_PORT	32	No	Yes	Physical or logical ingress port.
ETH_SRC	48	Yes	Yes	Ethernet source MAC
ETH_DST	48	Yes	Yes	Ethernet destination MAC
ETH_TYPE	16	No	Yes	Must be 0x86dd
VLAN_VID	16	Yes	Yes	VLAN ID. Cannot be masked for a VLAN bridging rule that redirects to a different L2 output group. Only applicable to VLAN flow entry types.
VLAN_PCP	3	No	Yes	802.1p priority field from VLAN tag. Always has a value, will be zero if packet did not have a VLAN tag.
IPV6_SRC	128	Yes	Yes	Matches IPv6 SIP
IPV6_DST	128	Yes	Yes	Matches IPv6 DIP
IP_PROTO	8	No	Yes	Matches IPv6 Next header
IPV6_FLABEL	20	No	Yes	Matches IPv6 flow label
IP_DSCP	6	No	Yes	Bits 0 through 5 of the IP ToS Field as defined in RFC 2474 if Ethertype = 0x86dd
IP_ECN	2	No	Yes	Bits 6 through 7 of the IP ToS Field as defined in RFC 3168 if Ethertype = 0x86dd
TCP_SRC	16	No	Yes	If Ethertype = 0x86dd and IP_PROTO = 6
UDP_SRC	16	No	Yes	If Ethertype = 0x86dd and IP_PROTO = 17
SCTP_SRC	16	No	Yes	If Ethertype = 0x86dd and IP_PROTO = 132
ICMPV6_TYPE	8	No	Yes	If Ethertype = 0x86dd and IP_PROTO = 58

Table 48-10. Policy ACL Flow Table IPv6 Match Fields (Continued)

Field	Bits	Maskable	Optional	Description
TCP_DST	16	No	Yes	If Ethertype = 0x86dd 00 and IP_PROTO = 6
UDP_DST	16	No	Yes	If Ethertype = 0x86dd and IP_PROTO = 17
SCTP_DST	16	No	Yes	If Ethertype = 0x86dd and IP_PROTO = 132
ICMPv6_CODE	8	No	Yes	If Ethertype = 0x86dd and IP_PROTO = 58

Notes:

The following table lists OpenFlow 1.3 match criteria that are NOT supported.

Table 48-11. Match Criteria Not Supported

Field	Description
IN_PHY_PORT	Switch physical input port.
METADATA	Metadata passed between tables.
MPLS_LABEL	MPLS label.
MPLS_TC	MPLS TC.
MPLS_BOS	MPLS BoS bit.
PBB_ISID	PBB I-SID.
TUNNEL_ID	Logical Port Metadata.
ARP_OP	ARP opcode.
ARP_SPA	ARP source IPv4 address.
ARP_TPA	ARP target IPv4 address.
ARP_SHA	ARP source hardware address.
ARP_THA	ARP target hardware address.
IPV6_ND_TARGET	Target address for ND.
IPV6_ND_SLL	Source link-layer for ND.

Table 48-11. Match Criteria Not Supported (Continued)

Field	Description
IPV6_ND_TLL	Target link-layer for ND.
IPV6_EXTHDR	IPv6 Extension Header pseudo-field

Action Set Actions

The Policy ACL Flow Table action set supports the actions listed in Table 48-12.

Table 48-12. Policy ACL Flow Table Flow Entry Action Set

Name	Argument	Description
Group	Group	Sets output group entry for processing the packet after this table. Group must exist, be consistent with the type of rule and packet;, and can be any of: L3 Unicast, L3 Multicast, or L3 ECMP; must respect VLAN ID naming conventions. Specifies the group to which to send this packet. Egress ports can be specified explicitly via groups. This action can be specified at the same time as the "Output" action only if the output port is CONTROLLER.
Output	ifNum	Specifies a port to which to send the packet. Possible values for this action are any valid switch port numbers and the reserved ports "NORMAL" and "CONTROLLER". "NORMAL" reserved port can be either the only action in the list, or specified along with "CONTROLLER" port. "CONTROLLER" reserved port can also be used with switch port numbers.
Drop	–	The packet is dropped
Set Field	–	Support marking DSCP field in IPv4 and IPv6 packets.
Set Field	–	Support marking 802.1p priority in the VLAN tag.

Counters and Flow Expiration

The Policy ACL Flow Table counters are listed in Table 48-13.

Table 48-13. Policy ACL Flow Table Counters

Name	Bits	Type	Description
Active Entries	32	Table	Reference count of number of active entries in the table.
Duration (sec)	32	Per-entry	Seconds since this flow entry was installed
Received Packets	64	Per-entry	Number of packets that hit this flow entry.
Received Bytes	64	Per-entry	Number of bytes that hit this flow entry.

Policy ACL Flow Table expiry provisions are shown in Table 48-14. Each flow entry can have its own timeout values.

Table 48-14. Policy ACL Flow Table Expiry

Name	Bits	Description
Hard Timeout	32	Number of seconds after which flow entry is removed. Optional, entry does not age out if zero or not specified.
Idle Timeout	32	Number of seconds of inactivity, after which a flow entry is removed. Optional, entry does not age out if zero or not specified.

Group Table

The group abstraction enables OpenFlow to represent a set of ports as a single entity for forwarding packets. Different types of groups are provided, to represent different abstractions such as multicasting or multipathing. Each group is composed of a set group buckets, and each group bucket contains the set of actions to be applied before forwarding to the port. Groups buckets can also forward to other groups, enabling groups to be chained together.

- Group indirection to represent a set of ports.
- Group table with 4 types of groups:
  - All — used for multicast and flooding
  - Select — used for multipath
  - Indirect — simple indirection
  - Fast Failover — use first live port
- Group action to direct a flow to a group.
- Group buckets contains actions related to the individual port

A group table consists of group entries. The ability for a flow entry to point to a group enables OpenFlow to represent additional methods of forwarding (e.g., select and all).

Each group entry is identified by its group identifier and contains:

- group identifier: a 32 bit unsigned integer uniquely identifying the group on the OpenFlow switch.
- group type: to determine group semantics.
- counters: updated when packets are processed by a group.
- action buckets: an ordered list of action buckets, where each action bucket contains a set of actions to execute and associated parameters. The actions in a bucket are always applied as an action set.

Dell Networking OpenFlow Hybrid does not assign any special meaning to the group ID. The OpenFlow controller is free to use any valid group identifier. Dell Networking OpenFlow Hybrid determines the type of hardware group to create based on the group type passed from the OpenFlow controller.

- The "Indirect" group type simply creates a next-hop. (L3 Unicast group entry)

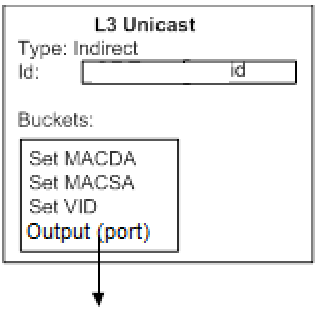
- The "All" group type creates an IPMC replication group that points to one or more next hops. Depending on the SA/DA/VLAN modifications actions, the next hops may be added to the IPMC group as routed or switches. (L3 Multicast group entry)
- The "Select" group type creates an ECMP group object which points to one or more next hops. (L3 ECMP group entry)
- The fast failover group type is unsupported.

The following sections provide additional details on each of these group types.

Indirect (L3 Unicast) Group Type

Indirect Group type (L3 Unicast Group) is used to supply the routing next hop and output interface for packet forwarding. To properly route a packet from the Policy ACL Flow Table, the forwarding flow entry must reference an L3 Unicast Group entry.

Figure 48-1. Indirect Group (L3 Unicast Group) Entry Usage



All packets must have a VLAN tag.

- Action Buckets

  The single action bucket is as shown in Table 48-15.

Table 48-15. Unicast Bucket Actions

Field	Argument	Description
Output	Port	Physical output port. Required

Table 48-15. Unicast Bucket Actions (Continued)

Field	Argument	Description
Set Field	MAC_DST	Write the next hop destination MAC. Optional.
Set Field	MAC_SRC	Write the source MAC corresponding to the L3 output interface. Optional.
Set Field	VLAN-id	Write the VLAN ID corresponding to the L3 output interface. Optional.

- Counters

  The L3 Unicast group entry counters are as shown in Table 48-16.

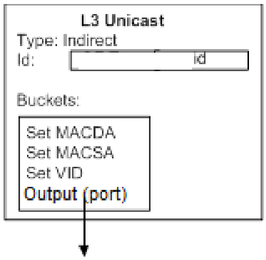
Table 48-16. L3 Unicast Group Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per-entry	Number of group entities currently referencing this group entry.
Duration (sec)	32	Per-entry	Seconds since this group entry was installed

All (L3 Multicast) Group Type

L3 Multicast group entries are of OpenFlow ALL type. The action buckets describe the interfaces to which multicast packet replicas are forwarded. Figure 48-2 illustrates L3 Multicast group entries.

Figure 48-2. L3 Multicast Group Entry Usage



IP multicast packets are forwarded differently depending on whether they are switched or routed. Packets must be switched in the VLAN in which they came, and cannot be output to IN_PORT.

- Action Buckets

  The action buckets contain the values shown in Table 48-17.

Table 48-17. L3 Multicast Bucket Actions

Field	Argument	Description
Set Field	Output Port	Write the L3 output interface. Required.
Set Field	MAC_DST	Write the next hop destination MAC. Optional.
Set Field	MAC_SRC	Write the source MAC corresponding to the L3 output interface. Optional.
Set Field	VLAN-id	Write the VLAN id corresponding to the L3 output interface. Optional.

NOTE: For replication of non-IP packets, all of (MAC-Src, MAC-dest, VLAN-ID) action bucket fields are to be left empty.

For replication of IP packets, at least one of (MAC-Src, MAC-dest and VLAN-ID) should be valid.

L2 multicast is supported. It is done using IPMC L2 replication when all of (MAC-Src, MAC-dest, VLAN-ID) action bucket fields are left empty. So an "All (L3 Multicast) Group" can have a mix of buckets — few with L3 replication and few with L2 replication. To use the L2 multicast, the user should not qualify the IP fields in flow match criteria.

- Counters

  The L3 Multicast group entry counters are as shown in Table 48-18.

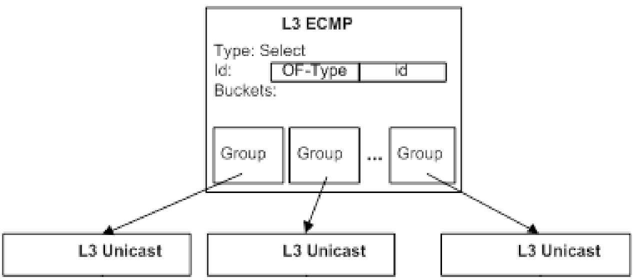
Table 48-18. L3 Multicast Group Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per-entry	Number of group entities currently referencing this group entry.
Duration (sec)	32	Per-entry	Seconds since this group entry was installed

Select (L3 ECMP) Group Type

L3 ECMP group entries are of OpenFlow type SELECT. For IP routing the action buckets reference the L3 Unicast group entries that are members of the multipath group for ECMP forwarding. Figure 48-3 illustrates this L3 ECMP Group entry usage.

Figure 48-3. L3 ECMP Group Entry Usage



An L3 ECMP Group entry can also be used in a Provider Edge Router.

An L3 ECMP Group entry can be specified as a routing target instead of an L3 Unicast Group entry. Selection of an action bucket for forwarding a particular packet is hardware specific.

- Action Buckets

  The action buckets contain the single value listed in Table 48-19.

Table 48-19. L3 ECMP Group Entry Bucket Actions

Field	Argument	Description
Group	Group-id	May chain to an L3 Unicast Group.

- Counters

  The L3 ECMP group entry counters are as shown in Table 48-20.

Table 48-20. L3 ECMP Group Entry Counters

Name	Bits	Type	Description
Reference Count	32	Per-entry	Number of group entities currently referencing this group entry.
Duration (sec)	32	Per-entry	Seconds since this group entry was installed.

Fast Failover

Fast Failover is not supported

Port Configuration, Status and Statistics

The Dell Networking OpenFlow Hybrid switch does not support any port configuration commands. The OFPT_PORT_MOD messages from the OpenFlow controller do not modify the port configuration.

The Dell Networking OpenFlow Hybrid switch reports port creation, removal, and status changes to the controller using the OFPT_PORT_STATUS message. The message is sent for physical ports and LAGs. All three reason codes, OFPPR_ADD, OFPPR_DELETE, and OFPPR_MODIFY, are supported for all port types.

The only trigger for the OFPPR_MODIFY reason code is a link status change for the port.

The desc field in the message contains port information. This field of type ofp_ port contains the following elements:

1. port_no — Set to the MIB-2 ifIndex field for the port.
2. hw_addr — All ports in the switch have the same MAC address. The switch reports the lowest MAC assigned to the unit. This address is typically printed on the MAC address label on the switch.
3. name — A unit/slot/port designation for physical ports and LAGs. The LAGs are also identified with the symbolic name lag-<n>.
4. config — Always set to 0.
5. state — The OFPPS_LINK_DOWN is set according to the link state for physical ports and LAGs. The remaining bits are always 0.
6. curr, advertised, supported, peer — These parameters are set to 0.
7. curr_speed – Current port bitrate in kbps.max_speed—Max port bitrate in kbps

The Dell Networking OpenFlow Hybrid switch reports port statistics in response to the OFPT_STATS_REQUEST message with type OFPST_PORT. The port statistics are reported using the structure of type ofp_port_stats. The packet and byte counters are reported for all interface types. The only supported error counters are rx_errors and tx_errors. The remaining error counters are always reported as 0xffffffffffffffff.

Queue Configuration and Status

The Dell Networking OpenFlow Hybrid switch supports eight queues for each physical port. The LAGs report aggregate statistics for member ports. The port queues are automatically created by the switch for each physical port.

The port queues are prioritized, with queue 7 having the highest priority. The traffic statistics reported for the port queues include OpenFlow traffic and non-OpenFlow traffic.

The OpenFlow Controller retrieves the list of port queues using the OFPT_QUEUE_GET_CONFIG_REQUEST message. Only physical ports and LAGs can be specified in the message.

The queue configuration reply message of type ofp_queue_get_config_reply includes an array of ofp_packet_queue structures. For each interface, the queues are numbered 0 to 7, with queue 7 representing the highest priority queue.

The port queues do not have any queue properties.

The OpenFlow Controller requests queue statistics using the OFPT_STATS_REQUEST message with type OFPST_QUEUE. Dell Networking OpenFlow Hybrid reports the tx_bytes, tx_packets, and tx_errors statistics for each queue.

Deploy OpenFlow Controller Flows

This section describes OpenFlow flow management within Dell Networking OpenFlow Hybrid enabled switches.

Dell Networking OpenFlow Hybrid Flow Database Organization and Manipulation

Dell Networking OpenFlow Hybrid supports multiple hardware flow tables, allowing the OpenFlow controller to specify into which table the flow should be added.

In some cases, the flows added by the OpenFlow Controller cannot be installed in the hardware even though the hardware has space available in its flow table. For example, this can happen if an interface specified in the flow match criteria has not been created yet in the driver. Similarly, the flows can be removed from the hardware when interfaces specified in the match criteria are removed. For example, in a modular system, a port card may be unplugged, causing the interfaces to go away.

When the OpenFlow protocol adds a flow to the Dell Networking OpenFlow Hybrid flow database, the flow is not immediately added to the hardware. The flow additions to the hardware are done by a separate task. Therefore, the OpenFlow Controller can add multiple flows very quickly without blocking, while waiting for flows to be added to the hardware.

Similarly, when the OpenFlow protocol removes flows from the Dell Networking OpenFlow Hybrid flow database, the flows are marked for deletion, but are removed from the hardware by a separate task. This enables the OpenFlow controller to not block while waiting for flows to be removed from the hardware.

To accommodate the scenario where the Flow Controller removes many flows and quickly adds many new flows, the OpenFlow flow database is twice the size of the hardware database. The extra headroom provides enough space to buffer the new flows before the old flows are removed from the hardware.

If the OpenFlow Controller adds a flow with the same match criteria as an existing flow, Dell Networking OpenFlow Hybrid treats the new flow as a flow modification action. The old flow is deleted from the hardware and the new flow is added to the hardware.

If a flow cannot be added to the hardware because the hardware reports that it is out of space, Dell Networking OpenFlow Hybrid sends a message to the OpenFlow controller indicating that the flow addition failed and removes the flow from the software table.

Each time the switch fails to add a flow to the hardware, it sends a syslog message indicating the flow XID.

The switch cannot always accommodate a flow in the hardware because the hardware space is shared between different flow types and is shared with other Dell Networking OpenFlow Hybrid components, and because the hardware usage depends on the flow match criteria.

An example of resource sharing among different flow types is the OpenFlow 1.0 Rule Table (24) and MAC Forwarding Table (25) that share IFP resources. Different flow types may require a different number of IFP slices. VFP-based flows (Source MAC VLAN Assignment (4)) have no common resources with IFP-based flows.

An example of resource sharing between components is the IFP. Both the QoS component and the OpenFlow component use IFP resources. The system does not reserve space in the IFP, but instead allocates resources as they are requested by the application.

The flow match criteria can affect hardware usage in a couple of ways. For entries that are added to the IFP or the VFP, the hardware table usage on multi-ASIC switches depends on the port match criteria in the flow. If the flow matches a physical port, the flow is inserted only into the IFP/VFP on the ASIC where the physical port is located. If the flow does not match a specific physical port, the flow is inserted in all ASICs. Since many flows use the ingress port as a match criterion, the overall flow table capacity depends on how the flows are distributed across the multiple ASICs.

Interaction between Flows and VLANs

The OpenFlow Controller can add flows for any VLAN ID. The VLANs for which flows are added are created in the Dell Networking OpenFlow Hybrid VLAN database as dynamic VLANs if they are not already configured on the switch. Learning is enabled on the dynamic VLAN. The switch never adds ports to OpenFlow dynamic VLANs, but instead disables ingress and egress filtering on the ports on which the OpenFlow flows are installed. This allows the OpenFlow traffic to be received and transmitted on those ports. The OpenFlow flows can also be added for VLANs that are statically created in the VLAN database. However, if the administrator removes a static VLAN with installed flows, then the traffic for those flows may not be forwarded correctly. The administrator should remove all flows on a static VLAN before deleting that VLAN.

VLANs dynamically created with the flows are not deleted when the flows are deleted. Dynamic VLANs are deleted only when the OpenFlow feature is disabled.

If the network administrator does not wish to mix OpenFlow and non-OpenFlow traffic on the same VLANs, then it is up to the administrator to ensure that the OpenFlow Controller is configured such that it does not add flows on VLANs used for non-OpenFlow traffic.

Since OpenFlow VLANs are created in hardware without any port members, the ports on which the OpenFlow traffic enters and exits the switch must disable egress filtering. Dell Networking OpenFlow Hybrid determines which ports are used for OpenFlow by examining the ingress port for flows with non-wildcard port match criteria and port numbers specified in the OFPAT_OUTPUT action. Once ingress/egress filtering is disabled, it is re-enabled only when the OpenFlow feature is disabled or the port is removed from the switch. Even if a flow previously using the port is removed and there are no other flows using the port, ingress/egress filtering remains disabled on that port.

Normally, traffic forwarded to ports with egress filtering disabled is always tagged. However the administrator may want to attach untagged clients to some of the ports. If the egress VLAN is explicitly created by the network administrator and the port is participating in the VLAN as untagged, then the switch settings take precedence over flow rules and traffic is transmitted untagged.

For the switch to receive the untagged traffic and map it to the appropriate VLAN, the OpenFlow controller can install a flow that maps the incoming MAC address to the VLAN. This is done with the flow type "Phase-1-Untagged-MAC" and action OFPAT_SET_VLAN_ID (see "Source MAC VLAN Assignment Table " on page 1626).

For the switch to transmit untagged traffic on the port for the untagged VLAN, the switch uses the VLAN translation table to configure the traffic that matches the VLAN and egress port to be sent untagged.

The switch strips the tag on the VLAN specified in the OFPAT_SET_VLAN_ID action for Phase-1-Untagged-MAC flows that use the magic VLAN ID 0xFFFF as a match criterion.

The pure OpenFlow 1.0 Controllers and the OpenFlow 1.3 controllers do not support Phase-1 flows. If such controllers are used in the network, then to map untagged ingress traffic to a specific VLAN on a specific port, the network administrator must configure the PVID for that port. To send traffic without tags, the network administrator must statically create the VLANs with untagged port members.

Interaction between Flows and Interfaces

Dell Networking OpenFlow Hybrid supports flows on physical ports and LAGs. For a flow to be installed in the hardware, the hardware must know about the interface. Ports that are members of link-up LAGs cannot be match ports or egress ports for flows. When a port becomes a LAG member, it becomes unknown to the OpenFlow application.

If a physical port is enabled for port-based routing, the port becomes unknown to the OpenFlow controller.

The OpenFlow Controller can install flows only on ports that are physically present. The OpenFlow Controller cannot install flows on preconfigured ports that are not physically present. It is possible, however, that the interface goes away after the flow is installed in the hardware. A race condition is also possible where a new flow is added while the port is physically present, but the port disappears before the switch has a chance to add the flow to the hardware.

If an unknown interface is used in the match criteria for a new flow, the flow is held in the application table until the interface is attached. Dell Networking OpenFlow Hybrid does not generate any error for the flow. Once the interface is attached, the flow is added to the hardware.

If the flow is already installed and the interface in the match criteria goes away, the flow is removed from the hardware. Dell Networking OpenFlow Hybrid keeps the flow in the application table and reinserts it into the hardware when the interface becomes attached again.

If an interface specified in the action list for the new flow is not attached, the flow is added to the hardware. If the missing interface is the only egress interface for the flow, the flow is configured to drop matching packets. If the missing interface is one of several egress interfaces, it is simply excluded from the egress interface list.

The OpenFlow application monitors for interface creation and removal events and modifies the flows as needed.

Flow Status and Statistics Collection

The OpenFlow Controller can ask the switch to send it the list of flows that match certain criteria. The switch sends the matching flows with one or more messages.

Flows support packet and byte counters. The switch polls the hardware counters periodically and stores the counter values in the application table. When sending messages to the controller, the switch retrieves the counter values from the application table. The switch does not read the hardware counters when sending flow statistics to the controller.

To avoid performance problems, the counter collection is rate-limited. The switch polls counters for 100 flows every 10 seconds. This means that if the flow table has one to 100 entries, the counters are updated every 10 seconds for all flows. If the flow table has 3000 entries, the counters are updated every 300 seconds for all flows.

The statistics poll rate and the number of flows per poll cycle are porting parameters that can be tuned as needed. When the switch is busy manipulating flows, the statistics update may take longer.

If a flow is removed from the hardware, the packet counters are reported as 0. If the flow is added back to the hardware, the counters start counting from 0.

Collect Port and Queue Status and Statistics

The OpenFlow Controller can collect status and statistics for ports and queues. When ports are created, Dell Networking OpenFlow Hybrid sends an OFPT_PORT_STATUS message to the OpenFlow Controller. The status message is triggered by creation of entries in the Physical Port Table. The same tables are used for reporting port status information.

The port status is updated by a separate task that periodically polls the status for all physical ports. To avoid performance issues, the statistics are polled every 10 seconds for a maximum of 100 interfaces. For physical ports, the switch also reads the queue statistics for all eight queues at the same time as it reads the port statistics.

The LAG statistics are reported as a sum of statistics for all active LAG member ports. This implies that when a port is removed from the LAG, the statistics counters for the LAG go down. Also, when all ports are removed from a LAG, the LAG statistics are reported as 0.

Usage Scenarios

The OpenFlow feature is mainly targeted for deployment in a data center network where devices located in different parts of the network require layer-2 connectivity.

The OpenFlow feature enables customers to avoid scaling problems and loops associated with the Layer-2 network.

The OpenFlow feature can also be used in a research environment, but there are two limitations that may make the "research" use case less attractive. First, there is only one OpenFlow instance, meaning that concurrent experiments are not supported at the switch level. Second, the OpenFlow controller has complete access to all ports and VLANs; therefore, using the switch for mixed production and experimental traffic is not advisable.

Eligible Interfaces

The OpenFlow application affects traffic forwarding on physical ports and LAGs.

OpenFlow Hybrid

The operation of the OpenFlow switch in a network largely depends on the functionality of the OpenFlow controller. The OpenFlow feature is a powerful tool that enables the OpenFlow controller to forward packets in the network without regard to the Layer-2 forwarding database and the IPv4 routing tables.

Refer to the OpenFlow Controller documentation to understand how the switch behaves in the customer network.

The one legacy networking rule the switch enforces is that the switch does not forward packets over ports that are in spanning-tree blocking state for the egress VLAN, even if the OpenFlow Controller has configured a rule to do so.

Example Configuration

This example configures the switch to operate with OpenFlow version 1.3 and to connect to the controller at IP address 1.2.3.4 on port 3435 with no security.

```
console(config)#vlan 10
console(config-vlan10)#interface vlan 10
console(config-if-vlan10)#ip address 1.2.3.1 255.255.0.0
console(config-if-vlan10)#interface gi1/0/1
console(config-if-Gi1/0/1)#switchport mode access
console(config-if-Gi1/0/1)#switchport access vlan 10
console(config-if-Gi1/0/1)#exit
console(config)#openflow

WARNING! OpenFlow does not operate on stack members. Enable
OpenFlow on stand-alone switches only.

console(config-of-switch)#controller ipv4 1.2.3.4 port 3435
security none
console(config-of-switch)#protocol-version 1.3
console(config-of-switch)#mode auto
```

Interaction with Other Switch Functions

The Dell Networking OpenFlow Hybrid component interacts with multiple Dell Networking switch components by either communicating with these components or sharing common resources with the components. The following sections describe these interactions.

OpenSSL

The OpenFlow component establishes SSL connections to the OpenFlow controllers and OpenFlow Managers. The total number of switch-initiated connection depends on the number of configured controllers and managers and can be in the order of 10 to 20 connections.

The OpenFlow component always initiates the SSL connections and does not accept SSL connections.

The OpenFlow component makes use of certificate-based authentication, mutual authentication, and encryption.

IP Stack

The OpenFlow component uses the IP stack for initiating SSL connections and TCP connections. The administrator can configure whether to connect to the OpenFlow Controllers using TCP or SSL. The administrator can also configure the IP port number to use for the connections. By default, the IP ports are 6632 and 6633.

For debugging, the switch accepts TCP connections to ports 6632 and 6633 when passive mode is enabled. Passive connection mode can be enabled using the command **openflow passive-mode** in Global Config mode.

VLANs

The OpenFlow component dynamically creates VLANs that it detects in the flow match criteria or the flow VLAN modification action.

LAGs

When physical ports become LAG members, the flows installed by the OpenFlow Controller on these ports are removed from the hardware and the flows that are installed for the LAG are activated for the new LAG member port. The reverse action takes place when the ports are removed from the LAG.

Ports

The OpenFlow component installs flows in the hardware and removes flows from the hardware as ports become attached and detached or join and leave the LAG.

When flows referencing a specific port in the match criteria or output actions are added to the hardware, the OpenFlow component operationally disables ingress filtering on the port.

Ingress filtering is be re-enabled on those ports when flows are removed or aged-out.

Network Interface ARP Table

In some cases, the OpenFlow component may trigger ARP resolution for a specific IP address.

Routing Interface ARP Table

In some cases, the OpenFlow component may trigger ARP resolution for a specific IP address.

QoS

The QoS component does not interact with OpenFlow directly, but it shares the ingress field processor and VLAN field processor hardware resources with the OpenFlow component.

The QoS component gracefully handles an out-of-resource condition and ensures that flows installed by the QoS component have precedence over the OpenFlow flows when the actions are in conflict.

IP Routing, IP Multicast, and Layer-2 Multicast

The OpenFlow component uses the same hardware resources as the routing and IP multicast components. Namely, the OpenFlow component uses the Next-Hop entries and Multicast Group entries in the hardware.

The routing and multicast Dell Networking OpenFlow Hybrid feature gracefully handles the out-of-resources errors.

Port Mirroring

The OpenFlow component is not active on probe ports. OpenFlow configuration is retained but not operational when a port is set as a probe port.

LLDP and Voice VLAN

The LLDP and Voice VLAN features do not interact with OpenFlow directly. The ODL controller proactively installs the flow to redirect all the incoming LLDP traffic to the controller. Then, it sends LLDP packet to the switch and installs the flows to flood all the incoming traffic to all the physical ports on the switch and to the ODL controller. As all the switches in the controller's network have the rule to redirect the LLDP traffic to the controller, the controller is able to build the network topology. To allow the switch to process incoming LLDP traffic and the ODL controller to build the topology, an overriding flow with "NORMAL" and "CONTROLLER" actions should be installed manually.

Limitations, Restrictions, and Assumptions

The following OpenFlow features are not supported:

1. Flow installation in the MAC Forwarding table.
2. Uplink Rate Limiting, including the flow installation in the Uplink Rate Limiter Table, traffic rate control, the rate limiter table, and the rate limiter statistics.
3. On the N4000 Series switches, flow installation is not supported if MAC ACLs exist.
4. OpenFlow functionality currently interoperates with the Open vSwitch command line utility ovs-ofctl2.3.0. Higher versions may have interoperability issues.

List of OpenFlow—Dell Networking Component Interferences

Table 48-21. OpenFlow-Dell Networking Component Interferences

Component	Behavior
DAI	If DAI is configured along with a MAC-address modification flow, DAI operates on the modified MAC address instead of original MAC address.
Static MAC filter	OpenFlow flow forwarding takes precedence over static MAC filtering. Traffic is forwarded according to flow rule.
IPv6 Destination Address Match	This is not supported on the N4000 due to hardware limitations. Disable iSCSI to install flows matching the IPv6 destination address.

OpenFlow Configuration Example

This example enables OpenFlow 1.3 on the switch and configures a connection to a controller at IPv4 address 172.16.0.3 over TCP port 3435 using no encryption on the out-of-band interface. This example presumes the out-of-band interface has obtained an IP address on the 172.16.0.X subnet.

```
console(config)#openflow

WARNING! OpenFlow does not operate on stack members. Enable
OpenFlow on stand-alone switches only.

console(config-of-switch)#protocol-version 1.3
console(config-of-switch)#controller ipv4 172.16.0.3 port 3435
security none
```

49

Dell Networking Python Support

Dell Networking switches support installation and execution of Python applications. Python applications that are to be executed on the switch must be developed and tested offline to the maximum degree possible. The switch does not offer interactive shell access for development of Python scripts, nor does the Dell Networking switch come with all of the normal Python "batteries included" modules. A list of the included packages is in the example below. Output from Python scripts is sent to the serial console, so a serial connection is mandatory when developing scripts.

An example Python script that prints some useful information is shown below. Explanation of Python syntax is beyond the scope of this document. Refer to the 2.7.10 version of Python documentation available elsewhere.

```
#!/usr/bin/env python
import sys
print "Hello World!\n"
print (sys.version)
help('modules')
```

To execute this script on the switch, save the lines above in a file named "app". Package the app script using the following commands on a Linux system. The package MUST be a gzipped tarball with a .tgz or .tar.gz extension. It is required that the permissions be set on the app file prior to packaging; i.e., user read, write, and execute must be set. Group and other permissions need not be set. Application names can be a maximum of 15 characters.

```
/home/jmclendo/tftpboot>chmod u+rwx app
/home/jmclendo/tftpboot>tar czf app.tgz app
```

Copy the resulting file to the switch using the copy command with the application target keyword. The application file name after the **application** keyword is required and must have a .tgz or .tar.gz extension or the tarball will not be extracted into the user-apps directory.

```
console#copy tftp://10.27.9.99/jmclendo/app.tgz application app.tgz

Transfer Mode.................................. TFTP
```

```
Server IP Address............................. 10.27.9.99
Source File Path.............................. jmclendo/
Source Filename............................... app.tgz
Data Type..................................... Application
Downloads application file

Management access will be blocked for the duration of the transfer
Are you sure you want to start? (y/n) y

File transfer in progress. Management access will be blocked for
the duration of the transfer. please wait...

215 bytes transferred

Application file download completed successfully.
```

Use the **dir** command to see the application. Applications reside in the user-apps subdirectory.

```
console#dir user-apps

Attr  Size(bytes)          Creation Time        Name
 drwx                 224 Jan 01 1970 00:24:56 .
 drwx                2824 Jan 01 1970 00:00:55 ..
 -rwx                  94 Nov 24 2015 15:56:30 app

Total Size: 216555520
Bytes Used: 3142
Bytes Free: 216552378
```

Install the application using the **application install** command in Global Configuration mode.

```
console#conf
console(config)#application install app
console(config)#show application

OpEN application table contains 2 entries.

Name             StartOnBoot AutoRestart CPU Sharing Max Memory
---------------- ----------- ----------- ----------- ----------
SupportAssist    Yes         Yes         0           0
app              No          No          0           0
```

△ **CAUTION:** The application install command has an **auto-restart** parameter. Do NOT use this parameter while debugging or on any short-lived application. The switch does NOT limit restarts and attempts to restart a failed application immediately. Installing a failing or short-lived application with auto-restart enabled will result in a switch that:
– cannot perform normal protocol operations at its advertised level.
– is difficult to access via the console.

The **auto-restart** parameter is recommended only for well-tested, stable applications.

Execute the application in Privileged Exec mode using the **application start** command and examine the results on the serial console.

```
console(config)#exit
console#application start app

Application started.

console#Hello World!

2.7.10 (default, Nov  6 2015, 14:45:45)
[GCC 4.8.2]

Please wait a moment while I gather a list of all available
modules...

BaseHTTPServer     audioop        io              rfc822
Bastion            base64         itertools       rlcompleter
CGIHTTPServer      bdb            json            robotparser
ConfigParser       binascii       keyword         runpy
Cookie             binhex         libopenclt      sched
DocXMLRPCServer    bisect         libospf         select
HTMLParser         bsddb          libpam          sets
MimeWriter         cPickle        libping         sgmllib
OpEN               cProfile       libproc_libs    sha
OpENUtil           cStringIO      librpcclt       shelve
OpEN_py            calendar       libsock_agent   shlex
Queue              cgi            libsshcompat    shutil
SimpleHTTPServer   cgitb          libsshpam       signal
SimpleXMLRPCServer chunk          libtraceroute   site
SocketServer       cmath          libvr_agent     smtpd
StringIO           cmd            libvrf_init     smtplib
UserDict           code           libz            sndhdr
UserList           codecs         linecache       socket
UserString         codeop         locale          spwd
```

```
_LWPCookieJar        collections         logging             sre
_MozillaCookieJar    colorsys            macpath             sre_compile
_OpEN                commands            macurl2path         sre_constants
__builtin__          compileall          mailbox             sre_parse
__future__           compiler            mailcap             ssl
_abcoll              contextlib          markupbase          stat
_ast                 cookielib           marshal             statvfs
_bisect              copy                math                string
_codecs              copy_reg            md5                 stringold
_codecs_cn           crypt               mhlib               stringprep
_codecs_hk           csv                 mimetools           strop
_codecs_iso2022      curses              mimetypes           struct
_codecs_jp           datetime            mimify              subprocess
_codecs_kr           dbhash              mmap                sunau
_codecs_tw           decimal             modulefinder        sunaudio
_collections         difflib             multifile           symbol
_csv                 dircache            multiprocessing     symtable
_ctypes              dis                 mutex               sys
_ctypes_test         distutils           netrc               sysconfig
_elementtree         doctest             new                 syslog
_functools           dumbdbm             nntplib             tabnanny
_heapq               dummy_thread        ntpath              tarfile
_hotshot             dummy_threading     nturl2path          telnetlib
_io                  email               numbers             tempfile
_json                encodings           opcode              termios
_locale              errno               operator            textwrap
_lsprof              exceptions          optparse            this
_md5                 fcntl               os                  thread
_multibytecodec      filecmp             os2emxpath          threading
_multiprocessing     fileinput           parser              time
_osx_support         fnmatch             pdb                 timeit
_pyio                formatter           pickle              toaiff
_random              fpformat            pickletools         token
_sha                 fractions           pipes               tokenize
_sha256              ftplib              pkgutil             trace
_sha512              functools           platform            traceback
_socket              future_builtins     plistlib            tty
_sre                 gc                  popen2              types
_ssl                 genericpath         poplib              unicodedata
_strptime            getopt              posix               urllib
_struct              getpass             posixfile           urllib2
_symtable            gettext             posixpath           urlparse
_sysconfigdata       glob                pprint              user
_testcapi            grp                 profile             uu
_threading_local     gzip                pstats              uuid
_warnings            hashlib             pty                 warnings
```

```
_weakref            heapq             pwd              wave
_weakrefset         hmac              py_compile       weakref
abc                 hotshot           pyclbr           webbrowser
aifc                htmlentitydefs    pydoc            whichdb
antigravity         htmllib           pydoc_data       wsgiref
anydbm              httplib           pyexpat          xdrlib
argparse            ihooks            quopri           xml
array               imaplib           random           xmllib
ast                 imghdr            re               xmlrpclib
asynchat            imp               repr             xxsubtype
asyncore            importlib         requests         zipfile
atexit              imputil           resource         zipimport
audiodev            inspect           rexec            zlib

Enter any module name to get more help.  Or, type "modules spam" to
search for modules whose descriptions contain the word "spam".
```

Note that the output of print statements only appears on the serial console.

One possible use for a Python script embedded on the switch is to perform configuration tasks. Such a Python script might use the telnetlib package to telnet into the switch console and perform some configuration. The following script provides the basic framework for a local telnet session to the switch console. The switch must be configured with an "admin" user, and telnet access must be allowed or this code will fail. Readers should look at the numerous articles on the Web for an explanation of the following Python code.

```
#!/usr/bin/env python
import telnetlib
import os
import re
import time
import string
import sys

HOST = '127.0.0.1'
PORT = 23
LOGIN_STRING = "Login:"
PASSWORD_STRING = "Password:"
TERMINAL_LEN_ZERO = "terminal length 0\n"
TERMINAL_MONITOR = "terminal monitor\n"
ENABLE_STRING = "enable\n"
CONFIG_STRING = "configure\n"
```

```
USERNAME = 'admin'
PASSWORD = 'password'
ENABLE_PASSWORD = ''
TIMEOUT = 3

def do_terminal_settings(tn):
    tn.write(TERMINAL_MONITOR)
    tn.read_until("#")
    tn.write(TERMINAL_LEN_ZERO)
    tn.read_until("#")

def do_login(tn):
    print "TN object created\n"
    tn.read_until(LOGIN_STRING, TIMEOUT)
    print "Read Login Prompt\n"
    tn.write(USERNAME + "\n")
    tn.read_until(PASSWORD_STRING, TIMEOUT)
    print "Read Password Prompt\n"
    tn.write(PASSWORD + "\n")
    tn.read_until(">", TIMEOUT)
    print "Received Exec Prompt\n"
    tn.write(ENABLE_STRING)
    tn.read_until("#", TIMEOUT)
    print "Received Enable Prompt\n"

def do_config(tn):
    tn.write(CONFIG_STRING)
    tn.read_until("#", TIMEOUT)
    print "Received Config Prompt\n"
    tn.write("ip routing\n");
    print "Enabled ip routing\n"
    tn.write("exit\n");
    tn.read_until("#")


def main():
    telnet =  telnetlib.Telnet(HOST,PORT)
    do_login(telnet)
    do_terminal_settings(telnet)
    do_config(telnet)
    telnet.close()
    sys.exit(0)

main()
```

A

Feature Limits and Platform Constants

Table A-1 lists the feature limits and Table A-2 lists the platform constants for the Dell Networking N-Series switches.

Certain platform constants may be adjusted by selecting a different SDM template. For example, both the Dell Networking N3000 Series switches and the Dell Networking N4000 Series switches support 16-wide ECMP using a non-default template.

Table A-1. Feature Limits

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
Switching features				
Spanning Tree				
MST Instances	31	15	15	15
RPVST VLANs	64	64	64	64
RPVST VLANs * Interfaces	1024	1024	1024	1024
Port Mirroring				
Number of monitor sessions	1	4	4	4
Max source ports in a session	192	624	624	816
RMON 1, 2, 3, 9				
Max Ether Stats entries	762	762	762	954
Max History entries	270	270	270	270
Max buckets per History entry	50	50	50	50
Max Alarm entries	32	32	32	32
Max Event entries	32	32	32	32
Max Log entries per Event entry	100	100	100	100
Management ACL (MACAL) Max Rules	64	64	64	64

Table A-1. Feature Limits (Continued)

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
Cut-through mode threshold (bytes)	–	–	–	768
VPC				
Max Number of VPCs	N/A	127	127	127
Max Member ports per VPC	N/A	8	8	8
DCPDP UDP Port Number	N/A	50000	50000	50000
Routing features				
IP Helper Max entries	64	64	512	512
VRF Max instances	–	–	14	52
Metro Ethernet features				
802.1ag				
Max number of domains	–	–	–	8
Max number of MAs per domain	–	–	–	32
Max number of MAs	–	–	–	32
Max number of RMEPs	–	–	–	64
Max number LTR entries	–	–	–	256
Management features				
HTTP Max Sessions	16	16	16	16
SSL/HTTPS Max Sessions	16	16	16	16
User management features				
User ID configuration				
Max number of configured users	8	8	8	8
Max user name length	64	64	64	64
Max password length	64	64	64	64
Max number of IAS users (internal user database)	100	100	100	100

Table A-1. Feature Limits (Continued)

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
Authentication login list				
Max Count	5	5	5	5
Max methods per list	6	6	6	6
Max name length	15	15	15	15
Authentication Enable lists				
Max Count	5	5	5	5
Max methods per list	6	6	6	6
Max name length	15	15	15	15
Authentication HTTP lists				
Max Count	1	1	1	1
Max methods per list	6	6	6	6
Max name length	15	15	15	15
Authentication HTTPS lists				
Max Count	1	1	1	1
Max methods per list	6	6	6	6
Max name length	15	15	15	15
Authentication Dot1x lists				
Max Count	1	1	1	1
Max methods per list	6	6	6	6
Max name length	15	15	15	15
Authorization Exec lists				
Max Count	5	5	5	5
Max methods per list	4	4	4	4
Max name length	20	20	20	20

Table A-1. Feature Limits (Continued)

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
Authorization Command lists				
Max Count	5	5	5	5
Max methods per list	4	4	4	4
Max name length	20	20	20	20
Accounting Exec lists				
Max Count	5	5	5	5
Max methods per list	2	2	2	2
Max name length	15	15	15	15
Accounting Commands lists				
Max Count	5	5	5	5
Max methods per list	1	1	1	1
Max name length	15	15	15	15
Login History	50	50	50	50
QoS features				
iSCSI				
Max Monitored TCP Ports/IP Addresses	N/A	16	16	16
Max Sessions	N/A	1024	1024	252
Max Connections	N/A	1024	1024	252
Stacking features				
Max physical units per stack	4	12	12	12
Max physical slots per unit	1	3	3	3
Max physical ports per slot	52	52	52	58
Max physical ports per unit	56	56	56	68
Max physical ports per stack	224	672	572	816
Max active stack ports per unit	4	2	2	8

Table A-2. Platform Constants

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
MAC addresses assigned per system	4	4	4	4
Reference CPU	ARM Cortex A9	ARM Cortex A9 Dual Core	ARM Cortex A9 Dual Core	NetLogic XLP308L Dual Core
Reference CPU speed	1 GHz	1 GHz	1 GHz	1.2 GHz
Reference RAM	1 Gbyte	1 Gbyte	1 Gbyte	2 Gbyte DDR3
Reference Flash	32 Mbyte	256 Mbyte	256 Mbyte	256 Mbyte
Maximum number of remote Telnet connections	4	4	4	4
Maximum number of remote SSH connections	5	5	5	5
Number of MAC addresses supported	16384	32768	32768	131072
Maximum Agetime in seconds	1000000	1000000	1000000	1000000
Number of VLANs	512	4096	4096	4096
Maximum VLAN ID	4093	4093	4093	4093
Number of 802.1p traffic classes	7	7	7	7
IEEE 802.1X				
Number of 802.1X clients per stack	1152	2496	2496	3264
Number of 802.1X clients per port	32	64	64	64
Number of LAGs (max. lags/ports/max dynamic LAG ports per system	64/8/144	128/8/144	128/8/144	128/8/144
Link Dependency Max Groups	24	72	72	72
Number of MAC-based VLANs supported	256	256	256	256
Number of network buffers	182	182	182	182
Number of records in log	400	400	400	400

Table A-2. Platform Constants (Continued)

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
Static filter entries				
Unicast MAC and source port	20	20	20	20
Multicast MAC and source port	20	20	20	20
Multicast MAC and destination port (only)	512	1024	1024	2048
Number of subnet-based VLANs supported	128	128	128	512
Protocol-based VLANs				
Max number of groups	128	128	128	128
Max protocols	16	16	16	16
Maximum MFDB entries	512	1024	1024	2048
IGMPv3/MLDv2 Snooping limits				
IGMPv3/MLDv2 HW entries when IP Multicast present	N/A	1024	1024	512
IGMPv3/MLDv2 HW entries when Routing w/o IP Multicast	512	4096	4096	2048
IGMPv3/MLDv2 HW entries when Switching only	512	8192	8192	4096
Jumbo frame support				
Max size supported	9216	9216	9216	9216
Number of IP Source Guard stations	253	1020	1020	378
Number of DHCP snooping bindings	16384	32768	32768	32768
Number of DHCP snooping static entries	128	1024	1024	1024

Table A-2. Platform Constants (Continued)

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
LLDP-MED				
Number of remote nodes	416	1296	1296	1632
LLDP Remote Management address buffers	1296	1296	1296	1632
LLDP Unknown TLV address buffers	100	100	100	100
LLDP Organizationally Defined TLV buffers	8424	8424	8424	10608
Port MAC locking				
Dynamic addresses per port	300	600	600	600
Static addresses per port	100	100	100	100
sFlow				
Number of samplers	224	672	672	816
Number of pollers	224	672	672	816
Number of receivers	8	8	8	8
RADIUS				
Max Authentication servers	8	32	32	32
Max Accounting servers	8	32	32	32
Number of routes (IPv4/IPv6)				
IPv4 only template	512	1024	12288	12288
IPv4/IPv6 template				
IPv4 routes				
	384	512	8160	8160
IPv6 routes				
RIP application route scaling				
	64	256	4096	4096
OSPF application route scaling				
	256	256	512	512
	N/A	N/A	8160	8160
Number of static routes (IPv4/IPv6)	256/ 64	256/ 128	1024/ 1024	1024/ 1024

Table A-2. Platform Constants (Continued)

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
OSPF				
Max OSPFv2 LSAs				
IPv4-only template	N/A	N/A	37064	36968
IPv4/IPv6 template	N/A	N/A	24680	24680
OSPFv2 max neighbors	N/A	N/A	400	400
Max OSPFv3 LSAs	N/A	N/A	12488	12488
OSPFv2 max neighbors	N/A	N/A	400	400
OSPFv3 max neighbors per interface	N/A	N/A	100	100
BGP				
BGP Route Scaling	N/A	N/A	8160	8160
BGP Peer Scaling	N/A	N/A	256	256
Tunnels				
Number of configured v6-over-v4 tunnels	N/A	N/A	8	8
Number of automatic (6to4) tunnels	N/A	N/A	1	1
Number 6to4 next hops	N/A	N/A	16	16
DHCP server				
Max number of pools	N/A	16	16	16
Total max leases	N/A	256	256	256
DNS client				
Concurrent requests	16	16	16	16
Name server entries	8	8	8	8
Search list entries	6	6	6	6
Static host entries	64	64	64	64
Cache entries	128	128	128	128
Domain search list entries	32	32	32	32

Table A-2. Platform Constants (Continued)

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
DHCPv6 Server				
Max number of pools	N/A	16	16	16
DNS domain names within a pool	N/A	5	5	5
DNS server addresses within a pool	N/A	8	8	8
Delegated prefix definitions within a pool	N/A	10	10	10
Number of VLAN routing interfaces	128	128	128	128
Number of ARP entries (Hosts)				
IPv4-only template	2042	4096	6144	6144
IPv4/IPv6 template (v4/v6)	1021/ 510	4096/ 512	4096/ 1024	4096/ 1024
Static v4 ARP entries	128	128	128	128
Number of ECMP next hops per route	0	0	4	4
Number of ECMP groups	0	0	64	1024
MLAG				
Maximum MLAGs	64	64	64	64
MRP				
MSRP Streams	–	–	–	64
MMRP MACs	–	–	–	64
MVR				
MVR Groups	64	64	64	64

Table A-2. Platform Constants (Continued)

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
IP Multicast				
Number of IPv4/IPv6 Multicast Forwarding Entries	N/A	N/A	2048 (1536 IPv4, 512 IPv6)	1024 (512 IPv4, 256 IPv6)
IGMP Group Memberships per system	N/A	N/A	2048 each for IPv4 and IPv6	2048 each for IPv4 and IPv6
DVMRP Neighbors	N/A	N/A	256	256
PIM-DM Neighbors	N/A	N/A	256	256
PIM-SM Neighbors	N/A	N/A	256	256
PIM-SM Static RP entries	N/A	N/A	5	5
PIM-SM Candidate RP Group Range entries	N/A	N/A	20	20
PIM-SM SSM range entries	N/A	N/A	5	5
IGMP Sources processed per group per message	N/A	N/A	73	73
Fan Out (max OIFs per group when all groups active)	N/A	N/A	16	16

Table A-2. Platform Constants (Continued)

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
ACL limits				
Maximum number of ACLs (any type)	100	100	100	100
Maximum number of configurable rules per list.	1023	1023	1023	1023
Maximum ACL Rules per Interface and Direction (IPv4/L2)	1012 ing 511 egr	1023 ing 1023 egr	2048 ing 1023 egr	1023 ing 1023 egr
Maximum ACL Rules per Interface and Direction (IPv6)	378 ing 253 egr	1023 ing 509 egr	1659 ing 509 egr	635 ing 509 egr
Maximum ACL Rules (system-wide)	3072	3914	3914	3060
Maximum VLAN interfaces with ACLs applied	24	24	24	24
Maximum ACL Logging Rules (system-wide)	128	128	128	128
CoS Device Characteristics				
Configurable Queues per port (stacking/nonstacking)	7	7	7	7
Configurable Drop Precedence levels	3	3	3	3

Table A-2. Platform Constants (Continued)

Feature	N1500 Series	N2000 Series	N3000 Series	N4000 Series
DiffServ Device Limits				
Number of queues (stacking/nonstacking)	7	7	7	7
Max Rules per Class	13	13	13	13
Max Instances per Policy	12	12	12	12
Max Attributes per Instance	3	3	3	3
Max Service Interfaces	800	800	800	944
Max table entries				
Class Table	32	32	32	32
Class Rule Table	416	416	416	416
Policy Table	64	64	64	64
Policy Instance Table	768	768	768	768
Policy Attribute Table	2304	2304	2304	2304
Max Nested Class Reference Chain Rule Count	26	26	26	26
Datacenter Device Limits				
PFC number of lossless priorities	N/A	N/A	N/A	2
ETS number of traffic class groups	N/A	N/A	N/A	3
AutoVoIP number of voice calls	16	16	16	16
Voice VLAN number of devices	192	192	672	816

B

System Process Definitions

The following process/thread definitions are intended to assist the end user in troubleshooting switch issues. Only the most often seen threads/processes are listed here. Other processes or threads may be seen occasionally but are not a cause for concern.

Table B-1. System Process Definitions

Name	Task Summary
aclClusterTask aclEventTask aclLogTask	ACL tasks
ARP Timer	ARP tasks
autoInstTask	Auto Install task - USB, etc.
bcmATP-RX bcmATP-TX	BCM system task: Acknowledged Transport Protocol
bcmCNTR.0	BCM system task: SDK Statistics collection
bcmDISC	BCM system task: SDK Discovery task
bcmDPC	BCM system task: SDK DPC task
bcmL2X.0	BCM system task: SDK L2 SOC shadow table maintenance
bcmLINK.0	BCM system task: SDK Physical link status monitor
bcmNHOP	BCM system task: SDK transport Next Hop task
bcmRLINK	BCM system task:SDK Remote registration last
bcmRPC	BCM system task:SDK Remote registration last
bcmRX	BCM system task: SDK Control plane packet receiver/dispatcher
bcmTUNQ	BCM system task: SDK transport queueing task
bcmTX	BCM system task: SDK Control plane packet transmitter

Table B-1. System Process Definitions (Continued)

Name	Task Summary
bcmXGS3AsyncTask	BCM system task: SDK XGX3 hw task
BootP	Boot Loader
boxs Req	Box Services Request (temperature, power, fan)
boxs Resp	Box Services Response (temperature, power, fan)
boxs Timer	Box Services Response (temperature, power, fan)
cdaFftpTask	Code Distribution Administrator FTP task
cdaStatusTask	Code Distribution Administrator Status task
cdaUpdateTask	Code Distribution Administrator Update task
cliWebIORedirectTask	CLI Web IO Redirection Task
cmgrInsertTask	Card Manager Insertion Handler
cmgrTask	Card Manager Status (built-in and plug-in card configuration processing)
Cnfgr_Thread	Configurator (startup manager)
CP Wired If	Captive Portal
cpuUtilMonitorTask	CPU Utilities monitor
DapiDebugTask	Device API debug processing
DHCP Server Processing Task DHCP snoop dhcpsPingTask DHCPv4 Client Task DHCPv6 Client Task DHCPv6 Server Task	DHCP Tasks
dnsRxTask dnsTask	DNS tasks
dosTask	Denial of Service task
dot1qTask	VLAN routing task

Table B-1. System Process Definitions (Continued)

Name	Task Summary
Dot1s transport task dot1s_helper_task dot1s_task dot1s_timer_task	Spanning Tree tasks
dot1xTask dot1xTimerTask	802.1x authentication tasks
dot3ad_core_task dot3ad_core_ac_task dot3ad_helper_task dot3ad_timer_task	Link aggregation tasks
dtlAddrTask dtlTask	Device Transform Layer - Silicon Integration Layer
dvmrpMapTask	DVMRP Mapping Layer
Dynamic ARP Inspection	Dynamic ARP Inspection task
EDB	Entity MIB Processing task
EDB Trap	Entity MIB Trap task
emWeb	UI processing task
envMonTask	Environment Monitor (fans, power supplies, temperature, ...)
fdbTask	Forwarding Data Base Manager
fftpTask	FTP processing
gccp_t	GARP Central Control Point task (dot 1d)

Table B-1. System Process Definitions (Continued)

Name	Task Summary
hapiBpduTxTask hapiL2AsyncTask hapiL2FlushTask hapiL3AsyncTask hapiLinkStatusTask hapiMcAsyncTask hapiRxTask hapiTxTask	High Level API - SDK Integration Layer
hpcBroadRpcTask	SDK Remote messaging task.
ip6MapExceptionDataTask ip6MapLocalDataTask ip6MapNbrDiscTask ip6MapProcessingTask ip6MapRadvdTask ipcom_sysl IpHelperTask ipMapForwardingTask ipMapProcessingTask ipnetd	IP Stack
iscsiTask	ISCSI task
isdpTask	ISDP task
lldpTask	LLDP task
LOG	System LOG processing
LOGC	System LOG processing
MAC Age Task	MAC address table aging
MAC Send Task	MAC address table learning
macalTask	Management ACL packet processing

Table B-1. System Process Definitions (Continued)

Name	Task Summary
mcastMapTask mgmdMapTask	Multicast Mapping Tasks
mvrTask	MVR Message Handler
nim_t	Network Interface Manager
osapiMonTask	System Task Monitor
osapiTimer	Application timer service
osapiWdTask	Hardware watchdog timer service
OSPF mapping Task OSPF Proto OSPFV3 mapping Task OSPFV3 recvmsg Task OSPFv3 Proto	OSPF tasks
pimdmMapTask	PIMDM task
pimsmMapTask	PIMSM task
pingAsync	Ping response processing
pktRcvrTask	Multicast control plane packet receiver/dispatch
pmlTask	Port MAC Locking management task
portAggTask	Port Aggregator task
radius_rx_task radius_task	RADIUS server tasks
ripMapProcessingTask	RIP Mapping layer
RLIM cnfgr task	VRRP configuration
RLIM task	VRRP message processing
RMONTask	RMON Statistics Collection
serialInput	Serial Input task
sFlowTask	sFlow task
SimAddrConflictTask	System Interface Manager Address Conflict Task

Table B-1. System Process Definitions (Continued)

Name	Task Summary
simPts_task	System Interface Manager (time zone, system name, service port config, file transfers, ...)
SNMPCTTask SNMPSaveCfgTask SNMPTask SNMPTrapTask	SNMP Tasks
snoopTask	IGMP/MLD Snooping packet processing
SNTP SNTPC	SNTP tasks
spmTask	Stack port manager - stacking control plane packet processing
sshdEvTask	SSH task
ssltTask	SSL task
Stk Mgr Task	Stack Manager Task
tacacs_rx_task tacacs_task	TACACS tasks
tArpCallback tArpReissue	ARP tasks
tArpTimerExp	ARP Timer Expiry
tCpktSvc	NSF Processing
tCptvPrtl	Captive portal control plane processing
tDhcp6sTask tDhcpsTask	DHCP Tasks
tEmWeb	Web page server
TimeRange Processing Task	ACL Time Ranges
tIomEvtMon	CMC Communication
tL7Timer0	System Timer
tLogTask	System LOG processing

Table B-1. System Process Definitions (Continued)

Name	Task Summary
TransferTask	TFTP Processing
trapTask	Trap handler
tRipTask	RIP Routing
tRtrDiscProcessingTask	Router Discovery packet processing
usbFlashDriveTask	USB Flash driver processing
umCfgUpdateTask umWorkerTask unitMgrTask	Stack Management: Unit Manager tasks
USL Worker Task	USL Message processing (primarily MAC address table CLI commands)
UtilTask	Mgmt. UI login/logout processing
voipTask	Voice Over IP
VRRPdaemon	VRRP task

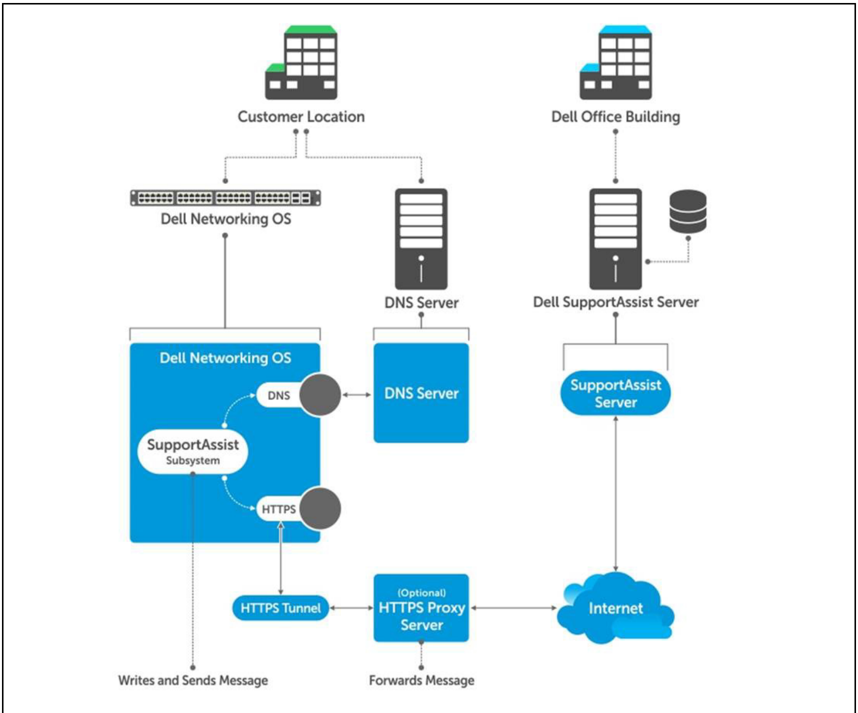
C

Dell SupportAssist

Dell SupportAssist sends troubleshooting data securely to Dell. SupportAssist in this Dell Networking OS release does not support automated email notification at the time of hardware fault alert, automatic case creation, automatic part dispatch, or reports. SupportAssist requires Dell Networking OS 6.3 and SupportAssist Package XYZ or later to be installed on the Dell Networking device.

Dell SupportAssist is enabled by default on all Dell Networking switches. To disable Dell SupportAssist, enter the command **eula-consent support-assist reject** in Global Configuration mode and then save the configuration.

Figure C-1. Dell Support Assist



Dell SupportAssist operates by periodically reporting switch identity (service tag and serial number), configuration, logs, status, and diagnostic information to an external SupportAssist server operated by Dell, Inc. Information is logged periodically on the SupportAssist server.

It is recommended that Dell Networking customers utilizing Dell SupportAssist configure the appropriate contact information using the **contact-person** and **contact-company** commands in Support-Assist Configuration mode.

The Dell SupportAssist EULA is printed here:

> I accept the terms of the license agreement. You can reject the license agreement by configuring this command 'eula-consent support-assist reject'.
>
> By installing SupportAssist, you allow Dell to save your contact information (e.g. name, phone number and/or email address) which would be used to provide technical support for your Dell products and services. Dell may use the information for providing recommendations to improve your IT infrastructure.
> Dell SupportAssist also collects and stores machine diagnostic information, which may include but is not limited to configuration information, user supplied contact information, names of data volumes, IP addresses, access control lists, diagnostics & performance information, network configuration information, host/server configuration & performance information and related data (Collected Data) and transmits this information to Dell. By downloading SupportAssist and agreeing to be bound by these terms and the Dell end user license agreement, available at: http://www.dell.com/aeula, you agree to allow Dell to provide remote monitoring services of your IT environment and you give Dell the right to collect the Collected Data in accordance with Dell's Privacy Policy, available at: http://www.dell.com/privacypolicycountryspecific, in order to enable the performance of all of the various functions of SupportAssist during your entitlement to receive related repair services from Dell. You further agree to allow Dell to transmit and store the Collected Data from SupportAssist in accordance with these terms. You agree that the provision of SupportAssist may involve international transfers of data from you to Dell and/or to Dell's affiliates, subcontractors or business partners. When making such transfers, Dell shall ensure appropriate protection is in place to safeguard the Collected Data being transferred in connection with SupportAssist. If you are downloading SupportAssist on behalf

of a company or other legal entity, you are further certifying to Dell that you have appropriate authority to provide this consent on behalf of that entity. If you do not consent to the collection, transmission and/or use of the Collected Data, you may not download, install or otherwise use SupportAssist.

Index

Numerics

A

B

C

D

G

H

I

J

L

M

N

O

P

Q

R

S

T

U

V

W