Lenovo RackSwitch G8124-E

Application Guide

for Lenovo Networking OS 8.3

Lenovo™

Note: Before using this information and the product it supports, read the general information in the *Safety information and Environmental Notices and User Guide* documents on the Lenovo *Documentation* CD and the *Warranty Information* document that comes with the product.

Contents

Preface

This *Application Guide* describes how to configure and use the Lenovo Networking OS 8.3 software on the Lenovo RackSwitch G8124-E (referred to as G8124-E throughout this document). For documentation on installing the switch physically, see the *Installation Guide* for your G8124-E.

Who Should Use This Guide

This guide is intended for network installers and system administrators engaged in configuring and maintaining a network. The administrator should be familiar with Ethernet concepts, IP addressing, Spanning Tree Protocol, and SNMP configuration parameters.

What You'll Find in This Guide

This guide will help you plan, implement, and administer Lenovo N/OS software. Where possible, each section provides feature overviews, usage examples, and configuration instructions. The following material is included:

Part 1: Getting Started

This material is intended to help those new to N/OS products with the basics of switch management. This part includes the following chapters:

- Chapter 1, "Switch Administration," describes how to access the G8124-E to configure the switch and view switch information and statistics. This chapter discusses a variety of manual administration interfaces, including local management via the switch console, and remote administration via Telnet, a web browser, or via SNMP.

- Chapter 2, "Initial Setup," describes how to use the built-in Setup utility to perform first-time configuration of the switch.

- Chapter 3, "Switch Software Management," describes how to update the N/OS software operating on the switch.

Part 2: Securing the Switch

- Chapter 4, "Securing Administration," describes methods for using Secure Shell for administration connections, and configuring end-user access control.

- Chapter 5, "Authentication & Authorization Protocols," describes different secure administration for remote administrators. This includes using Remote Authentication Dial-in User Service (RADIUS), as well as TACACS+ and LDAP.

- Chapter 6, "Access Control Lists," describes how to use filters to permit or deny specific types of traffic, based on a variety of source, destination, and packet attributes.

Part 3: Switch Basics

- Chapter 7, "VLANs," describes how to configure Virtual Local Area Networks (VLANs) for creating separate network segments, including how to use VLAN tagging for devices that use multiple VLANs. This chapter also describes Protocol-based VLANs, and Private VLANs.

- Chapter 8, "Ports and Link Aggregation," describes how to group multiple physical ports together to aggregate the bandwidth between large-scale network devices.

- Chapter 9, "Spanning Tree Protocols," discusses how Spanning Tree Protocol (STP) configures the network so that the switch selects the most efficient path when multiple paths exist. Covers Rapid Spanning Tree Protocol (RSTP), Per-VLAN Rapid Spanning Tree (PVRST), and Multiple Spanning Tree Protocol (MSTP).

- Chapter 10, "Virtual Link Aggregation Groups," describes using Virtual Link Aggregation Groups (VLAGs) to form LAGs spanning multiple VLAG-capable aggregator switches.

- Chapter 11, "Quality of Service," discusses Quality of Service (QoS) features, including IP filtering using Access Control Lists (ACLs), Differentiated Services, and IEEE 802.1p priority values.

Part 4: Advanced Switching Features

- Chapter 12, "Deployment Profiles," describes how the G8124-E can operate in different modes for different deployment scenarios, adjusting switch capacity levels to optimize performance for different types of networks.

- Chapter 13, "Virtualization," provides an overview of allocating resources based on the logical needs of the data center, rather than on the strict, physical nature of components.

- Chapter 14, "Virtual NICs," discusses using virtual NIC (vNIC) technology to divide NICs into multiple logical, independent instances.

- Chapter 15, "VMready," discusses virtual machine (VM) support on the G8124-E.

- Chapter 16, "FCoE and CEE," discusses using various Converged Enhanced Ethernet (CEE) features such as Priority-based Flow Control (PFC), Enhanced Transmission Selection (ETS), and FIP Snooping for solutions such as Fibre Channel over Ethernet (FCoE).

- Chapter 18, "Dynamic ARP Inspection." discusses this security feature that lets a switch intercept and examine all ARP request and response packets in a subnet, discarding those packets with invalid IP to MAC address bindings. This capability protects the network from man-in-the-middle attacks.

Part 5: IP Routing

- Chapter 19, "Basic IP Routing," describes how to configure the G8124-E for IP routing using IP subnets, BOOTP, and DHCP Relay.

- Chapter 20, "Routed Ports describes how to configure a switch port to forward Layer 3 traffic.

- Chapter 21, "Internet Protocol Version 6," describes how to configure the G8124-E for IPv6 host management.

- Chapter 22, "IPsec with IPv6," describes how to configure Internet Protocol Security (IPsec) for securing IP communications by authenticating and encrypting IP packets, with emphasis on Internet Key Exchange version 2, and authentication/confidentiality for OSPFv3.

- Chapter 23, "Routing Information Protocol," describes how the N/OS software implements standard Routing Information Protocol (RIP) for exchanging TCP/IP route information with other routers.

- Chapter 24, "Internet Group Management Protocol," describes how the N/OS software implements IGMP Snooping or IGMP Relay to conserve bandwidth in a multicast-switching environment.

- Chapter 25, "Multicast Listener Discovery," describes how Multicast Listener Discovery (MLD) is used with IPv6 to support host users requests for multicast data for a multicast group.

- Chapter 26, "Border Gateway Protocol," describes Border Gateway Protocol (BGP) concepts and features supported in N/OS.

- Chapter 27, "Open Shortest Path First," describes key Open Shortest Path First (OSPF) concepts and their implemented in N/OS, and provides examples of how to configure your switch for OSPF support.

- Chapter 28, "Protocol Independent Multicast," describes how multicast routing can be efficiently accomplished using the Protocol Independent Multicast (PIM) feature.

Part 6: High Availability Fundamentals

- Chapter 29, "Basic Redundancy," describes how the G8124-E supports redundancy through LAGs and hotlinks.

- Chapter 30, "Layer 2 Failover," describes how the G8124-E supports high-availability network topologies using Layer 2 Failover.

- Chapter 31, "Virtual Router Redundancy Protocol," describes how the G8124-E supports high-availability network topologies using Virtual Router Redundancy Protocol (VRRP).

Part 7: Network Management

- Chapter 32, "Link Layer Discovery Protocol," describes how Link Layer Discovery Protocol helps neighboring network devices learn about each others' ports and capabilities.

- Chapter 33, "Simple Network Management Protocol," describes how to configure the switch for management through an SNMP client.

- Chapter 34, "NETCONF," describes how to manage the G8124-E using Network Configuration Protocol (NETCONF), a mechanism based on the Extensible Markup Language (XML).

Part 8: Monitoring

- Chapter 35, "Remote Monitoring," describes how to configure the RMON agent on the switch, so that the switch can exchange network monitoring data.

- Chapter 36, "sFlow, described how to use the embedded sFlow agent for sampling network traffic and providing continuous monitoring information to a central sFlow analyzer.

- Chapter 37, "Port Mirroring," discusses tools how copy selected port traffic to a monitor port for network analysis.

Part 9: Appendices

- Appendix A, "Glossary," describes common terms and concepts used throughout this guide.

- Appendix A, "Getting help and technical assistance," provides details on where to go for additional information about Lenovo and Lenovo products.

- Appendix B, "Notices," contains safety and environmental notices.

Additional References

Additional information about installing and configuring the G8124-E is available in the following guides:

- *Lenovo RackSwitch G8124-E Installation Guide*
- *Lenovo RackSwitch G8124-E ISCLI Command Reference for Networking OS 8.3*
- *Lenovo RackSwitch G8124-E Release Notes for Networking OS 8.3*

Typographic Conventions

The following table describes the typographic styles used in this book.

Table 1. *Typographic Conventions*

Typeface or Symbol	Meaning	Example
ABC123	This type is used for names of commands, files, and directories used within the text. It also depicts on-screen computer output and prompts.	View the readme.txt file. Main#
ABC123	This bold type appears in command examples. It shows text that must be typed in exactly as shown.	Main# **sys**
<ABC123>	This italicized type appears in command examples as a parameter placeholder. Replace the indicated text with the appropriate real name or value when using the command. Do not type the brackets. This also shows book titles, special terms, or words to be emphasized.	To establish a Telnet session, enter: host# **telnet** *<IP address>* Read your *User's Guide* thoroughly.
[ ]	Command items shown inside brackets are optional and can be used or excluded as the situation demands. Do not type the brackets.	host# **ls** [**-a**]
|	The vertical bar (|) is used in command examples to separate choices where multiple options exist. Select only one of the listed options. Do not type the vertical bar.	host# **set left|right**
AaBbCc123	This block type depicts menus, buttons, and other controls that appear in Web browsers and other graphical interfaces.	Click the **Save** button.

●

Part 1: Getting Started

Chapter 1. Switch Administration

Your RackSwitch G8124-E (G8124-E) is ready to perform basic switching functions right out of the box. Some of the more advanced features, however, require some administrative configuration before they can be used effectively.

The extensive Lenovo Networking OS switching software included in the G8124-E provides a variety of options for accessing the switch to perform configuration, and to view switch information and statistics.

This chapter discusses the various methods that can be used to administer the switch.

Administration Interfaces

Lenovo N/OS provides a variety of user-interfaces for administration. These interfaces vary in character and in the methods used to access them: some are text-based, and some are graphical; some are available by default, and some require configuration; some can be accessed by local connection to the switch, and others are accessed remotely using various client applications. For example, administration can be performed using any of the following:

● A built-in, text-based command-line interface and menu system for access via serial-port connection or an optional Telnet or SSH session

● The built-in Browser-Based Interface (BBI) available using a standard web-browser

● SNMP support for access through network management software such as IBM Director or HP OpenView

The specific interface chosen for an administrative session depends on user preferences, as well as the switch configuration and the available client tools.

In all cases, administration requires that the switch hardware is properly installed and turned on. (see the *RackSwitch G8124-E Installation Guide*).

Industry Standard Command Line Interface

The Industry Standard Command Line Interface (ISCLI) provides a simple, direct method for switch administration. Using a basic terminal, you can issue commands that allow you to view detailed information and statistics about the switch, and to perform any necessary configuration and switch software maintenance.

You can establish a connection to the ISCLI in any of the following ways:
● Serial connection via the serial port on the G8124-E (this option is always available)
● Telnet connection over the network
● SSH connection over the network

Browser-Based Interface

The Browser-based Interface (BBI) provides access to the common configuration, management and operation features of the G8124-E through your Web browser.

Establishing a Connection

The factory default settings permit initial switch administration through *only* the built-in serial port. All other forms of access require additional switch configuration before they can be used.

Remote access using the network requires the accessing terminal to have a valid, routable connection to the switch interface. The client IP address may be configured manually, or an IPv4 address can be provided automatically through the switch using a service such as DHCP or BOOTP relay (see "BOOTP/DHCP Client IP Address Services" on page 38), or an IPv6 address can be obtained using IPv6 stateless address configuration.

Note: Throughout this manual, *IP address* is used in places where either an IPv4 or IPv6 address is allowed. IPv4 addresses are entered in dotted-decimal notation (for example, 10.10.10.1), while IPv6 addresses are entered in hexadecimal notation (for example, 2001:db8:85a3::8a2e:370:7334). In places where only one type of address is allowed, *IPv4 address* or *IPv6 address* is specified.

Using the Switch Management Ports

To manage the switch through the management ports, you must configure an IP interface for each management interface. Configure the IPv4 address/mask and default gateway address:

1. Log on to the switch.

2. Enter Global Configuration mode.

```
RS G8124-E> enable
RS G8124-E# configure terminal
```

3. Configure a management IP address and mask.

   The switch reserves four management interfaces:

   ● Using IPv4:

```
RS G8124-E(config)# interface ip [127|128]
RS G8124-E(config-ip-if)# ip address <management interface IPv4 address>
RS G8124-E(config-ip-if)# ip netmask <IPv4 subnet mask>
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

   ○ IF 127 supports IPv4 management port A and uses IPv4 default gateway 3.

   ○ IF 128 supports IPv4 management port B and uses IPv4 default gateway 4.

- Using IPv6:

```
RS G8124-E(config)# interface ip [125|126]
RS G8124-E(config-ip-if)# ipv6 address <management interface IPv6 address>
RS G8124-E(config-ip-if)# ipv6 prefixlen <IPv6 prefix length>
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

- o IF 125 supports IPv6 management port A and uses IPv6 default gateway 3.
- o IF 126 supports IPv6 management port B and uses IPv6 default gateway 4.

4. Configure the appropriate default gateway.

- If using IPv4, IPv4 gateway 3 is required for IF 127, and IPv4 gateway 4 is required for IF 128.

```
RS G8124-E(config)# ip gateway [3|4] address <default gateway IPv4 address>
RS G8124-E(config)# ip gateway [3|4] enable
```

- If using IPv6, IPv6 gateway 3 is required for IF 125, and IPv4 gateway 4 is required for IF 126.

```
RS G8124-E(config)# ip gateway6 [3|4] address <default gateway IPv6 address>
RS G8124-E(config)# ip gateway6 [3|4] enable
```

Once you configure a management IP address for your switch, you can connect to a management port and use the Telnet program from an external management station to access and control the switch. The management port provides *out-of-band* management.

Using the Switch Data Ports

You also can configure *in-band* management through any of the switch data ports. To allow in-band management, use the following procedure:

1. Log on to the switch.

2. Enter IP interface mode.

```
RS G8124-E> enable
RS G8124-E# configure terminal
RS G8124-E(config)# interface ip <IP interface number>
```

Note: Interface 125 through 128 are reserved for out-of-band management (see).

3. Configure the management IP interface/mask.

- Using IPv4:

```
RS G8124-E(config-ip-if)# ip address <management interface IPv4 address>
RS G8124-E(config-ip-if)# ip netmask <IPv4 subnet mask>
```

- Using IPv6:

```
RS G8124-E(config-ip-if)# ipv6 address <management interface IPv6 address>
RS G8124-E(config-ip-if)# ipv6 prefixlen <IPv6 prefix length>
```

4. Configure the VLAN, and enable the interface.

```
RS G8124-E(config-ip-if)# vlan 1
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

5. Configure the default gateway.

- If using IPv4:

```
RS G8124-E(config)# ip gateway <gateway number> address <IPv4 address>
RS G8124-E(config)# ip gateway <gateway number> enable
```

- If using IPv6:

```
RS G8124-E(config)# ip gateway6 <gateway number> address <IPv6 address>
RS G8124-E(config)# ip gateway6 <gateway number> enable
```

Note: IPv4 gateway 1 and 2, and IPv6 gateway 1, are used for in-band data networks. IPv4 and IPv6 gateways 3 and 4 are reserved for out-of-band management ports (see "Using the Switch Management Ports" on page 29).

Once you configure the IP address and have a network connection, you can use the Telnet program from an external management station to access and control the switch. Once the default gateway is enabled, the management station and your switch do not need to be on the same IP subnet.

The G8124-E supports an industry standard command-line interface (ISCLI) that you can use to configure and control the switch over the network using the Telnet program. You can use the ISCLI to perform many basic network management functions. In addition, you can configure the switch for management using an SNMP-based network management system or a Web browser.

For more information, see the documents listed in "Additional References" on page 22.

Using Telnet

A Telnet connection offers the convenience of accessing the switch from a workstation connected to the network. Telnet access provides the same options for user and administrator access as those available through the console port.

By default, Telnet access is enabled. Use the following commands to disable or re-enable Telnet access:

```
RS G8124-E(config)# [no] access telnet enable
```

Once the switch is configured with an IP address and gateway, you can use Telnet to access switch administration from any workstation connected to the management network.

To establish a Telnet connection with the switch, run the Telnet program on your workstation and issue the following Telnet command:

```
telnet   <switch IPv4 or IPv6 address>
```

You will then be prompted to enter a password as explained .

Two attempts are allowed to log in to the switch. After the second unsuccessful attempt, the Telnet client is disconnected via TCP session closure.

Using Secure Shell

Although a remote network administrator can manage the configuration of a G8124-E via Telnet, this method does not provide a secure connection. The Secure Shell (SSH) protocol enables you to securely log into another device over a network to execute commands remotely. As a secure alternative to using Telnet to manage switch configuration, SSH ensures that all data sent over the network is encrypted and secure.

The switch can do only one session of key/cipher generation at a time. Thus, a SSH/SCP client will not be able to login if the switch is doing key generation at that time. Similarly, the system will fail to do the key generation if a SSH/SCP client is logging in at that time.

The supported SSH encryption and authentication methods are:

- Server Host Authentication: Client RSA-authenticates the switch when starting each connection

- Key Exchange: ecdh-sha2-nistp521, ecdh-sha2-nistp384, ecdh-sha2-nistp256, ecdh-sha2-nistp224, ecdh-sha2-nistp192, rsa2048-sha256, rsa1024-sha1, diffie-hellman-group-exchange-sha256, diffie-hellman-group-exchange-sha1, diffie-hellman-group14-sha1, diffie-hellman-group1-sha1

- Encryption: aes128-ctr, aes128-cbc, rijndael128-cbc, blowfish-cbc,3des-cbc, arcfour256, arcfour128, arcfour

- MAC: hmac-sha1, hmac-sha1-96, hmac-md5, hmac-md5-96

- User Authentication: Local password authentication, public key authentication, RADIUS, TACACS+

Lenovo Networking OS implements the SSH version 2.0 standard and is confirmed to work with SSH version 2.0-compliant clients such as the following:

- OpenSSH_5.4p1 for Linux

- Secure CRT Version 5.0.2 (build 1021)

- Putty SSH release 0.60

Using SSH with Password Authentication

By default, the SSH feature is disabled. Once the IP parameters are configured and the SSH service is enabled, you can access the command line interface using an SSH connection.

To establish an SSH connection with the switch, run the SSH program on your workstation by issuing the SSH command, followed by the switch IPv4 or IPv6 address:

```
# ssh <switch IP address>
```

You will then be prompted to enter a password as explained "Switch Login Levels" on page 43.

Using SSH with Public Key Authentication

SSH can also be used for switch authentication based on asymmetric cryptography. Public encryption keys can be uploaded on the switch and used to authenticate incoming login attempts based on the clients' private encryption key pairs. After a predefined number of failed public key login attempts, the switch reverts to password-based authentication.

To set up public key authentication:

1. Enable SSH:

```
RS G8124-E(config)# ssh enable
```

2. Import the public key file using SFTP or TFTP for the admin user account::

```
RS G8124-E(config)# copy {sftp|tftp} public-key
Port type ["DATA"/"MGT"]: mgt
Address or name of remote host: 9.43.101.151
Source file name: 11.key
Username of the public key: admin
Confirm download operation (y/n) ? y
```

Notes:

- When prompted to input a username, a valid user account name must be entered. If no username is entered, the key is stored on the switch, and can be assigned to a user account later.
- **Note:** A user account can have up to 100 public keys set up on the switch.

3. Configure a maximum number of 3 failed public key authentication attempts before the system reverts to password-based authentication:

```
RS G8124-E(config)# ssh maxauthattempts 3
```

Once the public key is configured on the switch, the client can use SSH to login from a system where the private key pair is set up:

```
# ssh <switch IP address>
```

Using a Web Browser

The switch provides a Browser-Based Interface (BBI) for accessing the common configuration, management, and operation features of the G8124-E through your Web browser.

By default, BBI access via HTTP is enabled on the switch.

You can also access the BBI directly from an open Web browser window. Enter the URL using the IP address of the switch interface (for example, http://*<IPv4 or IPv6 address>*).

Configuring HTTP Access to the BBI

By default, BBI access via HTTP is enabled on the switch.

To disable or re-enable HTTP access to the switch BBI, use the following commands:

```
RS G8124-E(config)# access http enable (Enable HTTP access)
   -or-
RS G8124-E(config)# no access http enable   (Disable HTTP access)
```

The default HTTP web server port to access the BBI is port 80. However, you can change the default Web server port with the following command:

```
RS G8124-E(config)# access http port <TCP port number>
```

To access the BBI from a workstation, open a Web browser window and type in the URL using the IP address of the switch interface (for example, http://*<IPv4 or IPv6 address>*).

Configuring HTTPS Access to the BBI

The BBI can also be accessed via a secure HTTPS connection over management and data ports.

1. Enable HTTPS.

   By default, BBI access is enabled via both HTTP and HTTPS on the switch. If HTTPS access has been disabled, use the following command to enable BBI Access via HTTPS:

```
RS G8124-E(config)# access https enable
```

2. Set the HTTPS server port number (optional).

   To change the HTTPS Web server port number from the default port 443, use the following command:

```
RS G8124-E(config)# access https port <x>
```

3. Generate the HTTPS certificate.

Accessing the BBI via HTTPS requires that you generate a certificate to be used during the key exchange. A default certificate is created the first time HTTPS is enabled, but you can create a new certificate defining the information you want to be used in the various fields.

```
RS G8124-E(config)# access https generate-certificate
Country Name (2 letter code) [US]:
State or Province Name (full name) [CA]:
Locality Name (eg, city) [Santa Clara]:
Organization Name (eg, company) [Lenovo Networking Operating System]:
Organizational Unit Name (eg, section) [Network Engineering]:
Common Name (eg, YOUR name) [0.0.0.0]:
Email (eg, email address) []:
Confirm generating certificate? [y/n]: y
Generating certificate. Please wait (approx 30 seconds)
restarting SSL agent
```

4. Save the HTTPS certificate.

The certificate is valid only until the switch is rebooted. To save the certificate so it is retained beyond reboot or power cycles, use the following command:

```
RS G8124-E(config)# access https save-certificate
```

When a client (such as a web browser) connects to the switch, the client is asked to accept the certificate and verify that the fields match what is expected. Once BBI access is granted to the client, the BBI can be used.

Browser-Based Interface Summary

The BBI is organized at a high level as follows:

Context buttons—These buttons allow you to select the type of action you wish to perform. The *Configuration* button provides access to the configuration elements for the entire switch. The *Statistics* button provides access to the switch statistics and state information. The *Dashboard* button allows you to display the settings and operating status of a variety of switch features.

Navigation Window—Provides a menu of switch features and functions:

- **System**—Provides access to the configuration elements for the entire switch.
- **Switch Ports**—Configure each of the physical ports on the switch.
- **Port-Based Port Mirroring**—Configure port mirroring behavior.
- **Layer 2**—Configure Layer 2 features for the switch.
- **RMON Menu**—Configure Remote Monitoring features for the switch.
- **Layer 3**—Configure Layer 3 features for the switch.
- **QoS**—Configure Quality of Service features for the switch.
- **Access Control**—Configure Access Control Lists to filter IP packets.
- **CEE**Configure Converged Enhanced Ethernet (CEE).
- **FCoE**Configure FibreChannel over Ethernet (FCoE).
- **Virtualization**—Configure vNICs and VMready for virtual machines (VMs).

- Dove Gateway – Configure Distributed Overlay Virtual Ethernet.

Using Simple Network Management Protocol

N/OS provides Simple Network Management Protocol (SNMP) version 1, version 2, and version 3 support for access through any network management software, such as IBM Director or HP-OpenView.

Note: SNMP read and write functions are enabled by default. For best security practices, if SNMP is not needed for your network, it is recommended that you disable these functions prior to connecting the switch to the network.

To access the SNMP agent on the G8124-E, the read and write community strings on the SNMP manager must be configured to match those on the switch. The default read community string on the switch is `public` and the default write community string is `private`.

The read and write community strings on the switch can be configured using the following commands:

```
RS G8124-E(config)# snmp-server read-community <1-32 characters>
    -and-
RS G8124-E(config)# snmp-server write-community <1-32 characters>
```

The SNMP manager must be able to reach any one of the IP interfaces on the switch.

For the SNMP manager to receive the SNMPv1 traps sent out by the SNMP agent on the switch, configure the trap host on the switch with the following commands:

```
RS G8124-E(config)# snmp-server trap-source <trap source IP interface>
RS G8124-E(config)# snmp-server host <IPv4 address> <trap host community string>
```

To restrict SNMP access to specific IPv4 subnets, use the following commands:

```
RS G8124-E(config)# access management-network <IPv4 address> <subnet mask>
snmp-ro
    -and-
RS G8124-E(config)# access management-network <IPv4 address> <subnet mask>
snmp-rw
```

For IPv6 networks, use:

```
RS G8124-E(config)# access management-network6 <IPv6 address> <IPv6 prefix length>
snmp-ro
    -and-
RS G8124-E(config)# access management-network6 <IPv6 address> <IPv6 prefix length>
snmp-rw
```

Note: Subnets allowed for SNMP read-only access must not overlap with subnets allowed for SNMP read-write access.

For more information on SNMP usage and configuration, see Chapter 33, "Simple Network Management Protocol."

BOOTP/DHCP Client IP Address Services

For remote switch administration, the client terminal device must have a valid IP address on the same network as a switch interface. The IP address on the client device may be configured manually, or obtained automatically using IPv6 stateless address configuration, or an IPv4 address may obtained automatically via BOOTP or DHCP relay as discussed in the next section.

The G8124-E can function as a relay agent for Bootstrap Protocol (BOOTP) or DHCP. This allows clients to be assigned an IPv4 address for a finite lease period, reassigning freed addresses later to other clients.

Acting as a relay agent, the switch can forward a client's IPv4 address request to up to five BOOTP/DHCP servers. In addition to the five global BOOTP/DHCP servers, up to five domain-specific BOOTP/DHCP servers can be configured for each of up to 10 VLANs.

When a switch receives a BOOTP/DHCP request from a client seeking an IPv4 address, the switch acts as a proxy for the client. The request is forwarded as a UDP Unicast MAC layer message to the BOOTP/DHCP servers configured for the client's VLAN, or to the global BOOTP/DHCP servers if no domain-specific BOOTP/DHCP servers are configured for the client's VLAN. The servers respond to the switch with a Unicast reply that contains the IPv4 default gateway and the IPv4 address for the client. The switch then forwards this reply back to the client.

DHCP is described in RFC 2131, and the DHCP relay agent supported on the G8124-E is described in RFC 1542. DHCP uses UDP as its transport protocol. The client sends messages to the server on port 67 and the server sends messages to the client on port 68.

BOOTP and DHCP relay are collectively configured using the BOOTP commands and menus on the G8124-E.

Global BOOTP Relay Agent Configuration

To enable the G8124-E to be a BOOTP (or DHCP) forwarder, enable the BOOTP relay feature, configure up to four global BOOTP server IPv4 addresses on the switch, and enable BOOTP relay on the interface(s) on which the client requests are expected.

Generally, it is best to configure BOOTP for the switch IP interface that is closest to the client, so that the BOOTP server knows from which IPv4 subnet the newly allocated IPv4 address will come.

In the G8124-E implementation, there are no primary or secondary BOOTP servers. The client request is forwarded to all the global BOOTP servers configured on the switch (if no domain-specific servers are configured). The use of multiple servers provides failover redundancy. However, no health checking is supported.

1. Use the following commands to configure global BOOTP relay servers:

```
RS G8124-E(config)# ip bootp-relay enable
RS G8124-E(config)# ip bootp-relay server <1-5> address <IPv4 address>
```

2. Enable BOOTP relay on the appropriate IP interfaces.

BOOTP/DHCP Relay functionality may be assigned on a per-interface basis using the following commands:

```
RS G8124-E(config)# interface ip <interface number>
RS G8124-E(config-ip-if)# relay
RS G8124-E(config-ip-if)# exit
```

Domain-Specific BOOTP Relay Agent Configuration

Use the following commands to configure up to five domain-specific BOOTP relay agents for each of up to 10 VLANs:

```
RS G8124-E(config)# ip bootp-relay bcast-domain <1-10> vlan <VLAN number>
RS G8124-E(config)# ip bootp-relay bcast-domain <1-10> server <1-5> address
     <IPv4 address>
RS G8124-E(config)# ip bootp-relay bcast-domain <1-10> enable
```

As with global relay agent servers, domain-specific BOOTP/DHCP functionality may be assigned on a per-interface basis (see).

DHCP Option 82

DHCP Option 82 provides a mechanism for generating IP addresses based on the client device's location in the network. When you enable the DHCP relay agent option on the switch, it inserts the relay agent information option 82 in the packet, and sends a unicast BOOTP request packet to the DHCP server. The DHCP server uses the option 82 field to assign an IP address, and sends the packet, with the original option 82 field included, back to the relay agent. DHCP relay agent strips off the option 82 field in the packet and sends the packet to the DHCP client.

Configuration of this feature is optional. The feature helps resolve several issues where untrusted hosts access the network. See RFC 3046 for details.

Given below are the commands to configure DHCP Option 82:

```
RS G8124-E(config)# ip bootp-relay information enable     (Enable Option 82)
RS G8124-E(config)# ip bootp-relay enable                (Enable DHCP
relay)
RS G8124-E(config)# ip bootp-relay server <1-5> address <IP address>
```

DHCP Snooping

DHCP snooping provides security by filtering untrusted DHCP packets and by building and maintaining a DHCP snooping binding table. This feature is applicable only to IPv4.

An untrusted interface is a port that is configured to receive packets from outside the network or firewall. A trusted interface receives packets only from within the network. By default, all DHCP ports are untrusted.

The DHCP snooping binding table contains the MAC address, IP address, lease time, binding type, VLAN number, and port number that correspond to the local untrusted interface on the switch; it does not contain information regarding hosts interconnected with a trusted interface.

By default, DHCP snooping is disabled on all VLANs. You can enable DHCP snooping on one or more VLANs. You must enable DHCP snooping globally. To enable this feature, enter the commands below:

```
RS G8124-E(config)# ip dhcp snooping vlan <vlan number(s)>
RS G8124-E(config)# ip dhcp snooping
```

Following is an example of DHCP snooping configuration, where the DHCP server and client are in VLAN 100, and the server connects using port 24.

```
RS G8124-E(config)# ip dhcp snooping vlan 100
RS G8124-E(config)# ip dhcp snooping
RS G8124-E(config)# interface port 24
RS G8124-E(config-if)# ip dhcp snooping trust(Optional; Set port as trusted)
RS G8124-E(config-if)# ip dhcp snooping information option-insert
                                    (Optional; add DHCP option 82)
RS G8124-E(config-if)# ip dhcp snooping limit rate 100
                                    (Optional; Set DHCP packet rate)
```

Easy Connect Wizard

Lenovo EasyConnect (EZC) is a feature designed to simplify switch configuration. A set of predefined configurations can be applied on the switch via ISCLI. By launching the EZC Wizard, you are prompted for a minimal set of input and the tool automatically customizes the switch software.

The EZC Wizard allows you to choose one of the following configuration modes:

- **Basic System** mode supports settings for hostname, static management port IP, netmask, and gateway.

- **Transparent** mode collects server and uplink port settings. vNIC groups are used to define the loop free domains.

   Note: You can either accept the static defaults or enter a different port list for uplink and/or server ports.

The EZC configuration will be applied immediately. Any existing configuration will be deleted, the current active or running configuration will not be merged or appended to the EZC configuration.

For any custom settings that are not included in the predefined configuration sets, the user has to do it manually.

Note: To support scripting, the feature also has a single-line format. For more information, please refer to *Lenovo Networking ISCLI Reference Guide*.

Configuring the Easy Connect Wizard

To launch the EZC Wizard, use the following command:

```
RS G8124-E# easyconnect
```

The wizard displays the available predefined configuration modes. You are prompted to select one of the following options:

```
RS G8124-E# easyconnect
Auto configures the switch into a set configuration based on the input
provided.
Current configuration will be overwritten with auto configuration
settings.
The wizard can be canceled anytime by pressing Ctrl+C.
Select which of the following features you want enabled:
#Configure Basic system (yes/no)?
#Configure Transparent mode (yes/no)?
```

Basic System Mode Configuration Example

This example shows the parameters available for configuration in Basic System mode:

```
RS G8124-E# easyconnect
Configure Basic system (yes/no)? y

Please enter "none" for no hostname.
Enter hostname(Default: None)? host

Please enter "dhcp" for dhcp IP.
Select management IP address (Current: 10.241.13.32)?
Enter management netmask(Current: 255.255.255.128)?
Enter management gateway:(Current: 10.241.13.1)?

Pending switch port configuration:

    Hostname: host
    Management interface:
        IP:      10.241.13.32
        Netmask: 255.255.255.128
        Gateway: 10.241.13.1
Confirm erasing current config to re-configure Easy Connect (yes/no)?
```

Note: You can either accept the default values or enter new parameters.

Transparent Mode Configuration Example

This example shows the parameters available for configuration in Transparent mode:

```
RS G8124-E# #easyconnect
Configure Transparent mode (yes/no)? y
Select Uplink Ports (Static Defaults:  1-4)?
The following Uplink ports will be enabled:
        Uplink ports(1G/10G):  1-4

Select Server Ports (Static Defaults: 5-24)?
The following Server ports will be enabled:
        Server ports(1G/10G):  5-24

Pending switch configuration:

    Uplink Ports:    1-4
    Server Ports:    5-24
    Disabled Ports:
Confirm erasing current config to re-configure Easy Connect (yes/no)?
```

Notes:

- If your selection for an uplink port group contains both 1G and 10G ports, the selection is not valid and you are guided to select other ports. Server ports can have both 1G and 10G selected at the same time.

- You can either accept the static defaults or enter a different port list for uplink and/or server ports.

Switch Login Levels

To enable better switch management and user accountability, three levels or *classes* of user access have been implemented on the G8124-E. Levels of access to CLI, Web management functions, and screens increase as needed to perform various switch management tasks. Conceptually, access classes are defined as follows:

- User interaction with the switch is completely passive—nothing can be changed on the G8124-E. Users may display information that has no security or privacy implications, such as switch statistics and current operational state information.

- Operators can only effect temporary changes on the G8124-E. These changes will be lost when the switch is rebooted/reset. Operators have access to the switch management features used for daily switch operations. Because any changes an operator makes are undone by a reset of the switch, operators cannot severely impact switch operation.

- Administrators are the only ones that may make permanent changes to the switch configuration—changes that are persistent across a reboot/reset of the switch. Administrators can access switch functions to configure and troubleshoot problems on the G8124-E. Because administrators can also make temporary (operator-level) changes as well, they must be aware of the interactions between temporary and permanent changes.

Access to switch functions is controlled through the use of unique user names and passwords. Once you are connected to the switch via console, remote Telnet, or SSH, you are prompted to enter a password. The default user names/password for each access level are listed in the following table.

Note: It is recommended that you change the default switch passwords after initial configuration and as regularly as required under your network security policies.

Table 2. *User Access Levels - Default Settings*

User Account	Password	Description and Tasks Performed	Status
user	user	The User has no direct responsibility for switch management. He or she can view all switch status information and statistics, but cannot make any configuration changes to the switch.	Disabled
oper	oper	The Operator manages all functions of the switch. The Operator can reset ports, except the management ports.	Disabled
admin	admin	The superuser Administrator has complete access to all menus, information, and configuration commands on the G8124-E, including the ability to change both the user and administrator passwords.	Enabled

Note: Access to each user level (except admin account) can be disabled by setting the password to an empty value. To disable admin account, use `no access user administrator-enable` command. Admin account can be disabled only if there is at least one user account enabled and configured with administrator privilege.

Setup vs. the Command Line

Once the administrator password is verified, you are given complete access to the switch. If the switch is still set to its factory default configuration, you will need to run Setup (see Chapter 2, "Initial Setup"), a utility designed to help you through the first-time configuration process. If the switch has already been configured, the command line is displayed instead.

Idle Disconnect

By default, the switch will disconnect your Telnet session after 10 minutes of inactivity. This function is controlled by the idle timeout parameter, which can be set from 0 to 60 minutes, where 0 means the session will never timeout.

Use the following command to set the idle timeout value:

```
RS G8124-E(config)# system idle <0-60>
```

Boot Strict Mode

The implementations specified in this section are compliant with National Institute of Standards and Technology (NIST) Special Publication (SP) 800-131A.

The RackSwitch G8124-E can operate in two boot modes:

- Compatibility mode (default): This is the default switch boot mode. This mode may use algorithms and key lengths that may not be allowed/acceptable by NIST SP 800-131A specification. This mode is useful in maintaining compatibility with previous releases and in environments that have lesser data security requirements.

- Strict mode: Encryption algorithms, protocols, and key lengths in strict mode are compliant with NIST SP 800-131A specification.

When in boot strict mode, the switch uses Secure Sockets Layer (SSL)/Transport Layer Security (TLS) 1.2 protocols to ensure confidentiality of the data to and from the switch.

Before enabling strict mode, ensure the following:

- The software version on all connected switches is Lenovo N/OS 8.3.

- The supported protocol versions and cryptographic cipher suites between clients and servers are compatible. For example: if using SSH to connect to the switch, ensure that the SSH client supports SSHv2 and a strong cipher suite that is compliant with the NIST standard.

- Compliant Web server certificate is installed on the switch, if using BBI.

- A new self-signed certificate is generated for the switch (RS G8124-E(config)# access https generate-certificate). The new certificate is generated using 2048-bit RSA key and SHA-256 digest.

- Protocols that are not NIST SP 800-131A compliant must be disabled or not used.

- Only SSHv2 or higher is used.

- The current configuration, if any, is saved in a location external to the switch. When the switch reboots, both the startup and running configuration are lost.

- Only protocols/algorithms compliant with NIST SP 800-131A specification are used/enabled on the switch. Please see the NIST SP 800-131A publication for details. The following table lists the acceptable protocols and algorithms:

Table 3. *Acceptable Protocols and Algorithms*

Protocol/Function	Strict Mode Algorithm	Compatibility Mode Algorithm
BGP	BGP does not comply with NIST SP 800-131A specification. When in strict mode, BGP is disabled. However, it can be enabled, if required.	Acceptable
Certificate Generation	RSA-2048 SHA-256	RSA 2048 SHA 256
Certificate Acceptance	RSA 2048 or higher SHA 224 or higher	RSA SHA, SHA2

Table 3. *Acceptable Protocols and Algorithms (continued)*

Protocol/Function	Strict Mode Algorithm	Compatibility Mode Algorithm
HTTPS	TLS 1.2 only See "Acceptable Cipher Suites" on page 50;	TLS 1.0, 1.1, 1.2 See "Acceptable Cipher Suites" on page 50;
IKE		
Key Exchange	DH Group 24	DH group 1, 2, 5, 14, 24
Encryption	3DES, AES-128-CBC	3DES, AES-128-CBC
Integrity	HMAC-SHA1	HMAC-SHA1, HMAC-MD5
IPSec		
AH	HMAC-SHA1	HMAC-SHA1, HMAC-MD5
ESP	3DES, AES-128-CBC, HMAC-SHA1	3DES, AES-128-CBC, HMAC-SHA1, HMAC-MD5
LDAP	LDAP does not comply with NIST SP 800-131A specification. When in strict mode, LDAP is disabled. However, it can be enabled, if required.	Acceptable
OSPF	OSPF does not comply with NIST SP 800-131A specification. When in strict mode, OSPF is disabled. However, it can be enabled, if required.	Acceptable
RADIUS	RADIUS does not comply with NIST SP 800-131A specification. When in strict mode, RADIUS is disabled. However, it can be enabled, if required.	Acceptable
Random Number Generator	NIST SP 800-90A AES CTR DRBG	NIST SP 800-90A AES CTR DRBG
Secure NTP	Secure NTP does not comply with NIST SP 800-131A specification. When in strict mode, secure NTP is disabled. However, it can be enabled, if required.	Acceptable
SLP	SHA-256 or higher RSA/DSA 2048 or higher	
SNMP	SNMPv3 only AES-128-CFB-128/SHA1 **Note:** Following algorithms are acceptable if you choose to support old SNMPv3 factory default users: AES-128-CFB/SHA1 DES/MD5 AES-128-CFB-128/SHA1	SNMPv1, SNMPv2, SNMPv3 DES/MD5, AES-128-CFB-128/SHA1

Table 3. *Acceptable Protocols and Algorithms (continued)*

Protocol/Function	Strict Mode Algorithm	Compatibility Mode Algorithm
SSH/SFTP		
Host Key	SSH-RSA	SSH-RSA
Key Exchange	ECDH-SHA2-NISTP521 ECDH-SHA2-NISTP384 ECDH-SHA2-NISTP256 ECDH-SHA2-NISTP224 RSA2048-SHA256 DIFFIE-HELLMAN-GROUP-EXCH ANGE-SHA256 DIFFIE-HELLMAN-GROUP-EXCH ANGE-SHA1	ECDH-SHA2-NISTP521 ECDH-SHA2-NISTP384 ECDH-SHA2-NISTP256 ECDH-SHA2-NISTP224 ECDH-SHA2-NISTP192 RSA2048-SHA256 RSA1024-SHA1 DIFFIE-HELLMAN-GROUP-EX CHANGE-SHA256 DIFFIE-HELLMAN-GROUP-EX CHANGE-SHA1 DIFFIE-HELLMAN-GROUP14-S HA1 DIFFIE-HELLMAN-GROUP1-S HA1
Encryption	AES128-CTR AES128-CBC 3DES-CBC	AES128-CTR AES128-CBC RIJNDAEL128-CBC BLOWFISH-CBC 3DES-CBC ARCFOUR256 ARCFOUR128 ARCFOUR
MAC	HMAC-SHA1 HMAC-SHA1-96	HMAC-SHA1 HMAC-SHA1-96 HMAC-MD5 HMAC-MD5-96
TACACS+	TACACS+ does not comply with NIST SP 800-131A specification. When in strict mode, TACACS+ is disabled. However, it can be enabled, if required.	Acceptable

Acceptable Cipher Suites

The following cipher suites are acceptable (listed in the order of preference) when the RackSwitch G8124-E is in compatibility mode:

Table 4. *List of Acceptable Cipher Suites in Compatibility Mode*

Cipher ID	Key Exchange	Authenti-cation	Encryption	MAC	Cipher Name
0xC027	ECDHE	RSA	AES_128_CBC	SHA256	TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA256
0xC013	ECDHE	RSA	AES_128_CBC	SHA1	TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA
0xC012	ECDHE	RSA	3DES	SHA1	SSL_ECDHE_RSA_WITH_3DES_EDE_CBC_SHA
0xC011	ECDHE	RSA	RC4	SHA1	SSL_ECDHE_RSA_WITH_RC4_128_SHA
0x002F	RSA	RSA	AES_128_CBC	SHA1	TLS_RSA_WITH_AES_128_CBC_SHA
0x003C	RSA	RSA	AES_128_CBC	SHA256	TLS_RSA_WITH_AES_128_CBC_SHA256
0x0005	RSA	RSA	RC4	SHA1	SSL_RSA_WITH_RC4_128_SHA
0x000A	RSA	RSA	3DES	SHA1	SSL_RSA_WITH_3DES_EDE_CBC_SHA
0x0033	DHE	RSA	AES-128_CBC	SHA1	TLS_DHE_RSA_WITH_AES_128_CBC_SHA
0x0067	DHE	RSA	AES_128_CBC	SHA256	TLS_DHE_RSA_WITH_AES_128_CBC_SHA256
0x0016	DHE	RSA	3DES	SHA1	SSL_DHE_RSA_WITH_3DES_EDE_CBC_SHA

The following cipher suites are acceptable (listed in the order of preference) when the RackSwitch G8124-E is in strict mode:

Table 5. *List of Acceptable Cipher Suites in Strict Mode*

Cipher ID	Key Exchange	Authenti-cation	Encryption	MAC	Cipher Name
0xC027	ECDHE	RSA	AES_128_CBC	SHA256	TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA256
0xC013	ECDHE	RSA	AES_128_CBC	SHA1	TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA
0xC012	ECDHE	RSA	3DES	SHA1	SSL_ECDHE_RSA_WITH_3DES_EDE_CBC_SHA
0x0033	DHE	RSA	AES-128_CBC	SHA1	TLS_DHE_RSA_WITH_AES_128_CBC_SHA
0x0067	DHE	RSA	AES_128_CBC	SHA256	TLS_DHE_RSA_WITH_AES_128_CBC_SHA256
0x0016	DHE	RSA	3DES	SHA1	SSL_DHE_RSA_WITH_3DES_EDE_CBC_SHA
0x002F	RSA	RSA	AES_128_CBC	SHA1	TLS_RSA_WITH_AES_128_CBC_SHA
0x003C	RSA	RSA	AES_128_CBC	SHA256	TLS_RSA_WITH_AES_128_CBC_SHA256
0x000A	RSA	RSA	3DES	SHA1	SSL_RSA_WITH_3DES_EDE_CBC_SHA

Configuring Strict Mode

To change the switch mode to boot strict mode, use the following command:

```
RS G8124-E(config)# [no] boot strict enable
```

When strict mode is enabled, you will see the following message:

```
Warning, security strict mode limits the cryptographic algorithms used by
secure protocols on this switch. Please see the documentation for full
details, and verify that peer devices support acceptable algorithms
before enabling this mode. The mode change will take effect after
reloading the switch and the configuration will be wiped during the
reload. System will enter security strict mode with default factory
configuration at next boot up.

Do you want SNMPV3 support old default users in strict mode (y/n)?
```

For SNMPv3 default users, see .

When strict mode is disabled, the following message is displayed:

```
Warning, disabling security strict mode. The mode change will take effect
after reloading the switch.
```

You must reboot the switch for the boot strict mode enable/disable to take effect.

Configuring No-Prompt Mode

If you expect to administer the switch using SNSC or another browser-based interface, you need to turn off confirmation prompts. To accomplish this, use the command:

```
RS G8124-E(config)# [no] terminal dont-ask
```

In no-prompt mode, confirmation prompts are disabled for this and future sessions.

SSL/TLS Version Limitation

Each of the following successive encryption protocol versions provide more security and less compatibility: SSLv3, TLS1.0, TLS1.1, TLS1.2. When negotiating the encryption protocol during the SSL handshake, the switch will accept, by default, the latest (and most secure) protocol version supported by the client equipment. To enforce a minimal level of security acceptable for the connections, use the following command:

```
RS G8124-E(config)# ssl minimum-version {ssl|tls10|tls11|tls12}
```

Limitations

In Lenovo N/OS 8.3, consider the following limitation/restrictions if you need to operate the switch in boot strict mode:

- Power ITEs and High-Availability features do not comply with NIST SP 800-131A specification.

- The G8124-E will not discover Platform agents/Common agents that are not in strict mode.

- Web browsers that do not use TLS 1.2 cannot be used.

- Limited functions of the switch managing Windows will be available.

Chapter 2. Initial Setup

To help with the initial process of configuring your switch, the Lenovo Networking OS software includes a Setup utility. The Setup utility prompts you step-by-step to enter all the necessary information for basic configuration of the switch.

Setup can be activated manually from the command line interface any time after login.

Information Needed for Setup

Setup requests the following information:

- Basic system information
  - Date & time
  - Whether to use Spanning Tree Group or not
- Optional configuration for each port
  - Speed, duplex, flow control, and negotiation mode (as appropriate)
  - Whether to use VLAN trunk mode/tagging or not (as appropriate)
- Optional configuration for each VLAN
  - Name of VLAN
  - Which ports are included in the VLAN
- Optional configuration of IP parameters
  - IP address/mask and VLAN for each IP interface
  - IP addresses for default gateway
  - Whether IP forwarding is enabled or not

Default Setup Options

You need to run the Setup utility to change the factory default settings. To accomplish this:

1. Connect to the switch.

   After connecting, the login prompt appears.

   ```
   Enter Password:
   ```

2. Enter **admin** as the default administrator password.

3. Start the Setup utility:

   ```
   RS G8124-E# setup
   ```

   Follow the instructions provided by the Setup utility.

Setting the Management Interface Default IP Address

To facilitate switch boot up, the in-band and out-of-band management interfaces are configured with factory default IP addresses. These are as follows:

- VLAN 1/ Interface 1: 192.168.49.50/24

- Out-of-band Management Port A: 192.168.50.50/24

- Out-of-band Management Port B: 192.168.51.50/24

If you configure static IP addresses or if DHCP/BOOTP addresses are assigned to these interfaces, the factory default IP addresses will not be applied. By default, DHCP and BOOTP are enabled on the management interfaces.

If you add interface 1 to another VLAN and do not configure any IP address, the factory default IP address will be automatically assigned to the interface.

We recommend that you disable the factory default IP address configuration after the switch boot up and configuration is complete. Use the following command:

```
RS G8124-E(config)# no system default-ip [data|mgtamgtb]
```

Stopping and Restarting Setup Manually

Stopping Setup

To abort the Setup utility, press <Ctrl-C> during any Setup question. When you abort Setup, the system will prompt:

```
Would you like to run from top again? [y/n]
```

Enter **n** to abort Setup, or **y** to restart the Setup program at the beginning.

Restarting Setup

You can restart the Setup utility manually at any time by entering the following command at the administrator prompt:

```
RS G8124-E(config)# setup
```

Setup Part 1: Basic System Configuration

When Setup is started, the system prompts:

```
"Set Up" will walk you through the configuration of
  System Date and Time, Spanning Tree, Port Speed/Mode,
  VLANs, and IP interfaces. [type Ctrl-C to abort "Set Up"]
```

1. Enter **y** if you will be configuring VLANs. Otherwise enter **n**.

   If you decide not to configure VLANs during this session, you can configure them later using the configuration menus, or by restarting the Setup facility. For more information on configuring VLANs, see the Lenovo Networking OS *Application Guide*.

   Next, the Setup utility prompts you to input basic system information.

2. Enter the year of the current date at the prompt:

```
System Date:
Enter year [2009]:
```

   Enter the four-digits that represent the year. To keep the current year, press <Enter>.

3. Enter the month of the current system date at the prompt:

```
System Date:
Enter month [1]:
```

   Enter the month as a number from 1 to 12. To keep the current month, press <Enter>.

4. Enter the day of the current date at the prompt:

```
Enter day [3]:
```

   Enter the date as a number from 1 to 31. To keep the current day, press <Enter>.

   The system displays the date and time settings:

```
System clock set to 18:55:36 Wed Jan 28, 2009.
```

5. Enter the hour of the current system time at the prompt:

```
System Time:
Enter hour in 24-hour format [18]:
```

   Enter the hour as a number from 00 to 23. To keep the current hour, press <Enter>.

6. Enter the minute of the current time at the prompt:

```
Enter minutes [55]:
```

Enter the minute as a number from 00 to 59. To keep the current minute, press <Enter>.

7. Enter the seconds of the current time at the prompt:

```
Enter seconds [37]:
```

Enter the seconds as a number from 00 to 59. To keep the current second, press <Enter>. The system then displays the date and time settings:

```
System clock set to 8:55:36 Wed Jan 28, 2009.
```

8. Turn Spanning Tree Protocol on or off at the prompt:

```
Spanning Tree:
Current Spanning Tree Group 1 setting: ON
Turn Spanning Tree Group 1 OFF? [y/n]
```

Enter **y** to turn off Spanning Tree, or enter **n** to leave Spanning Tree on.

Setup Part 2: Port Configuration

> **Note:** When configuring port options for your switch, some prompts and options may be different.

1. Select whether you will configure VLANs and VLAN trunk mode/tagging for ports:

```
Port Config:
Will you configure VLANs and VLAN Tagging/Trunk-Mode for ports? [y/n]
```

If you wish to change settings for VLANs, enter y, or enter n to skip VLAN configuration.

> **Note:** The sample screens that appear in this document might differ slightly from the screens displayed by your system. Screen content varies based on the firmware versions and options that are installed.

2. Select the port to configure, or skip port configuration at the prompt:

If you wish to change settings for individual ports, enter the number of the port you wish to configure. To skip port configuration, press <Enter> without specifying any port and go to .

3. Configure Gigabit Ethernet port flow parameters.

The system prompts:

```
Gig Link Configuration:
Port Flow Control:
Current Port EXT1 flow control setting:     both
Enter new value ["rx"/"tx"/"both"/"none"]:
```

Enter **rx** to enable receive flow control, **tx** for transmit flow control, **both** to enable both, or **none** to turn flow control off for the port. To keep the current setting, press <Enter>.

4. Configure Gigabit Ethernet port autonegotiation mode.

If you selected a port that has a Gigabit Ethernet connector, the system prompts:

```
Port Auto Negotiation:
Current Port autonegotiation:        on
Enter new value ["on"/"off"]:
```

Enter **on** to enable port autonegotiation, **off** to disable it, or press <Enter> to keep the current setting.

5. If configuring VLANs, enable or disable VLAN trunk mode/tagging for the port.

If you have selected to configure VLANs back in Part 1, the system prompts:

```
Port VLAN tagging/trunk mode config (tagged/trunk mode port can be a member
of multiple VLANs)
Current VLAN tagging/trunk mode support:            disabled
Enter new VLAN tagging/trunk mode support [d/e]:
```

Enter **d** to disable VLAN trunk mode/tagging for the port or enter **e** to enable VLAN tagging for the port. To keep the current setting, press <Enter>.

6. The system prompts you to configure the next port:

```
Enter port (1-24, MGTA, MGTB):
```

When you are through configuring ports, press <Enter> without specifying any port. Otherwise, repeat the steps in this section.

Setup Part 3: VLANs

If you chose to skip VLANs configuration back in Part 2, skip to .

1. Select the VLAN to configure, or skip VLAN configuration at the prompt:

```
VLAN Config:
Enter VLAN number from 2 to 4094, NULL at end:
```

If you wish to change settings for individual VLANs, enter the number of the VLAN you wish to configure. To skip VLAN configuration, press <Enter> without typing a VLAN number and go to .

2. Enter the new VLAN name at the prompt:

```
Current VLAN name: VLAN 2
Enter new VLAN name:
```

Entering a new VLAN name is optional. To use the pending new VLAN name, press <Enter>.

3. Enter the VLAN port numbers:

```
Define Ports in VLAN:
Current VLAN 2:  empty
Enter ports one per line, NULL at end:
```

Enter each port, by port number or port alias, and confirm placement of the port into this VLAN. When you are finished adding ports to this VLAN, press <Enter> without specifying any port.

4. Configure Spanning Tree Group membership for the VLAN:

```
Spanning Tree Group membership:
Enter new Spanning Tree Group index [1-127]:
```

5. The system prompts you to configure the next VLAN:

```
VLAN Config:
Enter VLAN number from 2 to 4094, NULL at end:
```

Repeat the steps in this section until all VLANs have been configured. When all VLANs have been configured, press <Enter> without specifying any VLAN.

Setup Part 4: IP Configuration

The system prompts for IPv4 parameters.

Although the switch supports both IPv4 and IPv6 networks, the Setup utility permits only IPv4 configuration. For IPv6 configuration, see Chapter 21, "Internet Protocol Version 6."

IP Interfaces

IP interfaces are used for defining the networks to which the switch belongs.

Up to 128 IP interfaces can be configured on the RackSwitch G8124-E (G8124-E). The IP address assigned to each IP interface provides the switch with an IP presence on your network. No two IP interfaces can be on the same IP network. The interfaces can be used for connecting to the switch for remote configuration, and for routing between subnets and VLANs (if used).

Note: Two interfaces are reserved for IPv4 management: interface 127 (MGTA) and 128 (MGTB). If the IPv6 feature is enabled, two additional interfaces are reserved to IPv6 management: interface 125 (MGTA) and 126 (MGTB).

1. Select the IP interface to configure, or skip interface configuration at the prompt:

```
IP Config:

IP interfaces:
Enter interface number: (1-128)
```

If you wish to configure individual IP interfaces, enter the number of the IP interface you wish to configure. To skip IP interface configuration, press <Enter> without typing an interface number and go to "Default Gateways" on page 64.

2. For the specified IP interface, enter the IP address in IPv4 dotted decimal notation:

```
Current IP address:     0.0.0.0
Enter new IP address:
```

To keep the current setting, press <Enter>.

3. At the prompt, enter the IPv4 subnet mask in dotted decimal notation:

```
Current subnet mask:         0.0.0.0
Enter new subnet mask:
```

To keep the current setting, press <Enter>.If configuring VLANs, specify a VLAN for the interface.

This prompt appears if you selected to configure VLANs back in Part 1:

```
Current VLAN:     1
Enter new VLAN [1-4094]:
```

Enter the number for the VLAN to which the interface belongs, or press <Enter> without specifying a VLAN number to accept the current setting.

4. At the prompt, enter **y** to enable the IP interface, or **n** to leave it disabled:

```
Enable IP interface? [y/n]
```

5. The system prompts you to configure another interface:

```
Enter interface number: (1-128)
```

Repeat the steps in this section until all IP interfaces have been configured. When all interfaces have been configured, press <Enter> without specifying any interface number.

Default Gateways

To set up a default gateway:

1. At the prompt, select an IP default gateway for configuration, or skip default gateway configuration:

```
IP default gateways:
Enter default gateway number: (1-4)
```

Enter the number for the IP default gateway to be configured. To skip default gateway configuration, press <Enter> without typing a gateway number and go to .

2. At the prompt, enter the IPv4 address for the selected default gateway:

```
Current IP address:     0.0.0.0
Enter new IP address:
```

Enter the IPv4 address in dotted decimal notation, or press <Enter> without specifying an address to accept the current setting.

3. At the prompt, enter **y** to enable the default gateway, or **n** to leave it disabled:

```
Enable default gateway? [y/n]
```

4. The system prompts you to configure another default gateway:

```
Enter default gateway number: (1-4)
```

Repeat the steps in this section until all default gateways have been configured. When all default gateways have been configured, press <Enter> without specifying any number.

IP Routing

When IP interfaces are configured for the various IP subnets attached to your switch, IP routing between them can be performed entirely within the switch. This eliminates the need to send inter-subnet communication to an external router device. Routing on more complex networks, where subnets may not have a direct presence on the G8124-E, can be accomplished through configuring static routes or by letting the switch learn routes dynamically.

This part of the Setup program prompts you to configure the various routing parameters.

At the prompt, enable or disable forwarding for IP Routing:

```
Enable IP forwarding? [y/n]
```

Enter **y** to enable IP forwarding. To disable IP forwarding, enter **n**. To keep the current setting, press <Enter>.

Setup Part 5: Final Steps

1. When prompted, decide whether to restart Setup or continue:

   ```
   Would you like to run from top again? [y/n]
   ```

   Enter **y** to restart the Setup utility from the beginning, or **n** to continue.

2. When prompted, decide whether you wish to review the configuration changes:

   ```
   Review the changes made? [y/n]
   ```

   Enter **y** to review the changes made during this session of the Setup utility. Enter **n** to continue without reviewing the changes. We recommend that you review the changes.

3. Next, decide whether to apply the changes at the prompt:

   ```
   Apply the changes? [y/n]
   ```

   Enter **y** to apply the changes, or **n** to continue without applying. Changes are normally applied.

4. At the prompt, decide whether to make the changes permanent:

   ```
   Save changes to flash? [y/n]
   ```

   Enter **y** to save the changes to flash. Enter **n** to continue without saving the changes. Changes are normally saved at this point.

5. If you do not apply or save the changes, the system prompts whether to abort them:

   ```
   Abort all changes? [y/n]
   ```

   Enter **y** to discard the changes. Enter **n** to return to the "Apply the changes?" prompt.

   Note: After initial configuration is complete, it is recommended that you change the default passwords.

Optional Setup for Telnet Support

Note: This step is optional. Perform this procedure only if you are planning on connecting to the G8124-E through a remote Telnet connection.

Telnet is enabled by default. To change the setting, use the following command:

```
RS G8124-E(config)# no access telnet
```

Loopback Interfaces

A loopback interface provides an IP address, but is not otherwise associated with a physical port or network entity. Essentially, it is a virtual interface that is perceived as being "always available" for higher-layer protocols to use and advertise to the network, regardless of other connectivity.

Loopback interfaces improve switch access, increase reliability, security, and provide greater flexibility in Layer 3 network designs. They can be used for many different purposes, but are most commonly for management IP addresses, router IDs for various protocols, and persistent peer IDs for neighbor relationships.

In Lenovo N/OS 8.3, loopback interfaces have been expanded for use with routing protocols such as OSPF, PIM, and BGP. Loopback interfaces can also be specified as the source IP address for syslog, SNMP, RADIUS, TACACS+, NTP, and router IDs.

Loopback interfaces must be configured before they can be used in other features. Up to five loopback interfaces are currently supported. They can be configured using the following commands:

```
RS G8124-E(config)# interface loopback <1-5>
RS G8124-E(config-ip-loopback)# [no] ip address <IPv4 address> <mask> enable
RS G8124-E(config-ip-loopback)# exit
```

Using Loopback Interfaces for Source IP Addresses

The switch can use loopback interfaces to set the source IP addresses for a variety of protocols. This assists in server security, as the server for each protocol can be configured to accept protocol packets only from the expected loopback address block. It may also make is easier to locate or process protocol information, since packets have the source IP address of the loopback interface, rather than numerous egress interfaces.

Configured loopback interfaces can be applied to the following protocols:

- Syslogs

```
RS G8124-E(config)# logging source-interface loopback <1-5>
```

- SNMP traps

```
RS G8124-E(config)# snmp-server trap-source loopback <1-5>
```

- RADIUS

```
RS G8124-E(config)# ip radius source-interface loopback <1-5>
```

- TACACS+

```
RS G8124-E(config)# ip tacacs source-interface loopback <1-5>
```

- NTP

```
RS G8124-E(config)# ntp source loopback <1-5>
```

Loopback Interface Limitation

- ARP is not supported. Loopback interfaces will ignore ARP requests.
- Loopback interfaces cannot be assigned to a VLAN.

Chapter 3. Switch Software Management

The switch software image is the executable code running on the G8124-E. A version of the image comes pre-installed on the device. As new versions of the image are released, you can upgrade the software running on your switch. To get the latest version of software supported for your G8124-E, go to the following website:

http://www.lenovo.com/support/

To determine the software version currently used on the switch, use the following switch command:

```
RS G8124-E# show boot
```

The typical upgrade process for the software image consists of the following steps:

- Load a new software image and boot image onto an FTP, SFTP or TFTP server on your network.

- Transfer the new images to your switch.

- Specify the new software image as the one which will be loaded into switch memory the next time a switch reset occurs.

- Reset the switch.

For instructions on the typical upgrade process using the N/OS ISCLI, or BBI, see "Loading New Software to Your Switch" on page 72.

CAUTION:
Although the typical upgrade process is all that is necessary in most cases, upgrading from (or reverting to) some versions of Lenovo Networking OS requires special steps prior to or after the software installation process. Please be sure to follow all applicable instructions in the release notes document for the specific software release to ensure that your switch continues to operate as expected after installing new software.

Loading New Software to Your Switch

The G8124-E can store up to two different switch software images (called `image1` and `image2`) as well as special boot software (called `boot`). When you load new software, you must specify where it is placed: either into `image1`, `image2`, or `boot`.

For example, if your active image is currently loaded into `image1`, you would probably load the new image software into `image2`. This lets you test the new software and reload the original active image (stored in `image1`), if needed.

![warning icon]

CAUTION:
When you upgrade the switch software image, always load the new boot image and the new software image before you reset the switch. If you do not load a new boot image, your switch might not boot properly (To recover, see "Recovering from a Failed Software Upgrade" on page 74**).**

To load a new software image to your switch, you will need the following:

● The image and boot software loaded on an FTP, SFTP or TFTP server on your network.

   Note: Be sure to download both the new boot file and the new image file.

● The hostname or IP address of the FTP, SFTP or TFTP server

   Note: The DNS parameters must be configured if specifying hostnames.

● The name of the new software image or boot file

When the software requirements are met, use one of the following procedures to download the new software to your switch. You can use the ISCLI, or the BBI to download and activate new software.

Loading Software via the ISCLI

1. In Privileged EXEC mode, enter the following command:

```
Router# copy {tftp|ftp|sftp} {image1|image2|boot-image}
```

2. Enter the hostname or IP address of the FTP, SFTP or TFTP server.

```
Address or name of remote host: <name or IP address>
```

3. Enter the name of the new software file on the server.

```
Source file name: <filename>
```

The exact form of the name will vary by server. However, the file location is normally relative to the FTP, SFTP or TFTP directory (for example, `tftpboot`).

4. If required by the FTP, SFTP or TFTP server, enter the appropriate username and password.

5. The switch will prompt you to confirm your request.

   Once confirmed, the software will begin loading into the switch.

6. When loading is complete, use the following commands to enter Global Configuration mode to select which software image (`image1` or `image2`) you want to run in switch memory for the next reboot:

```
Router# configure terminal
Router(config)# boot image {image1|image2}
```

   The system will then verify which image is set to be loaded at the next reset:

```
Next boot will use switch software image1 instead of image2.
```

7. Reboot the switch to run the new software:

```
Router(config)# reload
```

   The system prompts you to confirm your request. Once confirmed, the switch will reboot to use the new software.

Loading Software via BBI

You can use the Browser-Based Interface to load software onto the G8124-E. The software image to load can reside in one of the following locations:

- FTP server

- TFTP server

- SFTP server

- Local computer

After you log onto the BBI, perform the following steps to load a software image:

1. Click the Configure context tab in the toolbar.

2. In the Navigation Window, select System > Config/Image Control.

   The Switch Image and Configuration Management page appears.

3. If you are loading software from your computer (HTTP client), skip this step and go to the next. Otherwise, if you are loading software from an FTP, SFTP, or TFTP server, enter the server's information in the FTP, SFTP, or TFTP Settings section.

4. In the Image Settings section, select the image version you want to replace (Image for Transfer).

- If you are loading software from an FTP, SFTP, or TFTP server, enter the file name and click **Get Image**.

- If you are loading software from your computer, click **Browse**.

- In the File Upload Dialog, select the file and click **OK**. Then click **Download via Browser**.

Once the image has loaded, the page refreshes to show the new software.

The Boot Management Menu

The Boot Management menu allows you to switch the software image, reset the switch to factory defaults, or to recover from a failed software download.

You can interrupt the boot process and enter the Boot Management menu from the serial console port. When the system displays Memory Test, press **<Shift B>**. The Boot Management menu appears.

```
Resetting the System ...
Memory Test ...............................

Boot Management Menu
I - Change booting image
C - Change configuration block
Q - Reboot
E - Exit

Please choose your menu option: I
Current boot image is 1. Enter image to boot: 1 or 2: 2
Booting from image 2
```

The Boot Management menu allows you to perform the following actions:

- To change the booting image, press **I** and follow the screen prompts.

- To change the configuration block, press **C**, and follow the screen prompts.

- To reboot the switch, press **Q**. The booting process restarts.

- To exit the Boot Management menu, press **E**. The booting process continues.

Recovering from a Failed Software Upgrade

Use the following procedure to recover from a failed software upgrade.

1. Connect a PC to the serial port of the switch.

2. Open a terminal emulator program that supports XModem Download (for example, HyperTerminal, CRT, PuTTY) and select the following serial port characteristics:

    - Speed: 9600 bps

    - Data Bits: 8

    - Stop Bits: 1

    - Parity: None

    - Flow Control: None

3. To access the Boot Management menu you must interrupt the boot process from the Console port. Boot the G8124-E, and when the system begins displaying Memory Test progress (a series of dots), press **<Shift B>**. The boot managment menu appears:

4. Select **R** for Boot in recovery mode. The following appears:

```
Entering Rescue Mode.
Please select one of the following options:
        T) Configure networking and tftp download an image
        X) Use xmodem 1K to serial download an image
        R) Reboot
        E) Exit
```

- If you choose option **X** (Xmodem serial download), go to Step 5.
- If you choose option **T** (TFTP download), go to Step 6.

5. **Xmodem download:** When you see the following message, change the Serial Port characteristics to 115200 bps:

```
Change the baud rate to 115200 bps and hit the <ENTER> key before
initiating the download.
```

a. Press <**Enter>** to set the system into download accept mode. When the readiness meter displays (a series of "C" characters), start XModem on your terminal emulator.

b. When you see the following message, change the Serial Port characteristics to 9600 bps:

```
Change the baud rate back to 9600 bps, hit the <ESC> key.
```

c. When you see the following prompt, enter the image number where you want to install the new software and press <**Enter**>:

```
Install image as image 1 or 2 (hit return to just boot image): 1
```

d. The following message is displayed when the image download is complete. Continue to step 7.

```
Entering Rescue Mode.
Please select one of the following options:
        T) Configure networking and tftp download an image
        X) Use xmodem 1K to serial download an image
        R) Reboot
        E) Exit

Option?:
```

6. **TFTP download:** The switch prompts you to enter the following information:

```
Performing TFTP rescue. Please answer the following questions (enter 'q'
to quit):
IP addr    :
Server addr:
Netmask    :
Gateway    :
Image Filename:
```

a. Enter the required information and press <**Enter**>.

b. You will see a display similar to the following:

```
         Host IP : 10.10.98.110
         Server IP : 10.10.98.100
         Netmask : 255.255.255.0
         Broadcast : 10.10.98.255
         Gateway : 10.10.98.254
Installing image G8124-E-8.3.1.0_OS.img from TFTP server 10.10.98.100
```

c. When you see the following prompt, enter the image number and press <**Enter**>:

```
Install image as image 1 or 2 (hit return to just boot image): 1
```

d. The following message is displayed when the image download is complete.
   Continue to step 7.

```
Installing image as image1...
Image1 updated successfully
Please select one of the following options:
        T) Configure networking and tftp download an image
        X) Use xmodem 1K to serial download an image
        R) Reboot
        E) Exit
```

7. Image recovery is complete. Perform one of the following steps:

- Press r to reboot the switch.

- Press e to exit the Boot Management menu

- Press the Escape key (<**Esc**>) to re-display the Boot Management menu.

Recovering from a Failed Boot Image

Use the following procedure to recover from a failed boot image upgrade.

1. Connect a PC to the serial port of the switch.

2. Open a terminal emulator program that supports Xmodem download (for
   example, HyperTerminal, CRT, PuTTY) and select the following serial port
   characteristics:

- Speed: 9600 bps

- Data Bits: 8

- Stop Bits: 1

- Parity: None

- Flow Control: None

3. Boot the switch and access the Boot Management menu by pressing <**Shift B**> while the Memory Test is in progress and the dots are being displayed.

4. Select **X** for Xmodem download. The following appears:

```
Perform xmodem download

To download an image use 1K Xmodem at 115200 bps.
```

5. When you see the following message, change the Serial Port characteristics to 115200 bps:

```
Change the baud rate to 115200 bps and hit the <ENTER> key before
initiating the download.
```

a. Press <**Enter**> to set the system into download accept mode. When the readiness meter displays (a series of "C" characters), start Xmodem on your terminal emulator.You will see a display similar to the following:

```
Extracting images ... Do *NOT* power cycle the switch.
**** RAMDISK ****
Un-Protected 38 sectors
Erasing Flash...
................................... done
Erased 38 sectors
Writing to Flash...9....8....7....6....5....4....3....2....1....done
Protected 38 sectors
**** KERNEL ****
Un-Protected 24 sectors
Erasing Flash...
....................... done
Erased 24 sectors
Writing to Flash...9....8....7....6....5....4....3....2....1....
```

b. When you see the following message, change the Serial Port characteristics to 9600 bps:

```
Change the baud rate back to 9600 bps, hit the <ESC> key.
```

Boot image recovery is complete.

Part 2: Securing the Switch

Chapter 4. Securing Administration

Secure switch management is needed for environments that perform significant management functions across the Internet. Common functions for secured management are described in the following sections:

- "Secure Shell and Secure Copy" on page 82
- "End User Access Control" on page 87

Note: SNMP read and write functions are enabled by default. For best security practices, if SNMP is not needed for your network, it is recommended that you disable these functions prior to connecting the switch to the network (see Chapter 33, "Simple Network Management Protocol).

Secure Shell and Secure Copy

Because using Telnet does not provide a secure connection for managing a G8124-E, Secure Shell (SSH) and Secure Copy (SCP) features have been included for G8124-E management. SSH and SCP use secure tunnels to encrypt and secure messages between a remote administrator and the switch.

SSH is a protocol that enables remote administrators to log securely into the G8124-E over a network to execute management commands.

SCP is typically used to copy files securely from one machine to another. SCP uses SSH for encryption of data on the network. On a G8124-E, SCP is used to download and upload the switch configuration via secure channels.

Although SSH and SCP are disabled by default, enabling and using these features provides the following benefits:

- Identifying the administrator using Name/Password

- Authentication of remote administrators

- Authorization of remote administrators

- Determining the permitted actions and customizing service for individual administrators

- Encryption of management messages

- Encrypting messages between the remote administrator and switch

- Secure copy support

Lenovo Networking OS implements the SSH version 2.0 standard and is confirmed to work with SSH version 2.0-compliant clients such as the following:

- OpenSSH_5.4p1 for Linux

- Secure CRT Version 5.0.2 (build 1021)

- Putty SSH release 0.60

Configuring SSH/SCP Features on the Switch

SSH and SCP features are disabled by default. To change the SSH/SCP settings, using the following procedures.

Note: To use SCP, you must first enable SSH.

To Enable or Disable the SSH Feature

Begin a Telnet session from the console port and enter the following command:

```
RS G8124-E(config)# [no] ssh enable
```

To Enable or Disable SCP Apply and Save

Enter the following command from the switch CLI to enable the SCP
`putcfg_apply` and `putcfg_apply_save` commands:

```
RS G8124-E(config)# [no] ssh scp-enable
```

Configuring the SCP Administrator Password

To configure the SCP-only administrator password, enter the following command
(the default password is `admin`):

```
RS G8124-E(config)# [no] ssh scp-password
Changing SCP-only Administrator password; validation required...
Enter current administrator password: <password>
Enter new SCP-only administrator password: <new password>
Re-enter new SCP-only administrator password: <new password>
New SCP-only administrator password accepted.
```

Using SSH and SCP Client Commands

This section shows the format for using some client commands. The following
examples use 205.178.15.157 as the IP address of a sample switch.

To Log In to the Switch

Syntax:

```
>> ssh [-4|-6] <switch IP address>
   -or-
>> ssh [-4|-6] <login name>@<switch IP address>
```

Note: The -4 option (the default) specifies that an IPv4 switch address will be
used. The -6 option specifies IPv6.

Example**:**

```
>> ssh scpadmin@205.178.15.157
```

To Copy the Switch Configuration File to the SCP Host

Syntax**:**

```
>> scp [-4|-6] <username>@<switch IP address>:getcfg <local filename>
```

Example:

```
>> scp scpadmin@205.178.15.157:getcfg ad4.cfg
```

To Load a Switch Configuration File from the SCP Host

Syntax:

```
>> scp [-4|-6] <local filename> <username>@<switch IP address>:putcfg
```

Example:

```
>> scp ad4.cfg scpadmin@205.178.15.157:putcfg
```

To Apply and Save the Configuration

When loading a configuration file to the switch, the apply and save commands are still required for the configuration commands to take effect. The apply and save commands may be entered manually on the switch, or by using SCP commands.

Syntax:

```
>> scp [-4|-6] <local filename> <username>@<switch IP address>:putcfg_apply
>> scp [-4|-6] <local filename> <username>@<switch IP address>:putcfg_apply_save
```

Example:

```
>> scp ad4.cfg scpadmin@205.178.15.157:putcfg_apply
>> scp ad4.cfg scpadmin@205.178.15.157:putcfg_apply_save
```

- The CLI diff command is automatically executed at the end of putcfg to notify the remote client of the difference between the new and the current configurations.

- putcfg_apply runs the apply command after the putcfg is done.

- putcfg_apply_save saves the new configuration to the flash after putcfg_apply is done.

- The putcfg_apply and putcfg_apply_save commands are provided because extra apply and save commands are usually required after a putcfg; however, an SCP session is not in an interactive mode.

To Copy the Switch Image and Boot Files to the SCP Host

Syntax:

```
>> scp [-4|-6] <username>@<switch IP address>:getimg1 <local filename>
>> scp [-4|-6] <username>@<switch IP address>:getimg2 <local filename>
>> scp [-4|-6] <username>@<switch IP address>:getboot <local filename>
```

Example:

```
>> scp scpadmin@205.178.15.157:getimg1 6.1.0_os.img
```

To Load Switch Configuration Files from the SCP Host

Syntax:

```
>> scp [-4|-6] <local filename> <username>@<switch IP address>:putimg1
>> scp [-4|-6] <local filename> <username>@<switch IP address>:putimg2
>> scp [-4|-6] <local filename> <username>@<switch IP address>:putboot
```

Example:

```
>> scp 6.1.0_os.img scpadmin@205.178.15.157:putimg1
```

SSH and SCP Encryption of Management Messages

The following encryption and authentication methods are supported for SSH and SCP:

- Server Host Authentication: Client RSA authenticates the switch at the beginning of every connection

- Key Exchange: RSA

- Encryption: 3DES-CBC, DES

- User Authentication: Local password authentication, RADIUS

Generating RSA Host Key for SSH Access

To support the SSH host feature, an RSA host key is required. The host key is 2048 bits and is used to identify the G8124-E.

To configure RSA host key, first connect to the G8124-E through the console port (commands are not available via external Telnet connection), and enter the following command to generate it manually.

```
RS G8124-E(config)# ssh generate-host-key
```

When the switch reboots, it will retrieve the host key from the FLASH memory.

Note: The switch will perform only one session of key/cipher generation at a time. Thus, an SSH/SCP client will not be able to log in if the switch is performing key generation at that time. Also, key generation will fail if an SSH/SCP client is logging in at that time.

SSH/SCP Integration with Radius Authentication

SSH/SCP is integrated with RADIUS authentication. After the RADIUS server is enabled on the switch, all subsequent SSH authentication requests will be redirected to the specified RADIUS servers for authentication. The redirection is transparent to the SSH clients.

SSH/SCP Integration with TACACS+ Authentication

SSH/SCP is integrated with TACACS+ authentication. After the TACACS+ server is enabled on the switch, all subsequent SSH authentication requests will be redirected to the specified TACACS+ servers for authentication. The redirection is transparent to the SSH clients.

End User Access Control

Lenovo N/OS allows an administrator to define end user accounts that permit end users to perform operation tasks via the switch CLI commands. Once end user accounts are configured and enabled, the switch requires username/password authentication.

For example, an administrator can assign a user, who can then log into the switch and perform operational commands (effective only until the next switch reboot).

Considerations for Configuring End User Accounts

Note the following considerations when you configure end user accounts:

- A maximum of 20 user IDs are supported on the switch.
- N/OS supports end user support for console, Telnet, BBI, and SSHv2 access to the switch.
- If RADIUS authentication is used, the user password on the Radius server will override the user password on the G8124-E. Also note that the password change command only modifies only the user password on the switch and has no effect on the user password on the Radius server. Radius authentication and user password cannot be used concurrently to access the switch.
- Passwords for end users can be up to 128 characters in length for TACACS, RADIUS, Telnet, SSH, Console, and Web access.

Strong Passwords

The administrator can require use of Strong Passwords for users to access the G8124-E. Strong Passwords enhance security because they make password guessing more difficult.

The following rules apply when Strong Passwords are enabled:

- Minimum length: 8 characters; maximum length: 64 characters
- Must contain at least one uppercase alphabet
- Must contain at least one lowercase alphabet
- Must contain at least one number
- Must contain at least one special character:

  Supported special characters: ! " # % & ' ( ) ; < = >> ? [\ ] * + , - . / : ^ _ { | } ~

- Cannot be same as the username
- No consecutive four characters can be the same as in the old password

When strong password is enabled, users can still access the switch using the old password but will be advised to change to a strong password at log-in.

Strong password requirement can be enabled using the following command:

```
RS G8124-E(config)# access user strong-password enable
```

The administrator can choose the number of days allowed before each password expires. When a strong password expires, the user is allowed to log in one last time (last time) to change the password. A warning provides advance notice for users to change the password.

User Access Control

The end-user access control commands allow you to configure end-user accounts.

Setting up User IDs

Up to 20 user IDs can be configured. Use the following commands to define any user name and set the user password at the resulting prompts:

```
RS G8124-E(config)# access user 1 name <1-64 characters>
RS G8124-E(config)# access user 1 password

Changing user1 password; validation required:
Enter current admin password: <current administrator password>
Enter new user1 password: <new user password>
Re-enter new user1 password: <new user password>
New user1 password accepted.
```

Defining a User's Access Level

The end user is by default assigned to the user access level (also known as class of service, or COS). COS for all user accounts have global access to all resources except for User COS, which has access to view only resources that the user owns. For more information, see Table 8 on page 97.

To change the user's level, select one of the following options:

```
RS G8124-E(config)# access user 1 level {user|operator|administrator}
```

Validating a User's Configuration

```
RS G8124-E# show access user uid 1
```

Enabling or Disabling a User

An end user account must be enabled before the switch recognizes and permits login under the account. Once enabled, the switch requires any user to enter both username and password.

```
RS G8124-E(config)# [no] access user 1 enable
```

Locking Accounts

To protect the switch from unauthorized access, the account lockout feature can be enabled. By default, account lockout is disabled. To enable this feature, ensure the strong password feature is enabled (*See* "Strong Passwords" on page 87). Then use the following command:

```
RS G8124-E(config)# access user strong-password lockout
```

After multiple failed login attempts, the switch locks the user account if lockout has been enabled on the switch.

Re-enabling Locked Accounts

The administrator can re-enable a locked account by reloading the switch or by using the following command:

```
RS G8124-E(config)# access user strong-password clear local user lockout
            username <user name>
```

However, the above command cannot be used to re-enable an account disabled by the administrator.

To re-enable all locked accounts, use the following command:

```
RS G8124-E(config)# access user strong-password clear local user lockout
all
```

Listing Current Users

The following command displays defined user accounts and whether or not each user is currently logged into the switch.

```
RS G8124-E# show access user

Usernames:
  user    - Enabled - offline
  oper    - Disabled - offline
  admin   - Always Enabled - online 1 session

Current User ID table:
1: name jane   , ena, cos user    , password valid, online 1 session
2: name john   , ena, cos user    , password valid, online 2 sessions
```

Logging into an End User Account

Once an end user account is configured and enabled, the user can login to the switch using the username/password combination. The level of switch access is determined by the COS established for the end user account.

Password Fix-Up Mode

Password Fix-Up Mode enables admin user account recovery if administrator access is lost. A user must connect to the switch over the serial console and log in using the "ForgetMe!" password. This enables the admin account if disabled and a new administrator password can be entered.

To disable the Password Fix-Up functionality, use the following command:

```
RS G8124-E(config)# no access user password-recovery
```

Chapter 5. Authentication & Authorization Protocols

Secure switch management is needed for environments that perform significant management functions across the Internet. The following are some of the functions for secured IPv4 management and device access:

- "RADIUS Authentication and Authorization" on page 92
- "TACACS+ Authentication" on page 96
- "LDAP Authentication and Authorization" on page 100

Note: Lenovo Networking OS 8.3 does not support IPv6 for RADIUS, TACACS+ or LDAP.

RADIUS Authentication and Authorization

Lenovo N/OS supports the RADIUS (Remote Authentication Dial-in User Service) method to authenticate and authorize remote administrators for managing the switch. This method is based on a client/server model. The Remote Access Server (RAS)—the switch—is a client to the back-end database server. A remote user (the remote administrator) interacts only with the RAS, not the back-end server and database.

RADIUS authentication consists of the following components:

- A protocol with a frame format that utilizes UDP over IP (based on RFC 2138 and 2866)
- A centralized server that stores all the user authorization information
- A client: in this case, the switch

The G8124-E—acting as the RADIUS client—communicates to the RADIUS server to authenticate and authorize a remote administrator using the protocol definitions specified in RFC 2138 and 2866. Transactions between the client and the RADIUS server are authenticated using a shared key that is not sent over the network. In addition, the remote administrator passwords are sent encrypted between the RADIUS client (the switch) and the back-end RADIUS server.

How RADIUS Authentication Works

The RADIUS authentication process follows these steps:

1. A remote administrator connects to the switch and provides a user name and password.

2. Using Authentication/Authorization protocol, the switch sends request to authentication server.

3. The authentication server checks the request against the user ID database.

4. Using RADIUS protocol, the authentication server instructs the switch to grant or deny administrative access.

Configuring RADIUS on the Switch

Use the following procedure to configure Radius authentication on your switch.

1. Configure the IPv4 addresses of the Primary and Secondary RADIUS servers, and enable RADIUS authentication.

```
RS G8124-E(config)# radius-server primary-host 10.10.1.1
RS G8124-E(config)# radius-server secondary-host 10.10.1.2
RS G8124-E(config)# radius-server enable
```

Note: You can use a configured loopback address as the source address so the RADIUS server accepts requests only from the expected loopback address block. Use the following command to specify the loopback interface:

```
RS G8124-E(config)# ip radius source-interface loopback <1-5>
```

2. Configure the RADIUS secret.

```
RS G8124-E(config)# radius-server primary-host 10.10.1.1 key
    <1-32 character secret>
RS G8124-E(config)# radius-server secondary-host 10.10.1.2 key
    <1-32 character secret>
```

3. If desired, you may change the default UDP port number used to listen to RADIUS.

   The well-known port for RADIUS is 1812.

```
RS G8124-E(config)# radius-server port <UDP port number>
```

4. Configure the number retry attempts for contacting the RADIUS server, and the timeout period.

```
RS G8124-E(config)# radius-server retransmit 3
RS G8124-E(config)# radius-server timeout 5
```

RADIUS Authentication Features in Lenovo N/OS

N/OS supports the following RADIUS authentication features:

● Supports RADIUS client on the switch, based on the protocol definitions in RFC 2138 and RFC 2866.

● Allows RADIUS secret password up to 32 bytes and less than 16 octets.

● Supports *secondary authentication server* so that when the primary authentication server is unreachable, the switch can send client authentication requests to the secondary authentication server. Use the following command to show the currently active RADIUS authentication server:

```
RS G8124-E# show radius-server
```

● Supports user-configurable RADIUS server retry and time-out values:
  o Time-out value = 1-10 seconds
  o Retries = 1-3

  The switch will time out if it does not receive a response from the RADIUS server in 1-3 retries. The switch will also automatically retry connecting to the RADIUS server before it declares the server down.

● Supports user-configurable RADIUS application port. The default is UDP port 1645. UDP port 1812, based on RFC 2138, is also supported.

● Allows network administrator to define privileges for one or more specific users to access the switch at the RADIUS user database.

Switch User Accounts

The user accounts listed in Table 6 can be defined in the RADIUS server dictionary file.

Table 6. *User Access Levels*

User Account	Description and Tasks Performed	Password
User	The User has no direct responsibility for switch management. They can view all switch status information and statistics but cannot make any configuration changes to the switch.	user
Operator	The Operator manages all functions of the switch. The Operator can reset ports, except the management port.	oper
Administrator	The super-user Administrator has complete access to all commands, information, and configuration commands on the switch, including the ability to change both the user and administrator passwords.	admin

RADIUS Attributes for Lenovo N/OS User Privileges

When the user logs in, the switch authenticates his/her level of access by sending the RADIUS access request, that is, the client authentication request, to the RADIUS authentication server.

If the remote user is successfully authenticated by the authentication server, the switch will verify the *privileges* of the remote user and authorize the appropriate access. The administrator has two options: to allow *backdoor* access via Telnet, SSH, HTTP, or HTTPS; to allow *secure backdoor* access via Telnet, SSH, or BBI. Backdoor and secure backdoor provides access to the switch when the RADIUS servers cannot be reached.

The default G8124-E setting for backdoor and secure backdoor access is `disabled`. Backdoor and secure backdoor access is always enabled on the console port.

Irrespective of backdoor/secure backdoor being enabled or not, you can always access the switch via the console port by using noradius as radius username. You can then enter the username and password configured on the switch. If you are trying to connect via SSH/Telnet/HTTP/HTTPS (not console port), there are two possibilities:

● Backdoor is enabled: The switch acts like it is connecting via console.

● Secure backdoor is enabled: You must enter the username: noradius. The switch checks if RADIUS server is reachable. If it is reachable, then you must authenticate via remote authentication server. Only if RADIUS server is not reachable, you will be prompted for local user/password to be authenticated against these local credentials.

All user privileges, other than those assigned to the Administrator, have to be defined in the RADIUS dictionary. RADIUS attribute 6 which is built into all RADIUS servers defines the administrator. The file name of the dictionary is RADIUS vendor-dependent. The following RADIUS attributes are defined for G8124-E user privileges levels:

Table 7. *Lenovo N/OS-proprietary Attributes for RADIUS*

User Name/Access	User-Service-Type	Value
User	*Vendor-supplied*	255
Operator	*Vendor-supplied*	252
Admin	*Vendor-supplied*	6

TACACS+ Authentication

N/OS supports authentication and authorization with networks using the Cisco Systems TACACS+ protocol. The G8124-E functions as the Network Access Server (NAS) by interacting with the remote client and initiating authentication and authorization sessions with the TACACS+ access server. The remote user is defined as someone requiring management access to the G8124-E either through a data port or management port.

TACACS+ offers the following advantages over RADIUS:

- TACACS+ uses TCP-based connection-oriented transport; whereas RADIUS is UDP-based. TCP offers a connection-oriented transport, while UDP offers best-effort delivery. RADIUS requires additional programmable variables such as re-transmit attempts and time-outs to compensate for best-effort transport, but it lacks the level of built-in support that a TCP transport offers.
- TACACS+ offers full packet encryption whereas RADIUS offers password-only encryption in authentication requests.
- TACACS+ separates authentication, authorization and accounting.

How TACACS+ Authentication Works

TACACS+ works much in the same way as RADIUS authentication as described on .

1. Remote administrator connects to the switch and provides user name and password.

2. Using Authentication/Authorization protocol, the switch sends request to authentication server.

3. Authentication server checks the request against the user ID database.

4. Using TACACS+ protocol, the authentication server instructs the switch to grant or deny administrative access.

   During a session, if additional authorization checking is needed, the switch checks with a TACACS+ server to determine if the user is granted permission to use a particular command.

TACACS+ Authentication Features in Lenovo N/OS

Authentication is the action of determining the identity of a user, and is generally done when the user first attempts to log in to a device or gain access to its services. N/OS supports ASCII inbound login to the device. PAP, CHAP and ARAP login methods, TACACS+ change password requests, and one-time password authentication are not supported.

Authorization

Authorization is the action of determining a user's privileges on the device, and usually takes place after authentication.

The default mapping between TACACS+ authorization levels and N/OS management access levels is shown in Table 8. The authorization levels must be defined on the TACACS+ server.

Table 8. *Default TACACS+ Authorization Levels*

N/OS User Access Level	TACACS+ level
user	0
oper	3
admin	6

Alternate mapping between TACACS+ authorization levels and N/OS management access levels is shown in Table 9. Use the following command to set the alternate TACACS+ authorization levels.

```
RS G8124-E(config)# tacacs-server privilege-mapping
```

Table 9. *Alternate TACACS+ Authorization Levels*

N/OS User Access Level	TACACS+ level
user	0 - 1
oper	6 - 8
admin	14 - 15

If the remote user is successfully authenticated by the authentication server, the switch verifies the *privileges* of the remote user and authorizes the appropriate access. The administrator has an option to allow *secure backdoor* access via Telnet/SSH. Secure backdoor provides switch access when the TACACS+ servers cannot be reached. You always can access the switch via the console port, by using notacacs and the administrator password, whether secure backdoor is enabled or not.

Note: To obtain the TACACS+ backdoor password for your G8124-E, contact Technical Support.

Accounting

Accounting is the action of recording a user's activities on the device for the purposes of billing and/or security. It follows the authentication and authorization actions. If the authentication and authorization is not performed via TACACS+, there are no TACACS+ accounting messages sent out.

You can use TACACS+ to record and track software login access, configuration changes, and interactive commands.

The G8124-E supports the following TACACS+ accounting attributes:

- protocol (console/Telnet/SSH/HTTP/HTTPS)

- start_time

- stop_time

- elapsed_time

- disc_cause

Note: When using the Browser-Based Interface, the TACACS+ Accounting Stop records are sent only if the **Logout** button on the browser is clicked.

Command Authorization and Logging

When TACACS+ Command Authorization is enabled, N/OS configuration commands are sent to the TACACS+ server for authorization. Use the following command to enable TACACS+ Command Authorization:

```
RS G8124-E(config)# tacacs-server command-authorization
```

When TACACS+ Command Logging is enabled, N/OS configuration commands are logged on the TACACS+ server. Use the following command to enable TACACS+ Command Logging:

```
RS G8124-E(config)# tacacs-server command-logging
```

The following examples illustrate the format of N/OS commands sent to the TACACS+ server:

```
authorization request, cmd=shell, cmd-arg=interface ip
accounting request, cmd=shell, cmd-arg=interface ip
authorization request, cmd=shell, cmd-arg=enable
accounting request, cmd=shell, cmd-arg=enable
```

Configuring TACACS+ Authentication on the Switch

1. Configure the IPv4 addresses of the Primary and Secondary TACACS+ servers, and enable TACACS authentication. Specify the interface port (optional).

```
RS G8124-E(config)# tacacs-server primary-host 10.10.1.1
RS G8124-E(config)# tacacs-server primary-host mgtb-port
RS G8124-E(config)# tacacs-server secondary-host 10.10.1.2
RS G8124-E(config)# tacacs-server secondary-host data-port
RS G8124-E(config)# tacacs-server enable
```

Note: You can use a configured loopback address as the source address so the TACACS+ server accepts requests only from the expected loopback address block. Use the following command to specify the loopback interface:
```
RS G8124-E(config)# ip tacacs source-interface loopback <1-5>
```

2. Configure the TACACS+ secret and second secret.

```
RS G8124-E(config)# tacacs-server primary-host 10.10.1.1 key
<1-32 character secret>
RS G8124-E(config)# tacacs-server secondary-host 10.10.1.2 key
      <1-32 character secret>
```

3. If desired, you may change the default TCP port number used to listen to TACACS+.

   The well-known port for TACACS+ is 49.

```
RS G8124-E(config)# tacacs-server port <TCP port number>
```

4. Configure the number of retry attempts, and the timeout period.

```
RS G8124-E(config)# tacacs-server retransmit 3
RS G8124-E(config)# tacacs-server timeout 5
```

LDAP Authentication and Authorization

N/OS supports the LDAP (Lightweight Directory Access Protocol) method to authenticate and authorize remote administrators to manage the switch. LDAP is based on a client/server model. The switch acts as a client to the LDAP server. A remote user (the remote administrator) interacts only with the switch, not the back-end server and database.

LDAP authentication consists of the following components:

- A protocol with a frame format that utilizes TCP over IP
- A centralized server that stores all the user authorization information
- A client: in this case, the switch

Each entry in the LDAP server is referenced by its Distinguished Name (DN). The DN consists of the user-account name concatenated with the LDAP domain name. If the user-account name is John, the following is an example DN:

```
uid=John,ou=people,dc=domain,dc=com
```

Configuring the LDAP Server

G8124-E user groups and user accounts must reside within the same domain. On the LDAP server, configure the domain to include G8124-E user groups and user accounts, as follows:

- User Accounts:

    Use the *uid* attribute to define each individual user account. If a custom attribute is used to define individual users, it must also be configured on the switch.

- User Groups:

    Use the *members* attribute in the *groupOfNames* object class to create the user groups. The first word of the common name for each user group must be equal to the user group names defined in the G8124-E, as follows:

    o `admin`

    o `oper`

    o `user`

Configuring LDAP Authentication on the Switch

1. Turn LDAP authentication on, then configure the IPv4 addresses of the Primary and Secondary LDAP servers. Specify the interface port (optional).

```
RS G8124-E(config)# ldap-server enable
RS G8124-E(config)# ldap-server primary-host 10.10.1.1 mgta-port
RS G8124-E(config)# ldap-server secondary-host 10.10.1.2 data-port
```

2. Configure the domain name.

```
RS G8124-E(config)# ldap-server domain <ou=people,dc=my-domain,dc=com>
```

3. You may change the default TCP port number used to listen to LDAP (optional).

The well-known port for LDAP is 389.

```
RS G8124-E(config)# ldap-server port <1-65000>
```

4. Configure the number of retry attempts for contacting the LDAP server, and the timeout period.

```
RS G8124-E(config)# ldap-server retransmit 3
RS G8124-E(config)# ldap-server timeout 10
```

5. You may change the default LDAP attribute (uid) or add a custom attribute. For instance, Microsoft's Active Directory requires the cn attribute.

```
RS G8124-E(config)# ldap-server attribute username <128 alpha-numeric characters>
```

Chapter 6. Access Control Lists

Access Control Lists (ACLs) are filters that permit or deny traffic for security purposes. They can also be used with QoS to classify and segment traffic to provide different levels of service to different traffic types. Each filter defines the conditions that must match for inclusion in the filter, and also the actions that are performed when a match is made.

Lenovo Networking OS 8.3 supports the following ACLs:

- IPv4 ACLs

  Up to 127 ACLs are supported for networks that use IPv4 addressing. IPv4 ACLs are configured using the following ISCLI command path:

  ```
  RS G8124-E(config)# access-control list <IPv4 ACL number> ?
  ```

- IPv6 ACLs

  Up to 128 ACLs are supported for networks that use IPv6 addressing. IPv6 ACLs are configured using the following ISCLI command path:

  ```
  RS G8124-E(config)# access-control list6 <IPv6 ACL number> ?
  ```

- VLAN Maps (VMaps)

  Up to 127 VLAN Maps are supported for attaching filters to VLANs rather than ports. See "VLAN Maps" on page 113 for details.

Note: The stated ACL capacity reflects the Default deployment mode. ACL support may differ under some deployments modes. In modes where ACLs are not supported, ACL configuration menus and commands are not available. See "Available Profiles" on page 206 for details.

Summary of Packet Classifiers

ACLs allow you to classify packets according to a variety of content in the packet header (such as the source address, destination address, source port number, destination port number, and others). Once classified, packet flows can be identified for more processing.

IPv4 ACLs, IPv6 ACLs, and VMaps allow you to classify packets based on the following packet attributes:

- Ethernet header options (for IPv4 ACLs and VMaps only)
  - Source MAC address
  - Destination MAC address
  - VLAN number and mask
  - Ethernet type (ARP, IP, IPv6, MPLS, RARP, etc.)
  - Ethernet Priority (the IEEE 802.1p Priority)
- IPv4 header options (for IPv4 ACLs and VMaps only)
  - Source IPv4 address and subnet mask
  - Destination IPv4 address and subnet mask
  - Type of Service value
  - IP protocol number or name as shown in Table 10:

Table 10. *Well-Known Protocol Types*

Number	Protocol Name
1	icmp
2	igmp
6	tcp
17	udp
89	ospf
112	vrrp

- IPv6 header options (for IPv6 ACLs only)
  - Source IPv6 address and prefix length
  - Destination IPv6 address and prefix length
  - Next Header value
  - Flow Label value
  - Traffic Class value

- TCP/UDP header options (for all ACLs)
  - TCP/UDP application source port and mask as shown in Table 11
  - TCP/UDP application destination port as shown in Table 11

Table 11. *Well-Known Application Ports*

Port	TCP/UDP Application	Port	TCP/UDP Application	Port	TCP/UDP Application
20	ftp-data	79	finger	179	bgp
21	ftp	80	http	194	irc
22	ssh	109	pop2	220	imap3
23	telnet	110	pop3	389	ldap
25	smtp	111	sunrpc	443	https
37	time	119	nntp	520	rip
42	name	123	ntp	554	rtsp
43	whois	143	imap	1645/1812	Radius
53	domain	144	news	1813	Radius
69	tftp	161	snmp	1985	Accounting
70	gopher	162	snmptrap		hsrp

  - TCP/UDP flag value as shown in Table 12

Table 12. *Well Known TCP Flag Values*

Flag	Value
URG	0x0020
ACK	0x0010
PSH	0x0008
RST	0x0004
SYN	0x0002
FIN	0x0001

- Packet format (for IPv4 ACLs and VMaps only)
  - Ethernet format (eth2, SNAP, LLC)
  - Ethernet tagging format
  - IP format (IPv4, IPv6)

Summary of ACL Actions

Once classified using ACLs, the identified packet flows can be processed differently. For each ACL, an *action* can be assigned. The action determines how the switch treats packets that match the classifiers assigned to the ACL. G8124-E ACL actions include the following:

- Pass or Drop the packet
- Re-mark the packet with a new DiffServ Code Point (DSCP)
- Re-mark the 802.1p field
- Set the COS queue
- Change ingress VLAN classification

Note: ACLs act only upon ingress traffic on a port, not egress traffic.

Assigning Individual ACLs to a Port

Once you configure an ACL, you must assign the ACL to the appropriate ports. Each port can accept multiple ACLs, and each ACL can be applied for multiple ports. ACLs can be assigned individually.

To assign an individual ACLs to a port, use the following IP Interface Mode commands:

```
RS G8124-E(config)# interface port <port>
RS G8124-E(config-if)# access-control list <IPv4 ACL number>
RS G8124-E(config-if)# access-control list6 <IPv6 ACL number>
```

When multiple ACLs are assigned to a port, higher-priority ACLs are considered first, and their action takes precedence over lower-priority ACLs. ACL order of precedence is discussed in the next section.

ACL Order of Precedence

When multiple ACLs are assigned to a port, they are evaluated in numeric sequence, based on the ACL number. Lower-numbered ACLs take precedence over higher-numbered ACLs. For example, ACL 1 (if assigned to the port) is evaluated first and has top priority.

If multiple ACLs match the port traffic, only the action of the one with the lowest ACL number is applied. The others are ignored.

If no assigned ACL matches the port traffic, no ACL action is applied.

ACL Metering and Re-Marking

You can define a profile for the aggregate traffic flowing through the G8124-E by configuring a QoS meter (if desired) and assigning ACLs to ports.

Note: When you add ACLs to a port, make sure they are ordered correctly in terms of precedence (see "ACL Order of Precedence" on page 106).

Actions taken by an ACL are called *In-Profile* actions. You can configure additional In-Profile and Out-of-Profile actions on a port. Data traffic can be metered, and re-marked to ensure that the traffic flow provides certain levels of service in terms of bandwidth for different types of network traffic.

Metering

QoS metering provides different levels of service to data streams through user-configurable parameters. A meter is used to measure the traffic stream against a traffic profile which you create. Thus, creating meters yields In-Profile and Out-of-Profile traffic for each ACL, as follows:

- **In-Profile**If there is no meter configured or if the packet conforms to the meter, the packet is classified as In-Profile.

- **Out-of-Profile**If a meter is configured and the packet does not conform to the meter (exceeds the committed rate or maximum burst rate of the meter), the packet is classified as Out-of-Profile.

Note: Metering is not supported for IPv6 ACLs. All traffic matching an IPv6 ACL is considered in-profile for re-marking purposes.

Using meters, you set a Committed Rate in Kbps (in multiples of 64 Mbps). All traffic within this Committed Rate is In-Profile. Additionally, you can set a Maximum Burst Size that specifies an allowed data burst larger than the Committed Rate for a brief period. These parameters define the In-Profile traffic.

Meters keep the sorted packets within certain parameters. You can configure a meter on an ACL, and perform actions on metered traffic, such as packet re-marking.

Re-Marking

Re-marking allows for the treatment of packets to be reset based on new network specifications or desired levels of service. You can configure the ACL to re-mark a packet as follows:

- Change the DSCP value of a packet, used to specify the service level that traffic receives.

- Change the 802.1p priority of a packet.

ACL Port Mirroring

For IPv4 ACLs and VMaps, packets that match the filter can be mirrored to another switch port for network diagnosis and monitoring.

The source port for the mirrored packets cannot be a portchannel, but may be a member of a portchannel.

The destination port to which packets are mirrored must be a physical port.

The action (permit, drop, etc.) of the ACL or VMap must be configured before assigning it to a port.

Use the following commands to add mirroring to an ACL:

● For IPv4 ACLs:

```
RS G8124-E(config)# access-control list <ACL number> mirror port
       <destination port>
```

The ACL must be also assigned to it target ports as usual (see "Assigning Individual ACLs to a Port" on page 106).

● For VMaps (see "VLAN Maps" on page 113):

```
RS G8124-E(config)# access-control vmap <VMap number> mirror port <monitor
       destination port>
```

See the configuration example on page 114.

Viewing ACL Statistics

ACL statistics display how many packets have "hit" (matched) each ACL. Use ACL statistics to check filter performance or to debug the ACL filter configuration.

You must enable statistics for each ACL that you wish to monitor:

```
RS G8124-E(config)# access-control list <ACL number> statistics
```

ACL Logging

ACLs are generally used to enhance port security. Traffic that matches the characteristics (source addresses, destination addresses, packet type, etc.) specified by the ACLs on specific ports is subject to the actions (chiefly permit or deny) defined by those ACLs. Although switch statistics show the number of times particular ACLs are matched, the ACL logging feature can provide additional insight into actual traffic patterns on the switch, providing packet details in the system log for network debugging or security purposes.

Enabling ACL Logging

By default, ACL logging is disabled. Enable or disable ACL logging on a per-ACL basis as follows:

```
RS G8124-E(config)# [no] access-control list <IPv4 ACL number> log
RS G8124-E(config)# [no] access-control list6 <IPv6 ACL number> log
```

Logged Information

When ACL logging is enabled on any particular ACL, the switch will collect information about packets that match the ACL. The information collected depends on the ACL type:

- For IP-based ACLs, information is collected regarding
  - Source IP address
  - Destination IP address
  - TCP/UDP port number
  - ACL action
  - Number of packets logged

  For example:

  ```
  Sep 27  4:20:28 DUT3 NOTICE  ACL-LOG: %IP ACCESS LOG: list
  ACL-IP-12-IN denied tcp 1.1.1.1 (0) -> 200.0.1.2 (0), 150
  packets.
  ```

- For MAC-based ACLs, information is collected regarding
  - Source MAC address
  - Source IP address
  - Destination IP address
  - TCP/UDP port number
  - ACL action
  - Number of packets logged

  For example:

  ```
  Sep 27  4:25:38 DUT3 NOTICE  ACL-LOG: %MAC ACCESS LOG: list
  ACL-MAC-12-IN permitted tcp 1.1.1.2 (0) (12,
  00:ff:d7:66:74:62) -> 200.0.1.2 (0) (00:18:73:ee:a7:c6), 32
  packets.
  ```

Rate Limiting Behavior

Because ACL logging can be CPU-intensive, logging is rate-limited. By default, the switch will log only 10 matching packets per second. This pool is shared by all log-enabled ACLs. The global rate limit can be changed as follows:

```
RS G8124-E(config)# access-control log rate-limit <1-1000>
```

Where the limit is specified in packets per second.

Log Interval

For each log-enabled ACL, the first packet that matches the ACL initiates an immediate message in the system log. Beyond that, additional matches are subject to the log interval. By default, the switch will buffer ACL log messages for a period of 300 seconds. At the end of that interval, all messages in the buffer are written to the system log. The global interval value can be changed as follows:

```
RS G8124-E(config)# access-control log interval <5-600>
```

Where the interval rate is specified in seconds.

In any given interval, packets that have identical log information are condensed into a single message. However, the packet count shown in the ACL log message represents only the logged messages, which due to rate-limiting, may be significantly less than the number of packets actually matched by the ACL.

Also, the switch is limited to 64 different ACL log messages in any interval. Once the threshold is reached, the oldest message will be discarded in favor of the new message, and an overflow message will be added to the system log.

ACL Logging Limitations

- When ACL logging is enabled, the ACL action for setting packet priority levels (to a value other than 1) is not supported. For instance, the following configuration is invalid:

```
RS G8124-E(config)# access-control list 1 action set-priority 2
```

- ACL logging reserves packet queue 1 for internal use. Features that allow remapping packet queues (such as QoS CPU rate limit) may not behave as expected if other packet flows are reconfigured to use queue 1.

ACL Configuration Examples

ACL Example 1

Use this configuration to block traffic to a specific host. All traffic that ingresses on port 1 is denied if it is destined for the host at IP address 100.10.1.1

1. Configure an Access Control List.

```
RS G8124-E(config)# access-control list 1 ipv4 destination-ip-address
    100.10.1.1
RS G8124-E(config)# access-control list 1 action deny
```

2. Add ACL 1 to port 1.

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# access-control list 1
RS G8124-E(config-if)# exit
```

ACL Example 2

Use this configuration to block traffic from a network destined for a specific host address. All traffic that ingresses in port 2 with source IP from class 100.10.1.0/24 and destination IP 200.20.2.2 is denied.

1. Configure an Access Control List.

```
RS G8124-E(config)# access-control list 2 ipv4 source-ip-address
    100.10.1.0 255.255.255.0
RS G8124-E(config)# access-control list 2 ipv4 destination-ip-address
    200.20.2.2 255.255.255.255
RS G8124-E(config)# access-control list 2 action deny
```

2. Add ACL 2 to port 2.

```
RS G8124-E(config)# interface port 2
RS G8124-E(config-if)# access-control list 2
RS G8124-E(config-if)# exit
```

ACL Example 3

Use this configuration to block traffic from a specific IPv6 source address. All traffic that ingresses in port 2 with source IP from class 2001:0:0:5:0:0:0:2/128 is denied.

1. Configure an Access Control List.

```
RS G8124-E(config)# access-control list6 3 ipv6 source-address
    2001:0:0:5:0:0:0:2 128
RS G8124-E(config)# access-control list6 3 action deny
```

2. Add ACL 2 to port 2.

```
RS G8124-E(config)# interface port 2
RS G8124-E(config-if)# access-control list6 3
RS G8124-E(config-if)# exit
```

ACL Example 4

Use this configuration to deny all ARP packets that ingress a port.

1. Configure an Access Control List.

```
RS G8124-E(config)# access-control list 2 ethernet ethernet-type arp
RS G8124-E(config)# access-control list 2 action deny
```

2. Add ACL 2 to port EXT2.

```
RS G8124-E(config)# interface port 2
RS G8124-E(config-if)# access-control list 2
RS G8124-E(config-if)# exit
```

ACL Example 5

Use the following configuration to permit access to hosts with destination MAC address that matches 11:05:00:10:00:00 FF:F5:FF:FF:FF:FF and deny access to all other hosts.

1. Configure Access Control Lists.

```
RS G8124-E(config)# access-control list 30 ethernet
    destination-mac-address 11:05:00:10:00:00 FF:F5:FF:FF:FF:FF
RS G8124-E(config)# access-control list 30 action permit
RS G8124-E(config)# access-control list 100 ethernet
    destination-mac-address 00:00:00:00:00:00 00:00:00:00:00:00
RS G8124-E(config)# access-control list 100 action deny
```

2. Add ACLs to a port.

```
RS G8124-E(config)# interface port 2
RS G8124-E(config-if)# access-control list 30
RS G8124-E(config-if)# access-control list 100
RS G8124-E(config-if)# exit
```

VLAN Maps

A VLAN map (VMap) is an ACL that can be assigned to a VLAN or VM group rather than to a switch port as with IPv4 ACLs. This is particularly useful in a virtualized environment where traffic filtering and metering policies must follow virtual machines (VMs) as they migrate between hypervisors.

Note: VLAN maps for VM groups are not supported simultaneously on the same ports as vNICs (see Chapter 14, "Virtual NICs").

The G8124-E supports up to 127 VMaps when the switch is operating in the default deployment mode (see "Available Profiles" on page 206). VMap menus and commands are not available in the Routing deployment mode.

Individual VMap filters are configured in the same fashion as IPv4 ACLs, except that VLANs cannot be specified as a filtering criteria (unnecessary, since the VMap are assigned to a specific VLAN or associated with a VM group VLAN).

VMaps are configured using the following ISCLI configuration command path:

```
RS G8124-E(config)# access-control vmap <VMap ID> ?
  action         Set filter action
  ethernet       Ethernet header options
  ipv4           IP version 4 header options
  meter          ACL metering configuration
  mirror         Mirror options
  packet-format  Set to filter specific packet format types
  re-mark        ACL re-mark configuration
  statistics     Enable access control list statistics
  tcp-udp        TCP and UDP filtering options
```

Once a VMap filter is created, it can be assigned or removed using the following configuration commands:

● For regular VLAN, use config-vlan mode:

```
RS G8124-E(config)# vlan <VLAN ID>
RS G8124-E(config-vlan)# [no] vmap <VMap ID> [serverports|
    non-serverports]
```

● For a VM group (see "VM Group Types" on page 224), use the global configuration mode:

```
RS G8124-E(config)# [no] virt vmgroup <ID> vmap <VMap ID>
    [serverports|non-serverports]
```

Note: Each VMap can be assigned to only one VLAN or VM group. However, each VLAN or VM group may have multiple VMaps assigned to it.

When the optional serverports or non-serverports parameter is specified, the action to add or remove the VMap is applied for either the switch server ports (serverports) or uplink ports (non-serverports). If omitted, the operation will be applied to all ports in the associated VLAN or VM group.

In this example, EtherType 2 traffic from VLAN 3 server ports is mirrored to a network monitor on port 4.

```
RS G8124-E(config)# access-control vmap 21 packet-format ethernet
ethernet-type2
RS G8124-E(config)# access-control vmap 21 mirror port 4
RS G8124-E(config)# access-control vmap 21 action permit
RS G8124-E(config)# vlan 3
RS G8124-E(config-vlan)# vmap 21 serverports
```

VLAN Classification

Packets received on ports may belong to multiple VLANs or may be untagged. VLAN classification using ACLs allows you to assign a new VLAN to such packets that can be flooded to a monitor port. These packets, when sent out from the monitor port, can be identified by their source ports.

The ingress and monitor ports must belong to the new VLAN. Tagging must be enabled on these ports.

Note: When you assign a new VLAN to the ingress packets, the packets will be forwarded based on the new VLAN.

By default, VLAN classification is disabled.

Consider the following example in which packets from ingress port 1 are assigned to VLAN 100. Packets are flooded to monitor port 20.

1. Create a new VLAN. Add ingress and monitor ports to the VLAN:

```
RS G8124-E(config)# vlan 100

VLAN number 100 with name "VLAN 100" created.
RS G8124-E(config-vlan)# no shutdown
RS G8124-E(config-vlan)# exit
```

2. Create an ACL to assign the new VLAN to packets received on ingress port:

```
RS G8124-E(config)# access-control list 1 action change-vlan 100
```

Note: You can add filters, such as source IP address, to the ACL.

3. Enable tagging on ingress and monitor ports:

```
RS G8124-E(config)# interface port 1,20
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 100
RS G8124-E(config-if)# exit
```

4. Apply the ACL to the ingress port:

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# access-control list 1
```

Notes:

- When you apply VLAN classification, VMAP, and IPv4 ACL at the same time, IPv4 ACL takes precedence over VMAP.

- When you apply VLAN classification, VMAP, and IPv6 ACL at the same time, VMAP ACL takes precedence over IPv6 ACL.

Restrictions

VLAN classification is primarily useful for monitoring packets. When you enable this feature, the switch considers only the new VLAN. The switch does not learn the original VLAN. We recommend that you use this feature only if the switch does not need to process regular switching and bidirectional traffic.

Using Storm Control Filters

Excessive transmission of broadcast or multicast traffic can result in a network storm. A network storm can overwhelm your network with constant broadcast or multicast traffic, and degrade network performance. Common symptoms of a network storm are denial-of-service (DoS) attacks, slow network response times, and network operations timing out.

The G8124-E provides filters that can limit the number of the following packet types transmitted by switch ports:

- Broadcast packets
- Multicast packets
- Unknown unicast packets (destination lookup failure)

Unicast packets whose destination MAC address is not in the Forwarding Database are *unknown unicasts*. When an unknown unicast is encountered, the switch handles it like a broadcast packet and floods it to all other ports in the VLAN (broadcast domain). A high rate of unknown unicast traffic can have the same negative effects as a broadcast storm.

Configuring Storm Control

Configure broadcast filters on each port that requires broadcast storm control. Set a threshold that defines the total number of broadcast packets transmitted (100-10000), in packets per second. When the threshold is reached, no more packets of the specified type are transmitted.

Up to nine storm filters can be configured on the G8124-E.

To filter broadcast packets on a port, use the following commands:

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# storm-control broadcast level rate <packets per second>
```

To filter multicast packets on a port, use the following commands:

```
RS G8124-E(config-if)# storm-control multicast level rate <packets per second>
```

To filter unknown unicast packets on a port, use the following commands:

```
RS G8124-E(config-if)# storm-control unicast level rate <packets per second>
RS G8124-E(config-if)# exit
```

Part 3: Switch Basics

This section discusses basic switching functions:

- VLANs

- Port Aggregation

- Spanning Tree Protocols (Spanning Tree Groups, Rapid Spanning Tree Protocol, and Multiple Spanning Tree Protocol)

- Virtual Link Aggregation Groups

- Quality of Service

Chapter 7. VLANs

This chapter describes network design and topology considerations for using Virtual Local Area Networks (VLANs). VLANs commonly are used to split up groups of network users into manageable broadcast domains, to create logical segmentation of workgroups, and to enforce security policies among logical segments. The following topics are discussed in this chapter:

  This section discusses how you can connect users and segments to a host that supports many logical segments or subnets by using the flexibility of the multiple VLAN system.

Note: VLANs can be configured from the Command Line Interface (see "VLAN Configuration" as well as "Port Configuration" in the *Command Reference*).

VLANs Overview

Setting up virtual LANs (VLANs) is a way to segment networks to increase network flexibility without changing the physical network topology. With network segmentation, each switch port connects to a segment that is a single broadcast domain. When a switch port is configured to be a member of a VLAN, it is added to a group of ports (workgroup) that belong to one broadcast domain.

Ports are grouped into broadcast domains by assigning them to the same VLAN. Frames received in one VLAN can only be forwarded within that VLAN, and multicast, broadcast, and unknown unicast frames are flooded only to ports in the same VLAN.

The RackSwitch G8124-E (G8124-E) supports jumbo frames with a Maximum Transmission Unit (MTU) of 9,216 bytes. Within each frame, 18 bytes are reserved for the Ethernet header and CRC trailer. The remaining space in the frame (up to 9,198 bytes) comprise the packet, which includes the payload of up to 9,000 bytes and any additional overhead, such as 802.1q or VLAN tags. Jumbo frame support is automatic: it is enabled by default, requires no manual configuration, and cannot be manually disabled.

VLANs and Port VLAN ID Numbers

VLAN Numbers

The G8124-E supports up to 4095 VLANs per switch. Each can be identified with any number between 1 and 4094. VLAN 1 is the default VLAN for the data ports. VLAN 4095 is used by the management network, which includes the management ports.

Use the following command to view VLAN information:

```
RS G8124-E# show vlan

VLAN          Name             Status          Ports
----  ------------------------ ------  ------------------------
1     Default VLAN             ena
2     VLAN 2                   dis     empty
4095  Mgmt VLAN                ena     MGMTA MGMTB
```

PVID/Native VLAN Numbers

Each port in the switch has a configurable default VLAN number, known as its *PVID*. By default, the PVID for all non-management ports is set to 1, which correlates to the default VLAN ID. The PVID for each port can be configured to any VLAN number between 1 and 4094.

Use the following command to view PVIDs:

```
RS G8124-E# show interface information
(or)
RS G8124-E# show interface trunk

Alias   Port Tag RMON Lrn Fld PVID    DESCRIPTION          VLAN(s)
            Trk              NVLAN
------- ---- --- ---- --- --- ----- --------------
-------------------------------
1       1    n   d    e   e   1                   1
2       2    n   d    e   e   1                   1
3       3    n   d    e   e   1                   1
4       4    n   d    e   e   1                   1
...     ...                                       ...
24      24   n   d    e   e   24                  24
MGTA    25   n   d    e   e   4095               4095
MGTB    26   n   d    e   e   4095               4095

* = PVID/Native-VLAN is tagged.
Trk  = Trunk mode
NVLAN = Native-VLAN
```

Use the following command to set the port PVID/Native VLAN:

```
Access Mode Port

RS G8124-E(config)# interface port <port number>
RS G8124-E(config-if)# switchport access vlan <VLAN ID>

For Trunk Mode Port


RS G8124-E(config)# interface port <port number>
RS G8124-E(config-if)# switchport trunk native vlan <VLAN ID>
```

Each port on the switch can belong to one or more VLANs, and each VLAN can have any number of switch ports in its membership. Any port that belongs to multiple VLANs, however, must have VLAN *tagging*/trunk mode enabled (see).

VLAN Tagging/Trunk Mode

Lenovo Networking OS software supports 802.1Q VLAN *tagging,* providing standards-based VLAN support for Ethernet systems.

Tagging places the VLAN identifier in the frame header of a packet, allowing each port to belong to multiple VLANs. When you add a port to multiple VLANs, you also must enable tagging on that port.

Since tagging fundamentally changes the format of frames transmitted on a tagged port, you must carefully plan network designs to prevent tagged frames from being transmitted to devices that do not support 802.1Q VLAN tags, or devices where tagging is not enabled.

Important terms used with the 802.1Q tagging feature are:

- VLAN identifier (VID)—the 12-bit portion of the VLAN tag in the frame header that identifies an explicit VLAN.

- Port VLAN identifier (PVID)—a classification mechanism that associates a port with a specific VLAN. For example, a port with a PVID of 3 (PVID =3) assigns all untagged frames received on this port to VLAN 3. Any untagged frames received by the switch are classified with the PVID of the receiving port.

- Tagged frame—a frame that carries VLAN tagging information in the header. This VLAN tagging information is a 32-bit field (VLAN tag) in the frame header that identifies the frame as belonging to a specific VLAN. Untagged frames are marked (tagged) with this classification as they leave the switch through a port that is configured as a tagged port.

- Untagged frame— a frame that does not carry any VLAN tagging information in the frame header.

- Untagged member—a port that has been configured as an untagged member of a specific VLAN. When an untagged frame exits the switch through an untagged member port, the frame header remains unchanged. When a tagged frame exits the switch through an untagged member port, the tag is stripped and the tagged frame is changed to an untagged frame.

- Tagged member—a port that has been configured as a tagged member of a specific VLAN. When an untagged frame exits the switch through a tagged member port, the frame header is modified to include the 32-bit tag associated with the PVID. When a tagged frame exits the switch through a tagged member port, the frame header remains unchanged (original VID remains). When an access port is set as a trunk, it is automatically added to all data VLANs. To change the allowed VLAN range, use the command:

  `switchport trunk allowed vlans` *<range>*

Note: If a 802.1Q tagged frame is received by a port that has VLAN-tagging disabled and the port VLAN ID (PVID) is different than the VLAN ID of the packet, then the frame is dropped at the ingress port.

Figure 1. Default VLAN settings
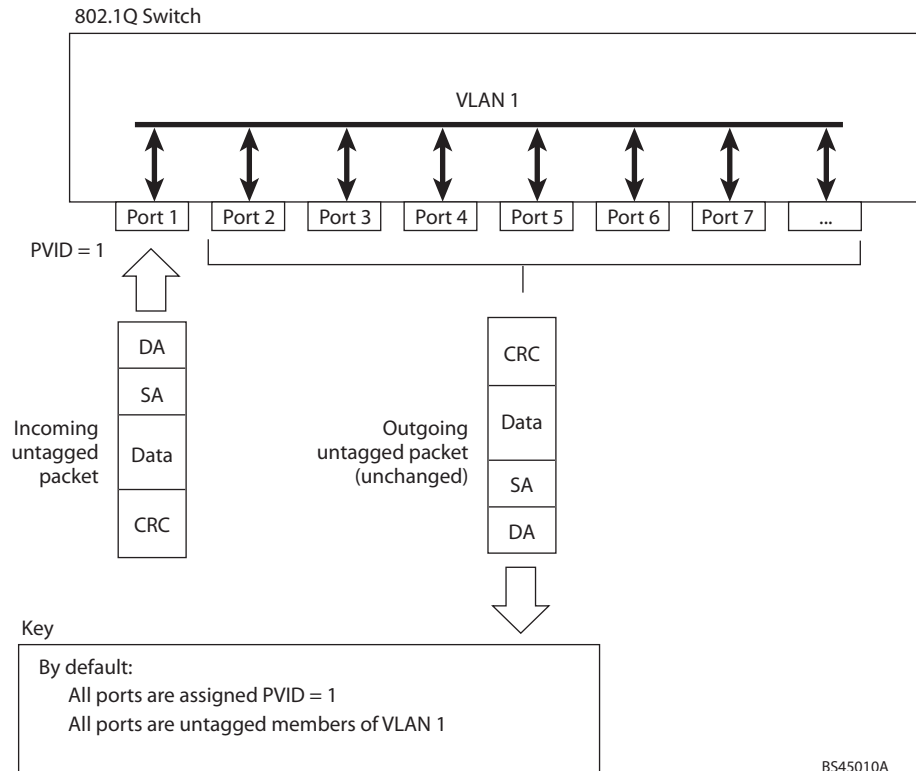
Note: The port numbers specified in these illustrations may not directly correspond to the physical port configuration of your switch model.

When a VLAN is configured, ports are added as members of the VLAN, and the ports are defined as either *tagged* or *untagged* (see Figure 2 through Figure 5).

The default configuration settings for the G8124-E has all ports set as untagged members of VLAN 1 with all ports configured as PVID = 1. In the default configuration example shown in Figure 1, all incoming packets are assigned to VLAN 1 by the default port VLAN identifier (PVID =1).

Figure 2 through Figure 5 illustrate generic examples of VLAN tagging. In Figure 2, untagged incoming packets are assigned directly to VLAN 2 (PVID = 2). Port 5 is configured as a *tagged* member of VLAN 2, and port 7 is configured as an *untagged* member of VLAN 2.

Note: The port assignments in the following figures are not meant to match the G8124-E.

Figure 2. Port-based VLAN assignment
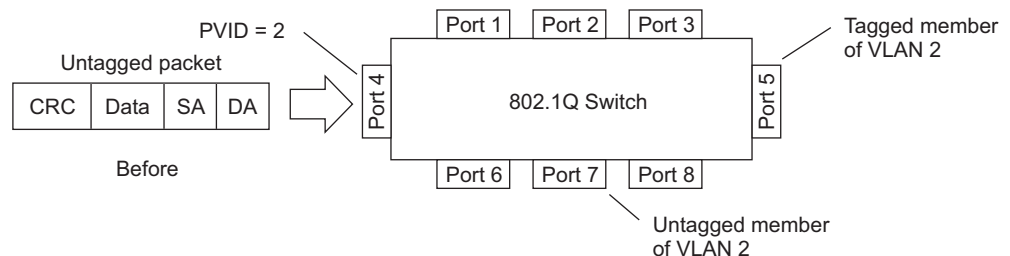
As shown in Figure 3, the untagged packet is marked (tagged) as it leaves the switch through port 5, which is configured as a tagged member of VLAN 2. The untagged packet remains unchanged as it leaves the switch through port 7, which is configured as an untagged member of VLAN 2.

Figure 3. 802.1Q tagging (after port-based VLAN assignment)
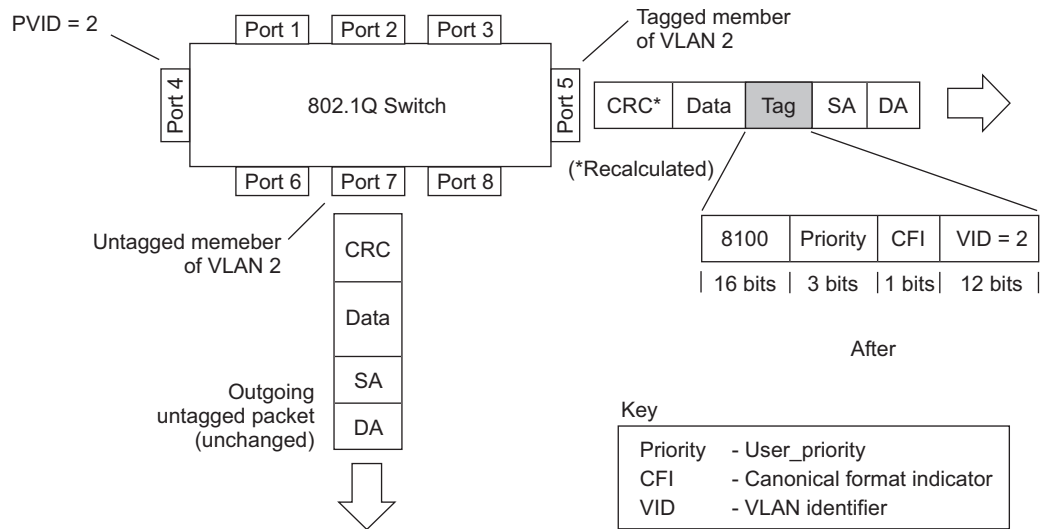
In Figure 4, tagged incoming packets are assigned directly to VLAN 2 because of the tag assignment in the packet. Port 5 is configured as a *tagged* member of VLAN 2, and port 7 is configured as an *untagged* member of VLAN 2.

Figure 4. 802.1Q tag assignment
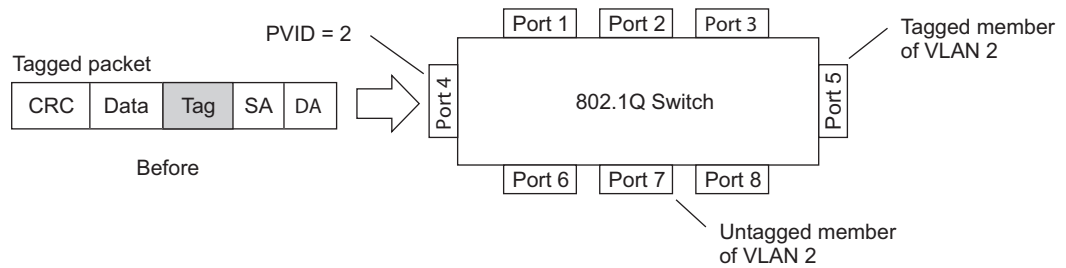
As shown in Figure 5, the tagged packet remains unchanged as it leaves the switch through port 5, which is configured as a tagged member of VLAN 2. However, the tagged packet is stripped (untagged) as it leaves the switch through port 7, which is configured as an untagged member of VLAN 2.

Figure 5. 802.1Q tagging (after 802.1Q tag assignment)
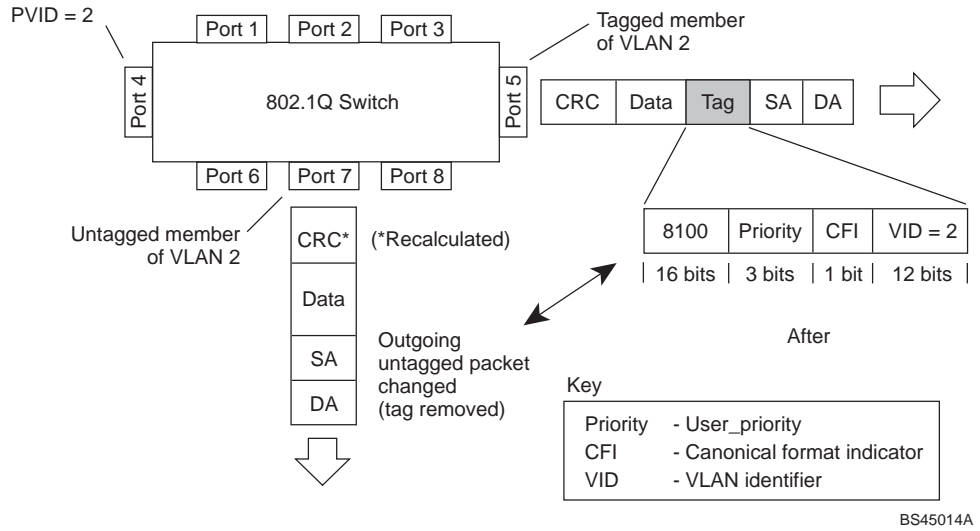
VLAN Topologies and Design Considerations

Note the following when working with VLAN topologies:

- By default, the G8124-E software is configured so that tagging/trunk mode is disabled on all ports.

- By default, the G8124-E software is configured so that all data ports are members of VLAN 1.

- By default, the Lenovo N/OS software is configured so that the management ports (MGTA and MGTB) are members of VLAN 4095 (the management VLAN).

- STG 128 is reserved for switch management.

- When using Spanning Tree, STG 2-128 may contain only one VLAN unless Multiple Spanning-Tree Protocol (MSTP) mode is used. With MSTP mode, STG 1 to 32 can include multiple VLANs.

- All ports involved in both aggregation and port mirroring must have the same VLAN configuration. If a port is on a LAG with a mirroring port, the VLAN configuration cannot be changed. For more information about aggregation, see Chapter 8, "Ports and Link Aggregation" and Chapter 37, "Port Mirroring."

Multiple VLANs with Tagging/Trunk Mode Adapters

illustrates a network topology described in and the configuration example on page .

Figure 6. Multiple VLANs with VLAN-Tagged Gigabit Adapters
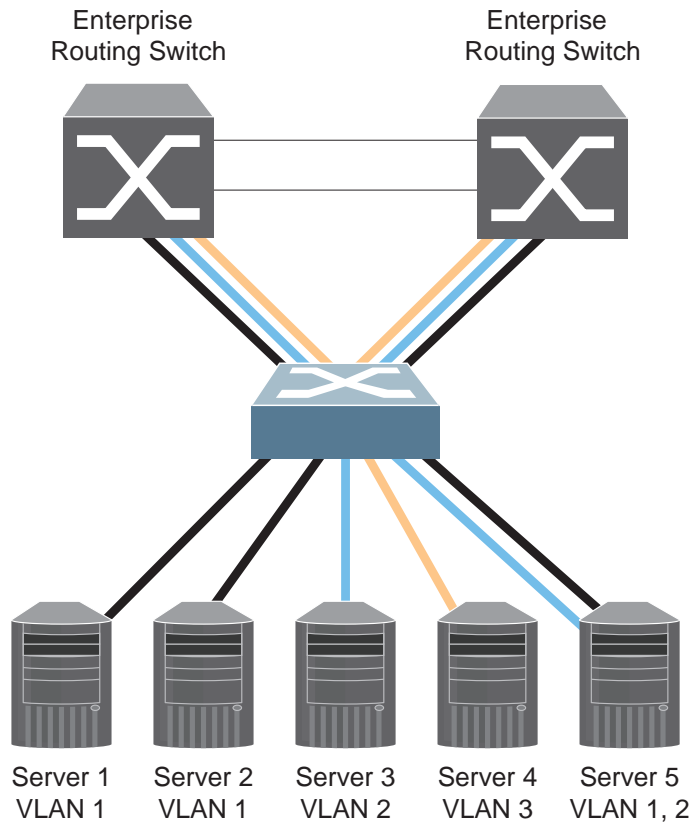
The features of this VLAN are described in the following table.

Table 13. *Multiple VLANs Example*

Component	Description
G8124-E switch	This switch is configured with three VLANs that represent three different IP subnets. Five ports are connected downstream to servers. Two ports are connected upstream to routing switches. Uplink ports are members of all three VLANs, with VLAN tagging/trunk mode enabled.
Server 1	This server is a member of VLAN 1 and has presence in only one IP subnet. The associated switch port is only a member of VLAN 1, so tagging/trunk mode is disabled.
Server 2	This server is a member of VLAN 1 and has presence in only one IP subnet. The associated switch port is only a member of VLAN 1, so tagging/trunk mode is disabled.
Server 3	This server belongs to VLAN 2, and it is logically in the same IP subnet as Server 5. The associated switch port has tagging/trunk mode disabled.
Server 4	A member of VLAN 3, this server can communicate only with other servers via a router. The associated switch port has tagging/trunk mode disabled.
Server 5	A member of VLAN 1 and VLAN 2, this server can communicate only with Server 1, Server 2, and Server 3. The associated switch port has tagging/trunk mode enabled.
Enterprise Routing switches	These switches must have all three VLANs (VLAN 1, 2, 3) configured. They can communicate with Server 1, Server 2, and Server 5 via VLAN 1. They can communicate with Server 3 and Server 5 via VLAN 2. They can communicate with Server 4 via VLAN 3. Tagging/trunk mode on switch ports is enabled.

Note: VLAN tagging/trunk mode is required only on ports that are connected to other switches or on ports that connect to tag-capable end-stations, such as servers with VLAN-tagging/trunk mode adapters.

To configure a specific VLAN on a trunk port, the following conditions must be met:

- The port must be in trunk mode.
- The VLAN must be in the trunk's allowed VLAN range. By default, the range includes all VLANs .
- The VLAN must be un-reserved.
- The VLAN must be created.

The order in which the conditions above are met is not relevant. However, all conditions must be met collectively. When all the conditions are met, the VLAN is enabled on the port. If one of the conditions is broken, the VLAN is disabled.

If a port's native VLAN is a private VLAN and its allowed VLAN range contains only invalid VLANs (either reserved VLANs or VLANs the port cannot belong to), removing the private VLAN mapping from the port will add the port to default VLAN and add the default VLAN to the allowed VLAN range.

When setting up multiple VLANs, ports configured in private VLAN mode are not added to private VLANs unless the private VLANs are also configured for those ports.

VLAN Configuration Example

Use the following procedure to configure the example network shown in Figure 6 on page 127.

1. Enable VLAN tagging/trunk mode on server ports that support multiple VLANs.

```
RS G8124-E(config)# interface port 5
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlans 1,2
RS G8124-E(config-if)# exit
```

2. Enable tagging/trunk mode on uplink ports that support multiple VLANs.

```
RS G8124-E(config)# interface port 19
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 20
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# exit
```

3. Configure server ports that belong to a single VLAN.

```
RS G8124-E(config)# interface port 4
RS G8124-E(config-if)# switchport access vlan 2
RS G8124-E(config-if)# exit
```

By default, all ports are members of VLAN 1, so configure only those ports that belong to other VLANs.

Private VLANs

Private VLANs provide Layer 2 isolation between the ports within the same broadcast domain. Private VLANs can control traffic within a VLAN domain, and provide port-based security for host servers.

Use Private VLANs to partition a VLAN domain into sub-domains. Each sub-domain is comprised of one primary VLAN and one or more secondary VLANs, as follows:

● Primary VLAN—carries unidirectional traffic downstream from promiscuous ports. Each Private VLAN configuration has only one primary VLAN. All ports in the Private VLAN are members of the primary VLAN.

● Secondary VLAN—Secondary VLANs are internal to a private VLAN domain, and are defined as follows:

　○ Isolated VLAN—carries unidirectional traffic upstream from the host servers toward ports in the primary VLAN. Each Private VLAN configuration can contain only one isolated VLAN.

　○ Community VLAN—carries upstream traffic from ports in the community VLAN to other ports in the same community, and to ports in the primary VLAN. Each Private VLAN configuration can contain multiple community VLANs.

After you define the primary VLAN and one or more secondary VLANs, you map the secondary VLAN(s) to the primary VLAN.

Private VLAN Ports

Private VLAN ports are defined as follows:

● Promiscuous—A promiscuous port is a port that belongs to the primary VLAN. The promiscuous port can communicate with all the interfaces, including ports in the secondary VLANs (Isolated VLAN and Community VLANs).

● Isolated—An isolated port is a host port that belongs to an isolated VLAN. Each isolated port has complete layer 2 separation from other ports within the same private VLAN (including other isolated ports), except for the promiscuous ports.

　○ Traffic sent to an isolated port is blocked by the Private VLAN, except the traffic from promiscuous ports.

　○ Traffic received from an isolated port is forwarded only to promiscuous ports.

● Community—A community port is a host port that belongs to a community VLAN. Community ports can communicate with other ports in the same community VLAN, and with promiscuous ports. These interfaces are isolated at layer 2 from all other interfaces in other communities and from isolated ports within the Private VLAN.

Configuration Guidelines

The following guidelines apply when configuring Private VLANs:

- Management VLANs cannot be Private VLANs. Management ports cannot be members of a Private VLAN.

- The default VLAN 1 cannot be a Private VLAN.

- IGMP Snooping must be disabled on Private VLANs.

- All VLANs that comprise the Private VLAN must belong to the same Spanning Tree Group.

- A secondary port's (isolated port and community port) PVID/Native VLAN must match its corresponding secondary VLAN ID.

- Ports in a secondary VLAN cannot be members of other VLANs.

Configuration Example

Follow this procedure to configure a Private VLAN.

1. Select a VLAN and define the Private VLAN type as primary.

```
RS G8124-E(config)# vlan 700
RS G8124-E(config-vlan)# private-vlan primary
RS G8124-E(config-vlan)# exit
```

2. Configure a promiscuous port for VLAN 700.

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# switchport mode private-vlan promiscuous
RS G8124-E(config-if)# switchport private-vlan mapping 700
RS G8124-E(config-if)# exit
```

3. Configure two secondary VLANs: isolated VLAN and community VLAN.

```
RS G8124-E(config)# vlan 701
RS G8124-E(config-vlan)# private-vlan isolated
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# vlan 702
RS G8124-E(config-vlan)# private-vlan community
RS G8124-E(config-vlan)# exit
```

4. Map secondary VLANs to primary VLAN.

```
RS G8124-E(config)# vlan 700-702
RS G8124-E(config-vlan)# stg 1
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# vlan 700
RS G8124-E(config-vlan)# private-vlan association 701,702
RS G8124-E(config-vlan)# exit
```

5. Configure host ports for secondary VLANs.

```
RS G8124-E(config)# interface port 2
RS G8124-E(config-if)# switchport mode private-vlan host
RS G8124-E(config-if)# switchport private-vlan association 700 701
RS G8124-E(config-if)# exit

RS G8124-E(config)# interface port 3
RS G8124-E(config-if)# switchport mode private-vlan host
RS G8124-E(config-if)# switchport private-vlan association 700 702
RS G8124-E(config-if)# exit
```

6. Verify the configuration.

```
RS G8124-E(config)# show vlan private-vlan

Private-VLAN    Type       Mapped-To   Status        Ports
------------    ---------  ----------  ----------    -----------------
700             primary    701,702     ena           1
701             isolated   700         ena           2
702             community  700         ena           3
```

Chapter 8. Ports and Link Aggregation

Link Aggregation (LAG) groups can provide super-bandwidth, multi-link connections between the G8124-E and other LAG-capable devices. A LAG is a group of ports that act together, combining their bandwidth to create a single, larger virtual link. This chapter provides configuration background and examples for aggregating multiple ports together:

- *"Aggregation Overview" on page 136"*
- *"Configuring a Static LAG" on page 138*
- *"Configurable LAG Hash Algorithm" on page 145*
- *"Link Aggregation Control Protocol" on page 140*

Aggregation Overview

When using LAGs between two switches, as shown in Figure 7, you can create a virtual link between the switches, operating with combined throughput levels that depends on how many physical ports are included.

The G8124-E supports two types of aggregation: static LAGs, and dynamic Link Aggregation Control Protocol (LACP) LAGs. You may configure up to 24 LAGs on the switch, with both types (static and LACP) sharing the same pool. Of the available configuration slots, any or all may be used for LACP LAGs, though only up to 12 may used for static LAGs. In addition, although up to a total of 24 LAGs may be configured and enabled, only a maximum of 16 may be operational at any given time.

For example, if you configure and enable 12 static LAGs (the maximum), up to 4 LACP LAGs may also be configured and enabled, for a total of 16 operational LAGs. If more than 16 LAGs are enabled at any given time, once the switch establishes the 16th LAG, any additional LAGs are automatically placed in a non-operational state. In this scenario, there is no administrative means to ensure which 16 LAGs are selected for operation.

Figure 7. Port LAG
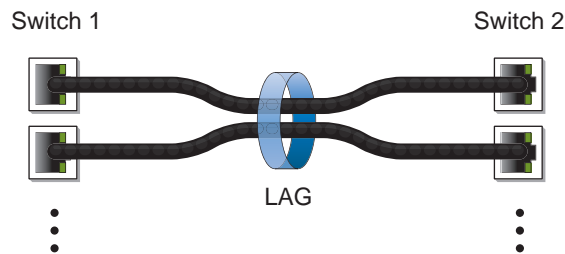
LAGs are also useful for connecting a G8124-E to third-party devices that support link aggregation, such as Cisco routers and switches with EtherChannel technology (*not* ISL aggregation technology) and Sun's Quad Fast Ethernet Adapter. LAG technology is compatible with these devices when they are configured manually.

LAG traffic is statistically distributed among the ports in a LAG, based on a variety of configurable options.

Also, since each LAG is comprised of multiple physical links, the LAG is inherently fault tolerant. As long as one connection between the switches is available, the trunk remains active and statistical load balancing is maintained whenever a port in a LAG is lost or returned to service.

Static LAGs

When you create and enable a static LAG, the LAG members (switch ports) take on certain settings necessary for correct operation of the aggregation feature.

Static LAG Requirements

Before you configure your LAG, you must consider these settings, along with specific configuration rules, as follows:

1. Read the configuration rules provided in the section, "Static Aggregation Configuration Rules" on page 137.

2. Determine which switch ports (up to 12) are to become *LAG members* (the specific ports making up the LAG).

3. Ensure that the chosen switch ports are set to `enabled`. LAG member ports must have the same VLAN and Spanning Tree configuration.

4. Consider how the existing Spanning Tree will react to the new LAG configuration. See Chapter 9, "Spanning Tree Protocols," for Spanning Tree Group configuration guidelines.

5. Consider how existing VLANs will be affected by the addition of a LAG.

Static Aggregation Configuration Rules

The aggregation feature operates according to specific configuration rules. When creating LAGs, consider the following rules that determine how a LAG reacts in any network topology:

- All links must originate from one logical device, and lead to one logical destination device. Usually, a LAG connects two physical devices together with multiple links. However, in some networks, a single logical device may include multiple physical devices or when using VLAGs (see Chapter 10, "Virtual Link Aggregation Groups). In such cases, links in a LAG are allowed to connect to multiple physical devices because they act as one logical device.
- Any physical switch port can belong to only one LAG.
- Aggregation from third-party devices must comply with Cisco® EtherChannel® technology.
- All ports in a LAG must have the same link configuration (speed, duplex, flow control), the same VLAN properties, and the same Spanning Tree, storm control, and ACL configuration. It is recommended that the ports in a LAG be members of the same VLAN.
- Each LAG inherits its port configuration (speed, flow control, tagging) from the first member port. As additional ports are added to the LAG, their settings must be changed to match the LAG configuration.
- When a port leaves a LAG, its configuration parameters are retained.
- You cannot configure a LAG member as a monitor port in a port-mirroring configuration.
- LAGs cannot be monitored by a monitor port; however, LAG members can be monitored.

Configuring a Static LAG

In the following example, three ports are aggregated between two switches.

Figure 8. LAG Configuration Example
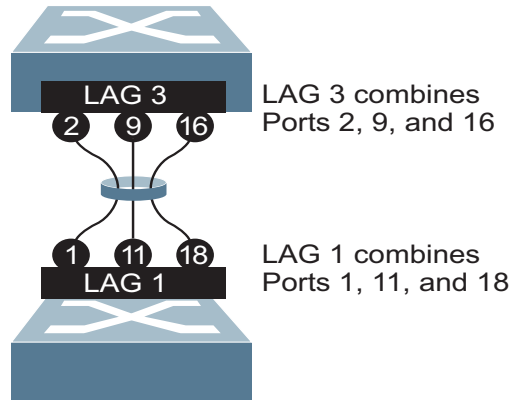
Prior to configuring each switch in this example, you must connect to the appropriate switches as the administrator.

Note: For details about accessing and using any of the commands described in this example, see the *RackSwitch G8124-E ISCLI Reference*.

1. Follow these steps on the G8124-E:
   a. Define a LAG.

   ```
   RS G8124-E(config)# portchannel 3 port 2,9,16
   RS G8124-E(config)# portchannel 3 enable
   ```

   b. Verify the configuration.

   ```
   # show portchannel information
   ```

   Examine the resulting information. If any settings are incorrect, make appropriate changes.

2. Repeat the process on the other switch.

   ```
   RS G8124-E(config)# portchannel 1 port 1,11,18
   RS G8124-E(config)# portchannel 1 enable
   ```

3. Connect the switch ports that will be members in the LAG.

   LAG 3 (on the G8124-E) is now connected to LAG 1 (on the other switch).

   Note: In this example, two G8124-E switches are used. If a third-party device supporting link aggregation is used (such as Cisco routers and switches with Ether-Channel technology or Sun's Quad Fast Ethernet Adapter), LAGs on the third-party device must be configured manually. Connection problems could arise when using automatic LAG negotiation on the third-party device.

4. Examine the aggregation information on each switch.

```
# show portchannel information
PortChannel 3: Enabled
Protocol—Static
port state:
  2: STG  1 forwarding
  9: STG  1 forwarding
 16: STG  1 forwarding
```

Information about each port in each configured LAG is displayed. Make sure that LAGs consist of the expected ports and that each port is in the expected state.

The following restrictions apply:

● Any physical switch port can belong to only one LAG.

● Up to 12 ports can belong to the same LAG.

● All ports in static LAGs must be have the same link configuration (speed, duplex, flow control).

● Aggregation with third-party devices must comply with Cisco® EtherChannel® technology.

Link Aggregation Control Protocol

Link Aggregation Control Protocol (LACP) is an IEEE 802.3ad standard for grouping several physical ports into one logical port (known as a Link Aggregation group) with any device that supports the standard. Please refer to IEEE 802.3ad-2002 for a full description of the standard.

The 802.3ad standard allows standard Ethernet links to form a single Layer 2 link using Link Aggregation Control Protocol (LACP). Link aggregation is a method of grouping physical link segments of the same media type and speed in full duplex, and treating them as if they were part of a single, logical link segment. If a link in an LACP LAG fails, traffic is reassigned dynamically to the remaining link(s) of the dynamic LAG.

Note: LACP implementation in the Lenovo Networking OS does not support the Churn machine, an option used to detect if the port is operable within a bounded time period between the actor and the partner. Only the Marker Responder is implemented, and there is no marker protocol generator.

A port's Link Aggregation Identifier (LAG ID) determines how the port can be aggregated. The Link Aggregation ID (LAG ID) is constructed mainly from the partner switch's *system ID* and the port's *admin key*, as follows:

- **System ID**: an integer value based on the partner switch's MAC address and the system priority assigned in the CLI.

- **Admin key**: a port's Admin key is an integer value (1-65535) that you can configure in the CLI. Each switch port that participates in the same LACP LAG must have the same *admin key* value. The Admin key is *local significant*, which means the partner switch does not need to use the same Admin key value.

For example, consider two switches, an Actor (the G8124-E) and a Partner (another switch), as shown in Table 14.

Table 14. *Actor vs. Partner LACP configuration*

Actor Switch	Partner Switch	LACP LAG
Port 7 (admin key = 100)	Port 1 (admin key = 50)	Primary LAG
Port 8 (admin key = 100)	Port 2 (admin key = 50)	Primary LAG
Port 9 (admin key = 100)	Port 3 (admin key = 70)	Secondary LAG
Port 10 (admin key = 100)	Port 4 (admin key = 70)	Secondary LAG

In the configuration shown in Table 14, Actor switch ports 7 and 8 aggregate to form an LACP LAG with Partner switch ports 1 and 2. Only ports with the same LAG ID are aggregated in the LAG. Actor switch ports 9 and 10 are not aggregated in the same LAG, because although they have the same admin key on the Actor switch, their LAG IDs are different (due to a different Partner switch admin key configuration). Instead, they form a secondary LAG with Partner switch ports 3 and 4.

To avoid the Actor switch ports (with the same admin key) from aggregating in another LAG, you can configure a group ID. Ports with the same admin key (although with different LAG IDs) compete to get aggregated in a LAG. The LAG

ID is decided based on the first port that is aggregated in the group. Ports with this LAG ID get aggregated and the other ports are placed in `suspended` mode. In the configuration shown in Table 14, if port 7 gets aggregated first, then the LAG ID of port 7 would be the LAG ID of the LAG. Port 9 would be placed in suspended mode. When in suspended mode, a port transmits only LACP data units (LACPDUs) and discards all other traffic.

A port may also be placed in suspended mode for the following reasons:

- When LACP is configured on the port but it stops receiving LACPDUs from the partner switch.

- When the port has a different LAG ID because of the partner switch MAC or port LACP key being different. For example: when a switch is connected to two partners.

The LAG ID can be configured using the following command:

```
RS G8124-E(config)# portchannel <13-36> lacp key <adminkey of the LAG>
suspend-individual
```

LACP automatically determines which member links can be aggregated and then aggregates them. It provides for the controlled addition and removal of physical links for the link aggregation.

Static LACP LAGs

To prevent switch ports with the same admin key from forming multiple LAGs, you can configure the LACP LAG as static. In a static LACP LAG, ports with the same admin key, but with different LAG IDs, compete to get aggregated in a LAG. The LAG ID for the LAG is decided based on the first port that is aggregated in the group. Ports with this LAG ID get aggregated and the other ports are placed in `suspended` mode. As per the configuration shown in Table 14 on page 140, if port 7 gets aggregated first, then the LAG ID of port 7 would be the LAG ID of the LAG. Port 8 will join the LAG while ports 9 and 10 would be placed in suspended mode. When in suspended mode, a port transmits only LACP data units (LACPDUs) and discards all other traffic.

LACP Port Modes

Each port on the switch can have one of the following LACP modes.

- off (default)
  The user can configure this port in to a regular static LAG.

- active
  The port is capable of forming an LACP LAG. This port sends LACPDU packets to partner system ports.

- passive
  The port is capable of forming an LACP LAG. This port only responds to the LACPDU packets sent from an LACP *active* port.

Each active LACP port transmits LACP data units (LACPDUs), while each passive LACP port listens for LACPDUs. During LACP negotiation, the admin key is exchanged. The LACP LAG is enabled as long as the information matches at both ends of the link. If the admin key value changes for a port at either end of the link, that port's association with the LACP LAG is lost.

When the system is initialized, all ports by default are in LACP *off* mode and are assigned unique *admin keys*. To make a group of ports aggregable, you assign them all the same *admin key*. You must set the port's LACP mode to *active* to activate LACP negotiation. You can set other port's LACP mode to passive, to reduce the amount of LACPDU traffic at the initial LAG-forming stage.

Use the following command to check whether the ports are aggregated:

```
RS G8124-E # show lacp information
```

LACP Individual

Ports with LACP enabled (active or passive) are prevented by default from forming individual links if they cannot join an LACP LAG. To override this behavior, use the following commands:

```
RS G8124-E(config) # interface port <port no. or range>
RS G8124-E(config-if) # no lacp suspend-individual
```

This allows the selected ports to be treated as normal link-up ports, which may forward data traffic according to STP, Hot Links or other applications, if they do not receive any LACPDUs.

To configure the LACP individual setting for all the ports in a static LACP LAG, use the following commands:

```
RS G8124-E(config)# interface portchannel lacp <adminkey of the LAG>
RS G8124-E(config-PortChannel)# [no] lacp suspend-individual
```

LACP Minimum Links Option

For dynamic LAGs that require a guaranteed amount of bandwidth to be considered useful, you can specify the minimum number of links for the LAG. If the specified minimum number of ports is not available, the LAG link will not be established. If an active LACP LAG loses one or more component links, the LAG will be placed in the down state if the number of links falls to less than the specified minimum. By default, the minimum number of links is 1, meaning that LACP LAGs will remain operational as long as at least one link is available.

The LACP minimum links setting can be configured as follows:

● Via interface configuration mode:

```
RS G8124-E(config)# interface port <port number or range>
RS G8124-E(config-if)# port-channel min-links <minimum links>
RS G8124-E(config-if)# exit
```

● Or via portchannel configuration mode:

```
RS G8124-E(config)# interface portchannel lacp <LACP key>
RS G8124-E(config-PortChannel)# port-channel min-links <minimum links>
RS G8124-E(config-if)# exit
```

Configuring LACP

Use the following procedure to configure LACP for ports 7, 8, 9 and 10 to participate in link aggregation.

1. Configure port parameters. All ports that participate in the LACP LAG must have the same settings, including VLAN membership.

2. Select the port range and define the admin key. Only ports with the same admin key can form an LACP LAG.

```
RS G8124-E(config)# interface port 7-10
RS G8124-E(config-if)# lacp key 100
```

3. Set the LACP mode.

```
RS G8124-E(config-if)# lacp mode active
```

4. Optionally allow member ports to individually participate in normal data traffic if no LACPDUs are received.

```
RS G8124-E(config-if)# no lacp suspend-individual
RS G8124-E(config-if)# exit
```

5. Set the link aggregation as static, by associating it with LAG ID 13:

```
RS G8124-E(config)# portchannel 13 lacp key 100
```

Configurable LAG Hash Algorithm

Traffic in a LAG is statistically distributed among member ports using a *hash* process where various address and attribute bits from each transmitted frame are recombined to specify the particular LAG port the frame will use.

The switch can be configured to use a variety of hashing options. To achieve the most even traffic distribution, select options that exhibit a wide range of values for your particular network. Avoid hashing on information that is not usually present in the expected traffic, or which does not vary.

The G8124-E supports the following hashing options:

- Layer 2 source MAC address:

```
RS G8124-E(config)# portchannel hash source-mac-address
```

- Layer 2 destination MAC address:

```
RS G8124-E(config)# portchannel hash destination-mac-address
```

- Layer 2 source and destination MAC address:

```
RS G8124-E(config)# portchannel hash source-destination-mac
```

- Layer 3 IPv4/IPv6 source IP address:

```
RS G8124-E(config)# portchannel hash source-ip-address
```

- Layer 3 IPv4/IPv6 destination IP address:

```
RS G8124-E(config)# portchannel hash destination-ip-address
```

- Layer 3 source and destination IPv4/IPv6 address (the default):

```
RS G8124-E(config)# portchannel hash source-destination-ip
```

Note: Layer 3 hashing options (source IP address and destination IP address) enabled either for port LAG hashing or for ECMP route hashing (ip route ecmphash) apply to both the aggregation and ECMP features (the enabled settings are cumulative). If LAG hashing behavior is not as expected, disable any unwanted options set in ECMP route hashing. Likewise, if ECMP route hashing behavior is not as expected, disable any unwanted options enabled in LAG hashing.

Chapter 9. Spanning Tree Protocols

When multiple paths exist between two points on a network, Spanning Tree Protocol (STP), or one of its enhanced variants, can prevent broadcast loops and ensure that the RackSwitch G8124-E uses only the most efficient network path.

This chapter covers the following topics:

Spanning Tree Protocol Modes

Lenovo Networking OS 8.3 supports the following STP modes:

- Rapid Spanning Tree Protocol (RSTP)

  IEEE 802.1D (2004) RSTP allows devices to detect and eliminate logical loops in a bridged or switched network. When multiple paths exist, STP configures the network so that only the most efficient path is used. If that path fails, STP automatically configures the best alternative active path on the network to sustain network operations. RSTP is an enhanced version of IEEE 802.1D (1998) STP, providing more rapid convergence of the Spanning Tree network path states on STG 1.

  See "Rapid Spanning Tree Protocol" on page 160 for details.

- Per-VLAN Rapid Spanning Tree (PVRST)

  PVRST mode is based on RSTP to provide rapid Spanning Tree convergence, but supports instances of Spanning Tree, allowing one STG per VLAN. PVRST mode is compatible with Cisco R-PVST/R-PVST+ mode.

  PVRST is the default Spanning Tree mode on the G8124-E. See "PVRST Mode" on page 148 for details.

- Multiple Spanning Tree Protocol (MSTP)

  IEEE 802.1Q (2003) MSTP provides both rapid convergence and load balancing in a VLAN environment. MSTP allows multiple STGs, with multiple VLANs in each.

  See "Multiple Spanning Tree Protocol" on page 162 for details.

Global STP Control

By default, the Spanning Tree feature is globally enabled on the switch, and is set for PVRST mode. Spanning Tree (and thus any currently configured STP mode) can be globally disabled using the following command:

```
RS G8124-E(config)# spanning-tree mode disable
```

Spanning Tree can be re-enabled by specifying the STP mode:

```
RS G8124-E(config)# spanning-tree mode {pvrst|rstp|mst}
```

where the command options represent the following modes:

- rstp: RSTP mode
- pvrst: PVRST mode
- mst: MSTP mode

PVRST Mode

Note: Per-VLAN Rapid Spanning Tree (PVRST) is enabled by default on the G8124-E.

Using STP, network devices detect and eliminate logical loops in a bridged or switched network. When multiple paths exist, Spanning Tree configures the network so that a switch uses only the most efficient path. If that path fails, Spanning Tree automatically sets up another active path on the network to sustain network operations.

N/OS PVRST mode is based on IEEE 802.1w RSTP. Like RSTP, PVRST mode provides rapid Spanning Tree convergence. However, PVRST mode is enhanced for multiple instances of Spanning Tree. In PVRST mode, each VLAN may be automatically or manually assigned to one of 127 available STGs. Each STG acts as an independent, simultaneous instance of STP. PVRST uses IEEE 802.1Q tagging to differentiate STP BPDUs and is compatible with Cisco R-PVST/R-PVST+ modes.

The relationship between ports, LAGs, VLANs, and Spanning Trees is shown in Table 15.

Table 15. *Ports, LAGs, and VLANs*

Switch Element	Belongs To
Port	LAG or one or more VLANs
LAG	One or more VLANs
VLAN (non-default)	• PVRST: One VLAN per STG • RSTP: All VLANs are in STG 1 • MSTP: Multiple VLANs per STG

Port States

The port state controls the forwarding and learning processes of Spanning Tree. In PVRST, the port state has been consolidated to the following: `discarding`, `learning`, and `forwarding`.

Due to the sequence involved in these STP states, considerable delays may occur while paths are being resolved. To mitigate delays, ports defined as *edge* ports ("Port Type and Link Type" on page 166) may bypass the `discarding` and `learning` states, and enter directly into the `forwarding` state.

Bridge Protocol Data Units

To create a Spanning Tree, the switch generates a configuration Bridge Protocol Data Unit (BPDU), which it then forwards out of its ports. All switches in the Layer 2 network participating in the Spanning Tree gather information about other switches in the network through an exchange of BPDUs.

A bridge sends BPDU packets at a configurable regular interval (2 seconds by default). The BPDU is used to establish a path, much like a hello packet in IP routing. BPDUs contain information about the transmitting bridge and its ports, including bridge MAC addresses, bridge priority, port priority, and path cost. If the ports are in trunk mode/tagged, each port sends out a special BPDU containing the tagged information.

The generic action of a switch on receiving a BPDU is to compare the received BPDU to its own BPDU that it will transmit. If the priority of the received BPDU is better than its own priority, it will replace its BPDU with the received BPDU. Then, the switch adds its own bridge ID number and increments the path cost of the BPDU. The switch uses this information to block any necessary ports.

Note: If STP is globally disabled, BPDUs from external devices will transit the switch transparently. If STP is globally enabled, for ports where STP is turned off, inbound BPDUs will instead be discarded.

Determining the Path for Forwarding BPDUs

When determining which port to use for forwarding and which port to block, the G8124-E uses information in the BPDU, including each bridge ID. A technique based on the "lowest root cost" is then computed to determine the most efficient path for forwarding.

Bridge Priority

The bridge priority parameter controls which bridge on the network is the STG root bridge. To make one switch become the root bridge, configure the bridge priority lower than all other switches and bridges on your network. The lower the value, the higher the bridge priority. Use the following command to configure the bridge priority:

```
RS G8124-E(config)# spanning-tree stp <STG number or range> bridge priority
<0-65535>
```

Port Priority

The port priority helps determine which bridge port becomes the root port or the designated port. The case for the root port is when two switches are connected using a minimum of two links with the same path-cost. The case for the designated port is in a network topology that has multiple bridge ports with the same path-cost connected to a single segment, the port with the lowest port priority becomes the designated port for the segment. Use the following command to configure the port priority:

```
RS G8124-E(config-if)# spanning-tree stp <STG number or range> priority <port priority>
```

where *priority value* is a number from 0 to 240, in increments of 16 (such as 0, 16, 32, and so on). If the specified priority value is not evenly divisible by 16, the value will be automatically rounded down to the nearest valid increment whenever manually changed in the configuration, or whenever a configuration file from a release prior to N/OS 6.5 is loaded.

Root Guard

The root guard feature provides a way to enforce the root bridge placement in the network. It keeps a new device from becoming root and thereby forcing STP re-convergence. If a root-guard enabled port detects a root device, that port will be placed in a blocked state.

You can configure the root guard at the port level using the following commands:

```
RS G8124-E(config)# interface port <port number>
RS G8124-E(config-if)# spanning-tree guard root
```

The default state is "none", i.e. disabled.

Loop Guard

In general, STP resolves redundant network topologies into loop-free topologies. The loop guard feature performs additional checking to detect loops that might not be found using Spanning Tree. STP loop guard ensures that a non-designated port does not become a designated port.

To globally enable loop guard, enter the following command:

```
RS G8124-E(config)# spanning-tree loopguard
```

Note: The global loop guard command will be effective on a port only if the port-level loop guard command is set to default as shown below:
```
RS G8124-E(config)# interface port <port number>
RS G8124-E(config-if)# no spanning-tree guard
```

To enable loop guard at the port level, enter the following command:

```
RS G8124-E(config)# interface port <port number>
RS G8124-E(config-if)# spanning-tree guard loop
```

The default state is "none" (disabled).

Port Path Cost

The port path cost assigns lower values to high-bandwidth ports, such as 10 Gigabit Ethernet, to encourage their use. The objective is to use the fastest links so that the route with the lowest cost is chosen. A value of 0 (the default) indicates that the default cost will be computed for an auto-negotiated link or LAG speed.

Use the following command to modify the port path cost:

```
RS G8124-E(config)# interface port <port number>
RS G8124-E(config-if)# spanning-tree stp <STG number or range> path-cost <path
cost value>
RS G8124-E(config-if)# exit
```

The port path cost can be a value from 1 to 200000000. Specify 0 for automatic path cost.

Simple STP Configuration

Figure 9 depicts a simple topology using a switch-to-switch link between two G8124-E 1 and 2.

Figure 9. Spanning Tree Blocking a Switch-to-Switch Link
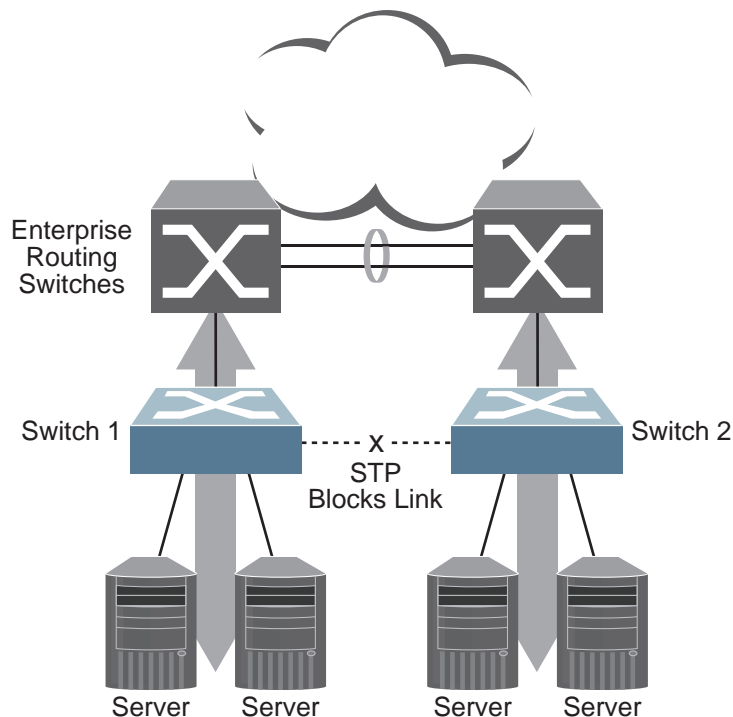
To prevent a network loop among the switches, STP must block one of the links between them. In this case, it is desired that STP block the link between the Lenovo switches, and not one of the G8124-E uplinks or the Enterprise switch LAG.

During operation, if one G8124-E experiences an uplink failure, STP will activate the Lenovo switch-to-switch link so that server traffic on the affected G8124-E may pass through to the active uplink on the other G8124-E, as shown in Figure 10.

Figure 10. Spanning Tree Restoring the Switch-to-Switch Link
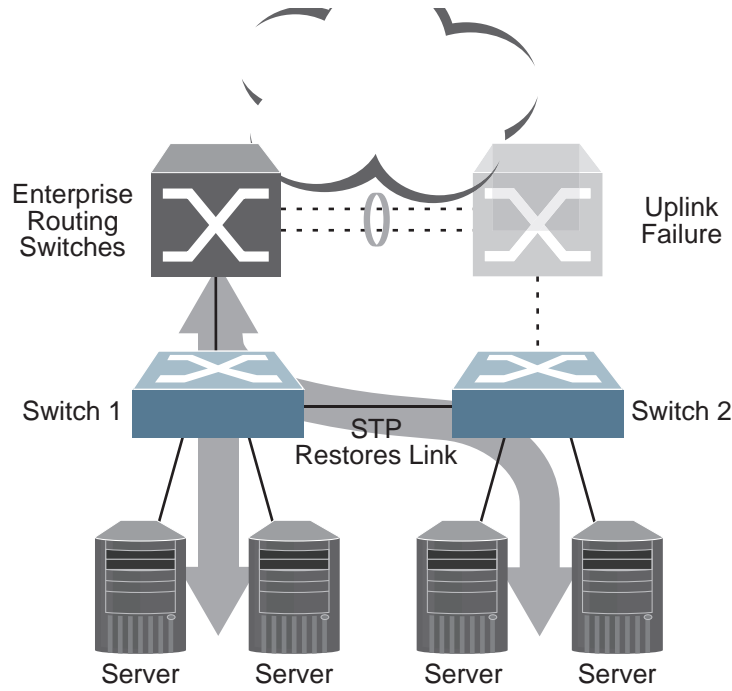
In this example, port 10 on each G8124-E is used for the switch-to-switch link. To ensure that the G8124-E switch-to-switch link is blocked during normal operation, the port path cost is set to a higher value than other paths in the network. To configure the port path cost on the switch-to-switch links in this example, use the following commands on each G8124-E.

```
RS G8124-E(config)# interface port 10
RS G8124-E(config-if)# spanning-tree stp 1 path-cost 60000
RS G8124-E(config-if)# exit
```

Per-VLAN Spanning Tree Groups

PVRST mode supports a maximum of 127 STGs, with each STG acting as an independent, simultaneous instance of STP.

Multiple STGs provide multiple data paths which can be used for load-balancing and redundancy. To enable load balancing between two G8124-Es using multiple STGs, configure each path with a different VLAN and then assign each VLAN to a separate STG. Since each STG is independent, they each send their own IEEE 802.1Q tagged Bridge Protocol Data Units (BPDUs).

Each STG behaves as a bridge group and forms a loop-free topology. The default STG 1 may contain multiple VLANs (typically until they can be assigned to another STG). STGs 2-127 may contain only one VLAN each.

Using Multiple STGs to Eliminate False Loops

Figure 11 shows a simple example of why multiple STGs are needed. In the figure, two ports on a G8124-E are connected to two ports on an application switch. Each of the links is configured for a different VLAN, preventing a network loop. However, in the first network, since a single instance of Spanning Tree is running on all the ports of the G8124-E, a physical loop is assumed to exist, and one of the VLANs is blocked, impacting connectivity even though no actual loop exists.

Figure 11. Using Multiple Instances of Spanning Tree Group
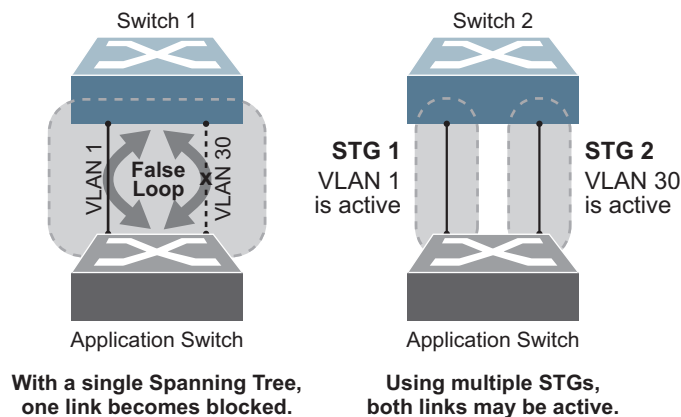
In the second network, the problem of improper link blocking is resolved when the VLANs are placed into different Spanning Tree Groups (STGs). Since each STG has its own independent instance of Spanning Tree, each STG is responsible only for the loops within its own VLAN. This eliminates the false loop, and allows both VLANs to forward packets between the switches at the same time.

VLANs and STG Assignment

In PVRST mode, up to 128 STGs are supported. Ports cannot be added directly to an STG. Instead, ports must be added as members of a VLAN, and the VLAN must then be assigned to the STG.

STG 1 is the default STG. Although VLANs can be added to or deleted from default STG 1, the STG itself cannot be deleted from the system. By default, STG 1 is enabled and includes VLAN 1, which by default includes all switch ports (except for management VLANs and management ports)

STG 128 is reserved for switch management. By default, STG 128 is disabled, but includes management VLAN 4095 and the management ports (MGMT-A and MGMT-B).

By default, all other STGs (STG 2 through 127) are enabled, though they initially include no member VLANs. VLANs must be assigned to STGs. By default, this is done automatically using VLAN Automatic STG Assignment (VASA), though it can also be done manually (see .

When VASA is enabled (as by default), each time a new VLAN is configured, the switch will automatically assign that new VLAN to its own STG. Conversely, when a VLAN is deleted, if its STG is not associated with any other VLAN, the STG is returned to the available pool.

The specific STG number to which the VLAN is assigned is based on the VLAN number itself. For low VLAN numbers (1 through 127), the switch will attempt to assign the VLAN to its matching STG number. For higher numbered VLANs, the STG assignment is based on a simple modulus calculation; the attempted STG number will "wrap around," starting back at the top of STG list each time the end of the list is reached. However, if the attempted STG is already in use, the switch will select the next available STG. If an empty STG is not available when creating a new VLAN, the VLAN is automatically assigned to default STG 1.

If ports are tagged, each tagged port sends out a special BPDU containing the tagged information. Also, when a tagged port belongs to more than one STG, the egress BPDUs are tagged to distinguish the BPDUs of one STG from those of another STG.

VASA is enabled by default, but can be disabled or re-enabled using the following commands:

```
RS G8124-E(config)# [no] spanning-tree stg-auto
```

If VASA is disabled, when you create a new VLAN, that VLAN automatically belongs to default STG 1. To place the VLAN in a different STG, assign it manually.

VASA applies only to PVRST mode and is ignored in RSTP and MSTP modes.

Manually Assigning STGs

The administrator may manually assign VLANs to specific STGs, whether or not VASA is enabled.

1. If no VLANs exist (other than default VLAN 1), see "Guidelines for Creating VLANs" on page 155 for information about creating VLANs and assigning ports to them.

2. Assign the VLAN to an STG using one of the following methods:

- From the global configuration mode:

```
RS G8124-E(config)# spanning-tree stp <STG numbers> vlan <VLANs>
```

- Or from within the VLAN configuration mode:

```
RS G8124-E(config)# vlan <VLAN numbers>
RS G8124-E(config-vlan)# stg <STG number>
RS G8124-E(config-vlan)# exit
```

When a VLAN is assigned to a new STG, the VLAN is automatically removed from its prior STG.

Note: For proper operation with switches that use Cisco PVST+, it is recommended that you create a separate STG for each VLAN.

Guidelines for Creating VLANs

Follow these guidelines when creating VLANs:

- When you create a new VLAN, if VASA is enabled (the default), that VLAN is automatically assigned its own STG. If VASA is disabled, the VLAN automatically belongs to STG 1, the default STG. To place the VLAN in a different STG, see "Manually Assigning STGs" on page 155. The VLAN is automatically removed from its old STG before being placed into the new STG.

- Each VLANs must be contained *within* a single STG; a VLAN cannot span multiple STGs. By confining VLANs within a single STG, you avoid problems with Spanning Tree blocking ports and causing a loss of connectivity within the VLAN. When a VLAN spans multiple switches, it is recommended that the VLAN remain within the same STG (be assigned the same STG ID) across all the switches.

- If ports are tagged, all aggregated ports can belong to multiple STGs.

- A port cannot be directly added to an STG. The port must first be added to a VLAN, and that VLAN added to the desired STG.

Rules for VLAN Tagged/Trunk Mode Ports

The following rules apply to VLAN tagged ports:

- Tagged/trunk mode ports can belong to more than one STG, but untagged/access mode ports can belong to only one STG.

- When a tagged/trunk mode port belongs to more than one STG, the egress BPDUs are tagged to distinguish the BPDUs of one STG from those of another STG.

Adding and Removing Ports from STGs

The following rules apply when you add ports to or remove ports from STGs:

- When you add a port to a VLAN that belongs to an STG, the port is also added to that STG. However, if the port you are adding is an untagged port and is already a member of another STG, that port will be removed from its current STG and added to the new STG. An untagged port cannot belong to more than one STG.

  For example: Assume that VLAN 1 belongs to STG 1, and that port 1 is untagged and does not belong to any STG. When you add port 1 to VLAN 1, port 1 will automatically become part of STG 1.

  However, if port 5 is untagged and is a member of VLAN 3 in STG 2, then adding port 5 to VLAN 1 in STG 1 will change the port PVID from 3 to 1:

  ```
  "Port 5 is an UNTAGGED/Access Mode port and its PVID/Native-VLAN
  changed from 3 to 1.
  ```

- When you remove a port from VLAN that belongs to an STG, that port will also be removed from the STG. However, if that port belongs to another VLAN in the same STG, the port remains in the STG.

  As an example, assume that port 2 belongs to only VLAN 2, and that VLAN 2 belongs to STG 2. When you remove port 2 from VLAN 2, the port is moved to default VLAN 1 and is removed from STG 2.

  However, if port 2 belongs to both VLAN 1 and VLAN 2, and both VLANs belong to STG 1, removing port 2 from VLAN 2 does not remove port 2 from STG 1 because the port is still a member of VLAN 1, which is still a member of STG 1.

- An STG cannot be deleted, only disabled. If you disable the STG while it still contains VLAN members, Spanning Tree will be off on all ports belonging to that VLAN.

The relationship between port, LAGs, VLANs, and Spanning Trees is shown in .

The Switch-Centric Model

PVRST is switch-centric: STGs are enforced only on the switch where they are configured. PVRST allows only one VLAN per STG, except for the default STG 1 to which multiple VLANs can be assigned. The STG ID is not transmitted in the Spanning Tree BPDU. Each Spanning Tree decision is based entirely on the configuration of the particular switch.

For example, in Figure 12, each switch is responsible for the proper configuration of its own ports, VLANs, and STGs. Switch A identifies its own port 17 as part of VLAN 2 on STG 2, and the Switch B identifies its own port 8 as part of VLAN 2 on STG 2.

Figure 12. Implementing PVRST
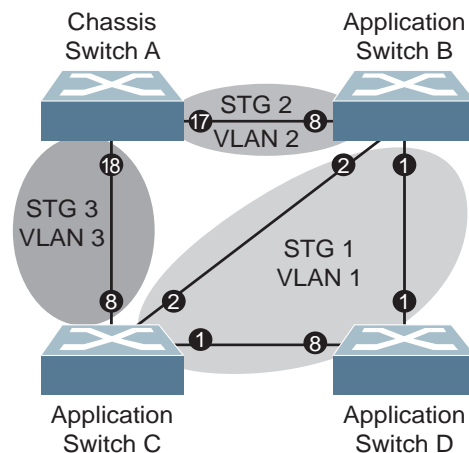
The VLAN participation for each Spanning Tree Group in Figure 12 on page 157 is as follows:

- VLAN 1 Participation

  Assuming Switch B to be the root bridge, Switch B transmits the BPDU for STG 1 on ports 1 and 2. Switch C receives the BPDU on port 2, and Switch D receives the BPDU on port 1. Because there is a network loop between the switches in VLAN 1, either Switch D will block port 8 or Switch C will block port 1, depending on the information provided in the BPDU.

- VLAN 2 Participation

  Switch B, the root bridge, generates a BPDU for STG 2 from port 8. Switch A receives this BPDU on port 17, which is assigned to VLAN 2, STG 2. Because switch B has no additional ports participating in STG 2, this BPDU is not forwarded to any additional ports and Switch B remains the designated root.

- VLAN 3 Participation

  For VLAN 3, Switch A or Switch C may be the root bridge. If Switch A is the root bridge for VLAN 3, STG 3, then Switch A transmits the BPDU from port 18. Switch C receives this BPDU on port 8 and is identified as participating in VLAN 3, STG 3. Since Switch C has no additional ports participating in STG 3, this BPDU is not forwarded to any additional ports and Switch A remains the designated root.

Configuring Multiple STGs

This configuration shows how to configure the three instances of STGs on the switches A, B, C, and D illustrated in .

Because VASA is enabled by default, each new VLAN is automatically assigned its own STG.

1. Set the Spanning Tree mode on each switch to PVRST.

```
RS G8124-E(config)# spanning-tree mode pvrst
```

Note: PVRST is the default mode on the G8124-E. This step is not required unless the STP mode has been previously changed, and is shown here merely as an example of manual configuration.

2. Configure the following on Switch A:

a. Enable VLAN 2 and VLAN 3.

```
RS G8124-E(config)# vlan 2
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# vlan 3
RS G8124-E(config-vlan)# exit

If VASA is disabled, enter the following commands:
RS G8124-E(config)# spanning-tree stp 2 vlan 2
RS G8124-E(config)# spanning-tree stp 3 vlan 3
```

b. Add port 17 to VLAN 2, port 18 to VLAN 3.

```
RS G8124-E(config)# interface port 17
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 2
RS G8124-E(config-if)# exit

RS G8124-E(config)# interface port 18
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 3
RS G8124-E(config-if)# exit
```

VLAN 2 and VLAN 3 are removed from STG 1.

Note: In PVRST mode, each instance of STG is enabled by default.

3. Configure the following on Switch B:

   a. Add port 8 to VLAN 2. Ports 1 and 2 are by default in VLAN 1 assigned to STG 1.

   ```
   RS G8124-E(config)# vlan 2
   RS G8124-E(config-vlan)# exit
   RS G8124-E(config)# interface port 8
   RS G8124-E(config-if)# switchport mode trunk
   RS G8124-E(config-if)# switchport trunk allowed vlan 2
   RS G8124-E(config-if)# exit

   If VASA is disabled, enter the following command:
   RS G8124-E(config)# spanning-tree stp 2 vlan 2
   ```

   b. VLAN 2 is automatically removed from STG 1. By default VLAN 1 remains in STG 1.

4. Configure the following on application switch C:

   a. Add port 8 to VLAN 3. Ports 1 and 2 are by default in VLAN 1 assigned to STG 1.

   ```
   RS G8124-E(config)# vlan 3
   RS G8124-E(config-vlan)# exit
   RS G8124-E(config)# interface port 8
   RS G8124-E(config-if)# switchport mode trunk
   RS G8124-E(config-if)# switchport trunk allowed vlan 3
   RS G8124-E(config-if)# exit

   If VASA is disabled, enter the following command:
   RS G8124-E(config)# spanning-tree stp 3 vlan 3
   ```

   b. VLAN 3 is automatically removed from STG 1. By default VLAN 1 remains in STG 1.

5. Switch D does not require any special configuration for multiple Spanning Trees. Switch D uses default STG 1 only.

Rapid Spanning Tree Protocol

RSTP provides rapid convergence of the Spanning Tree and provides the fast re-configuration critical for networks carrying delay-sensitive traffic such as voice and video. RSTP significantly reduces the time to reconfigure the active topology of the network when changes occur to the physical topology or its configuration parameters. RSTP reduces the bridged-LAN topology to a single Spanning Tree.

RSTP was originally defined in IEEE 802.1w (2001) and was later incorporated into IEEE 802.1D (2004), superseding the original STP standard.

RSTP parameters apply only to Spanning Tree Group (STG) 1. The PVRST mode STGs 2-128 are not used when the switch is placed in RSTP mode.

RSTP is compatible with devices that run IEEE 802.1D (1998) Spanning Tree Protocol. If the switch detects IEEE 802.1D (1998) BPDUs, it responds with IEEE 802.1D (1998)-compatible data units. RSTP is not compatible with Per-VLAN Rapid Spanning Tree (PVRST) protocol.

Port States

RSTP port state controls are the same as for PVRST: `discarding`, `learning`, and `forwarding`.

Due to the sequence involved in these STP states, considerable delays may occur while paths are being resolved. To mitigate delays, ports defined as *edge/portfast* ports () may bypass the `discarding` and `learning` states, and enter directly into the `forwarding` state.

RSTP Configuration Guidelines

This section provides important information about configuring RSTP. When RSTP is turned on, the following occurs:

- STP parameters apply only to STG 1.
- Only STG 1 is available. All other STGs are turned off.
- All VLANs, including management VLANs, are moved to STG 1.

RSTP Configuration Example

This section provides steps to configure RSTP.

1. Configure port and VLAN membership on the switch.

2. Set the Spanning Tree mode to Rapid Spanning Tree.

```
RS G8124-E(config)# spanning-tree mode rstp
```

3. Configure RSTP parameters.

```
RS G8124-E(config)# spanning-tree stp 1 bridge priority 8192
RS G8124-E(config)# spanning-tree stp 1 bridge hello-time 5
RS G8124-E(config)# spanning-tree stp 1 bridge forward-delay 20
RS G8124-E(config)# spanning-tree stp 1 bridge maximum-age 30
RS G8124-E(config)# no spanning-tree stp 1 enable
```

4. Configure port parameters:

```
RS G8124-E(config)# interface port 3
RS G8124-E(config-if)# spanning-tree stp 1 priority 240
RS G8124-E(config-if)# spanning-tree stp 1 path-cost 500
RS G8124-E(config-if)# no spanning-tree stp 1 enable
RS G8124-E(config-if)# exit
```

Multiple Spanning Tree Protocol

Multiple Spanning Tree Protocol (MSTP) extends Rapid Spanning Tree Protocol (RSTP), allowing multiple Spanning Tree Groups (STGs) which may each include multiple VLANs. MSTP was originally defined in IEEE 802.1s (2002) and was later included in IEEE 802.1Q (2003).

In MSTP mode, the G8124-E supports up to 32 instances of Spanning Tree, corresponding to STGs 1-32, with each STG acting as an independent, simultaneous instance of RSTP.

MSTP allows frames assigned to different VLANs to follow separate paths, with each path based on an independent Spanning Tree instance. This approach provides multiple forwarding paths for data traffic, thereby enabling load-balancing, and reducing the number of Spanning Tree instances required to support a large number of VLANs.

Due to Spanning Tree's sequence of discarding, learning, and forwarding, lengthy delays may occur while paths are being resolved. Ports defined as *edge/portfast* ports () bypass the Discarding and Learning states, and enter directly into the Forwarding state.

Note: In MSTP mode, Spanning Tree for the management ports is turned off by default.

MSTP Region

A group of interconnected bridges that share the same attributes is called an MST region. Each bridge within the region must share the following attributes:

● Alphanumeric name

● Revision number

● VLAN-to STG mapping scheme

MSTP provides rapid re-configuration, scalability and control due to the support of regions, and multiple Spanning-Tree instances support within each region.

Common Internal Spanning Tree

The Common Internal Spanning Tree (CIST) or MST0 provides a common form of Spanning Tree Protocol, with one Spanning-Tree instance that can be used throughout the MSTP region. CIST allows the switch to interoperate with legacy equipment, including devices that run IEEE 802.1D (1998) STP.

CIST allows the MSTP region to act as a virtual bridge to other bridges outside of the region, and provides a single Spanning-Tree instance to interact with them.

CIST port configuration includes Hello time, path-cost, and interface priority. These parameters do not affect Spanning Tree Groups 1-32. They apply only when the CIST is used.

MSTP Configuration Guidelines

This section provides important information about configuring Multiple Spanning Tree Groups:

- When the switch initially has PVRST mode enabled and VLANs 1-127 are configured and distributed to STGs 1-127, when you turn on MSTP, the switch moves VLAN 1 and VLANs 33-128 to the CIST. When MSTP is turned off, the switch moves VLAN 1 and VLANs 33-127 from the CIST to STG 1.
- When you enable MSTP, a default revision number of 1 and a blank region name are automatically configured.

MSTP Configuration Examples

Example 1

This section provides steps to configure MSTP on the G8124-E.

1. Configure port and VLAN membership on the switch.

2. Configure Multiple Spanning Tree region parameters, and set the mode to MSTP.

```
RS G8124-E(config)# spanning-tree mst configuration
                                        (Enter MST configuration mode)
RS G8124-E(config-mst)# name <name>      (Define the Region name)
RS G8124-E(config-mst)# revision <0 – 65535>(Define the Region revision number)
RS G8124-E(config-mst)# exit
RS G8124-E(config)# spanning-tree mode mst(Set mode to Multiple Spanning Trees)
```

3. Map VLANs to MSTP instances:

```
RS G8124-E(config)# spanning-tree mst configuration
                                        (Enter MST configuration mode)
RS G8124-E(config-mst)# instance <instance ID> vlan <vlan number or range>
```

Example 2

This configuration shows how to configure MSTP Groups on the switch, as shown in Figure 12.

Figure 13. Implementing Multiple Spanning Tree Groups
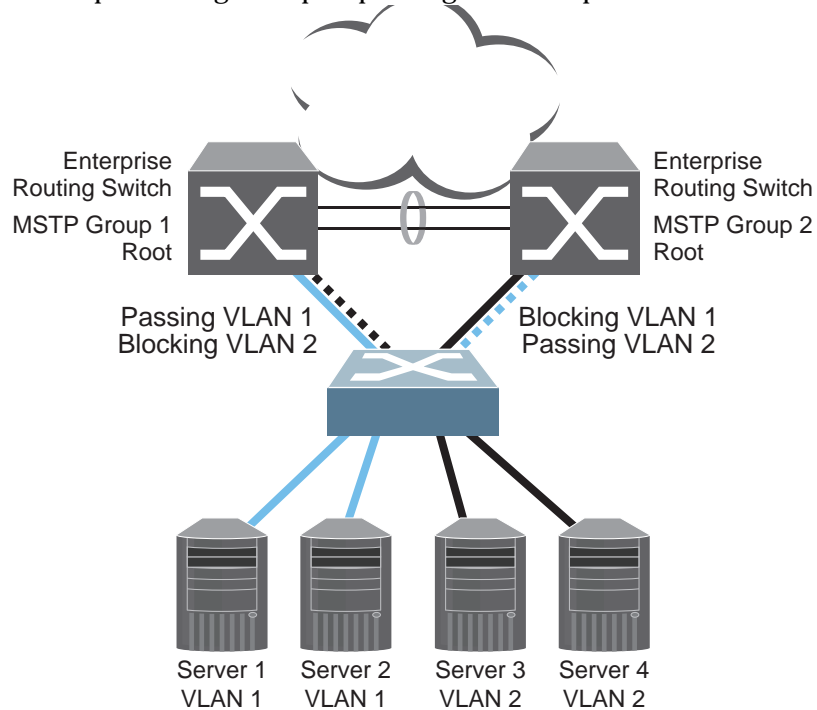
This example shows how multiple Spanning Trees can provide redundancy without wasting any uplink ports. In this example, the server ports are split between two separate VLANs. Both VLANs belong to two different MSTP groups. The Spanning Tree *priority* values are configured so that each routing switch is the root for a different MSTP instance. All of the uplinks are active, with each uplink port backing up the other.

1. Configure port membership and define the STGs for VLAN 1. Enable tagging on uplink ports that share VLANs. Port 19 and port 20 connect to the Enterprise Routing switches.

```
RS G8124-E(config)# interface port 19, 20
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# exit
```

2. Configure MSTP: Spanning Tree mode, region name, and version.

```
RS G8124-E(config)# spanning-tree mst configuration
RS G8124-E(config-mst)# name MyRegion      (Define the Region name)
RS G8124-E(config-mst)# revision 100     (Define the Revision level)
RS G8124-E(config-mst)# exit
RS G8124-E(config)# spanning-tree mode mst (Set mode to Multiple Spanning Trees)
```

3. Map VLANs to MSTP instances:

```
RS G8124-E(config)# spanning-tree mst configuration
RS G8124-E(config-mst)# instance 1 vlan 1
RS G8124-E(config-mst)# instance 2 vlan 2
```

4. Configure port membership and define the STGs for VLAN 2. Add server ports 3 and 4 to VLAN 2. Uplink ports 19 and 20 are automatically added to VLAN 2. Assign VLAN 2 to STG 2.

```
RS G8124-E(config)# interface port 3,4
RS G8124-E(config-if)# switchport access vlan 2
RS G8124-E(config-if)# exit
```

Note: Each STG is enabled by default.

Port Type and Link Type

Edge/Portfast Port

A port that does not connect to a bridge is called an *edge port*. Since edge ports are assumed to be connected to non-STP devices (such as directly to hosts or servers), they are placed in the forwarding state as soon as the link is up.

Edge ports send BPDUs to upstream STP devices like normal STP ports, but do not receive BPDUs. If a port with edge enabled does receive a BPDU, it immediately begins working as a normal (non-edge) port, and participates fully in Spanning Tree.

Use the following commands to define or clear a port as an edge port:

```
RS G8124-E(config)# interface port <port>
RS G8124-E(config-if)# [no] spanning-tree portfast
RS G8124-E(config-if)# exit
```

Note: When configuring a physical port as an STP edge port, you must shut down and reactivate ("interface port <number> shutdown" followed by "no interface port <number> shutdown") the port for the edge setting to take effect. Likewise, all the links of a port LAG or a VLAG (in both VLAG peer switches) must be shut down and reactivated before the LAG will function as an edge port.

Link Type

The link type determines how the port behaves in regard to Rapid Spanning Tree. Use the following commands to define the link type for the port:

```
RS G8124-E(config)# interface port <port>
RS G8124-E(config-if)# [no] spanning-tree link-type <type>
RS G8124-E(config-if)# exit
```

where *type* corresponds to the duplex mode of the port, as follows:

- p2p   A full-duplex link to another device (point-to-point)

- shared   A half-duplex link is a shared segment and can contain more than one device.

- auto   The switch dynamically configures the link type.

Note: Any STP port in full-duplex mode can be manually configured as a shared port when connected to a non-STP-aware shared device (such as a typical Layer 2 switch) used to interconnect multiple STP-aware devices.

Chapter 10. Virtual Link Aggregation Groups

In many data center environments, downstream servers or switches connect to upstream devices which consolidate traffic. For example, see Figure 14.

Figure 14. Typical Data Center Switching Layers with STP vs. VLAG
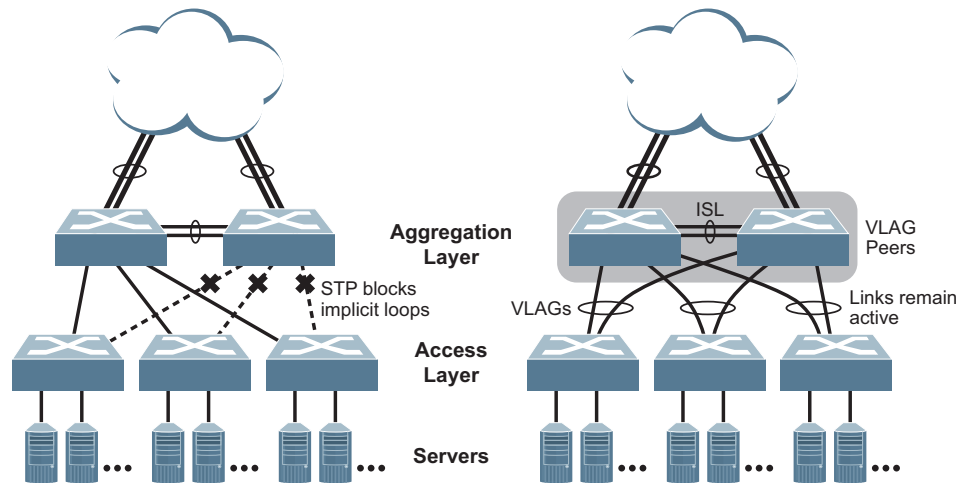
As shown in the example, a switch in the access layer may be connected to more than one switch in the aggregation layer to provide for network redundancy. Typically, Spanning Tree Protocol (RSTP, PVRST, or MSTP—see Chapter 9, "Spanning Tree Protocols") is used to prevent broadcast loops, blocking redundant uplink paths. This has the unwanted consequence of reducing the available bandwidth between the layers by as much as 50%. In addition, STP may be slow to resolve topology changes that occur during a link failure, and can result in considerable MAC address flooding.

Using Virtual Link Aggregation Groups (VLAGs), the redundant uplinks remain active, utilizing all available bandwidth.

Two switches are paired into VLAG peers, and act as a single virtual entity for the purpose of establishing a multi-port aggregation. Ports from both peers can be grouped into a VLAG and connected to the same LAG-capable target device. From the perspective of the target device, the ports connected to the VLAG peers appear to be a single LAG connecting to a single logical device. The target device uses the configured Tier ID to identify the VLAG peers as this single logical device. It is important that you use a unique Tier ID for each VLAG pair you configure. The VLAG-capable switches synchronize their logical view of the access layer port structure and internally prevent implicit loops. The VLAG topology also responds more quickly to link failure and does not result in unnecessary MAC flooding.

VLAGs are also useful in multi-layer environments for both uplink and downlink redundancy to any regular LAG-capable device. For example:

Figure 15. VLAG Application with Multiple Layers
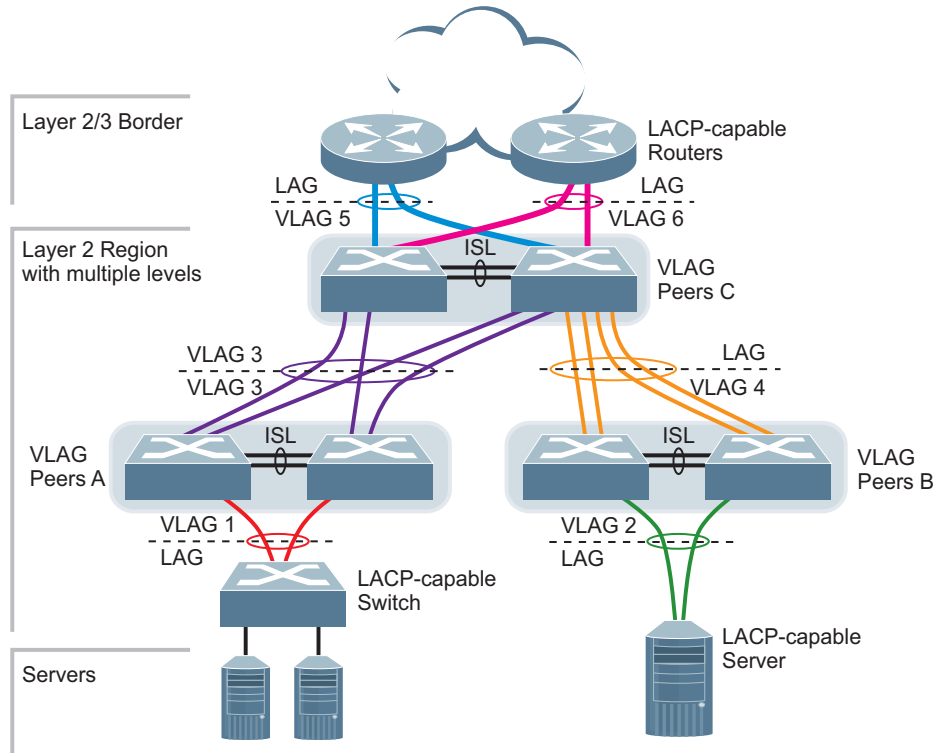
Wherever ports from *both* peered switches are aggregated to another device, the aggregated ports must be configured as a VLAG. For example, VLAGs 1 and 3 must be configured for both VLAG Peer A switches. VLAGs 2 and 4 must be configured for both VLAG Peer B switches.VLAGs 3, 5, and 6 must be configured on both VLAG Peer C switches. Other devices connecting to the VLAG peers are configured using regular static or dynamic LAGs.

Note: Do not configure a VLAG for connecting only one switch in the peer set to another device or peer set. For instance, in VLAG Peer C, a regular LAG is employed for the downlink connection to VLAG Peer B because only one of the VLAG Peer C switches is involved.

In addition, when used with VRRP, VLAGs can provide seamless active-active failover for network links. For example

Figure 16. VLAG Application with VRRP:
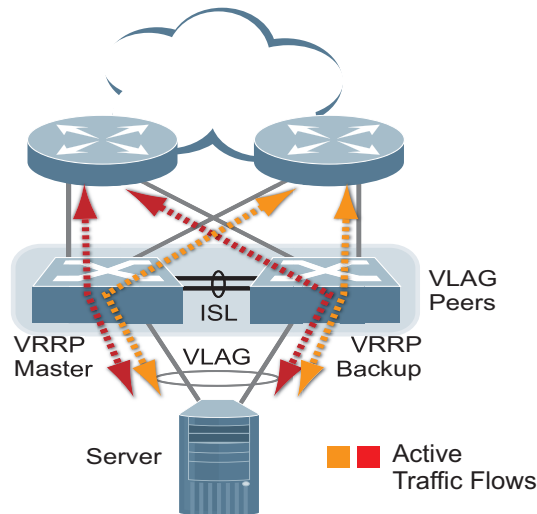
VLAG Capacities

Servers or switches that connect to the VLAG peers using a multi-port VLAG are considered VLAG clients. VLAG clients are not required to be VLAG-capable. The ports participating in the VLAG are configured as regular port LAGs on the VLAG client end.

On the VLAG peers, the VLAGs are configured similarly to regular port LAGs, using many of the same features and rules. See Chapter 8, "Ports and Link Aggregation" for general information concerning all port LAGs.

Each VLAG begins as a regular port LAG on each VLAG-peer switch. The VLAG may be either a static LAG (portchannel) or dynamic LACP LAG, and consumes one slot from the overall port LAG capacity pool. The LAG type must match that used on VLAG client devices. Additional configuration is then required to implement the VLAG on both VLAG peer switches.

You may configure up to 12 LAGs on the switch, with all types (regular or VLAG, static or LACP) sharing the same pool.

The maximum number of supported VLAG instances is as follows:

- **With STP off:** Maximum of 15 VLAG instances
- **With STP on:**
  - **PVRST/MSTP with one VLAG instance per VLAN/STG:** Maximum of 15 VLAG instances
  - **PVRST/MSTP with one VLAG instance belonging to multiple VLANs/STGs:** Maximum of 10 VLAG instances

Each LAG type can contain up to 12 member ports, depending on the port type and availability.

VLAGs versus Port LAGs

Though similar to regular port LAGs in many regards, VLAGs differ from regular port LAGs in a number of important ways:

- A VLAG can consist of multiple ports on two VLAG peers, which are connected to one logical client device such as a server, switch, or another VLAG device.

- The participating ports on the client device are configured as a regular port LAG.

- The VLAG peers must be the same model, and run the same software version.

- VLAG peers require a dedicated inter-switch link (ISL) for synchronization. The ports used to create the ISL must have the following properties:

  - ISL ports must have VLAN tagging turned on.

  - ISL ports must be configured for all VLAG VLANs.

  - ISL ports must be placed into a regular port LAG (dynamic or static).

  - A minimum of two ports on each switch are recommended for ISL use.

- Dynamic routing protocols, such as OSPF, cannot terminate on VLAGs.

- Routing over VLAGs is not supported. However, IP forwarding between subnets served by VLAGs can be accomplished using VRRP.

- VLAGs are configured using additional commands.

- It is recommended that end-devices connected to VLAG switches use NICs with dual-homing. This increases traffic efficiency, reduces ISL load, and provides faster link failover.

Configuring VLAGs

When configuring VLAG or making changes to your VLAG configuration, consider the following VLAG behavior:

- When adding a static Mrouter on VLAG links, ensure that you also add it on the ISL link to avoid VLAG link failure. If the VLAG link fails, traffic cannot be recovered through the ISL. Also ensure you add the same static entry on the peer VLAG switch for VLAG ports.

- If you have enabled VLAG on the switch, and you need to change the STP mode, ensure that you first disable VLAG and then change the STP mode.

- When VLAG is enabled, you may see two root ports on the secondary VLAG switch. One of these will be the actual root port for the secondary VLAG switch and the other will be a root port synced with the primary VLAG switch.

- The LACP key used must be unique for each VLAG in the entire topology.

- The STG to VLAN mapping on both VLAG peers must be identical.

The following parameters must be identically configured on the VLAG ports of both the VLAG peers:

- VLANs
- Native VLAN tagging
- Native VLAN/PVID
- STP mode
- BPDU Guard setting
- STP port setting
- MAC aging timers
- Static MAC entries
- ACL configuration parameters
- QoS configuration parameters

Basic VLAG Configuration

Figure 17 shows an example configuration where two VLAG peers are used for aggregating traffic from downstream devices.

Figure 17. Basic VLAGs
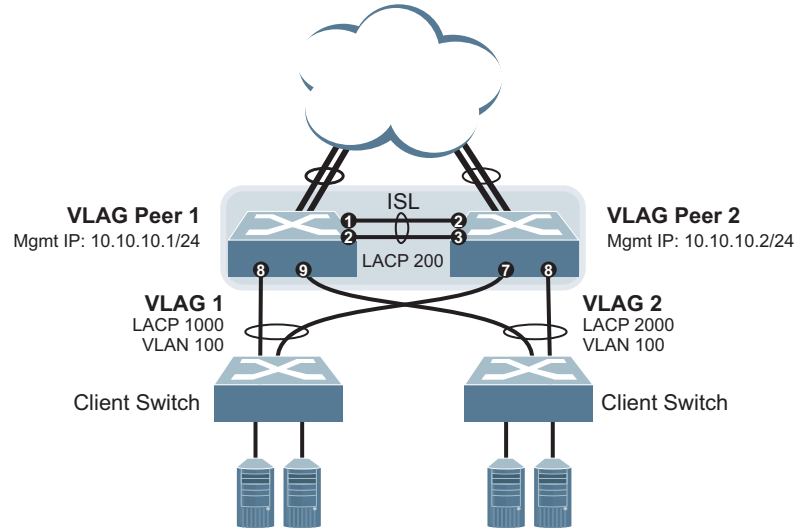
In this example, each client switch is connected to both VLAG peers. On each client switch, the ports connecting to the VLAG peers are configured as a dynamic LACP port LAG. The VLAG peer switches share a dedicated ISL for synchronizing VLAG information. On the individual VLAG peers, each port leading to a specific client switch (and part of the client switch's port LAG) is configured as a VLAG.

In the following example configuration, only the configuration for VLAG 1 on VLAG Peer 1 is shown. VLAG Peer 2 and all other VLAGs are configured in a similar fashion.

Configuring the ISL

The ISL connecting the VLAG peers is shared by all their VLAGs. The ISL needs to be configured only once on each VLAG peer.

1. Configure STP if required. Use PVRST or MSTP mode only:

```
RS G8124-E(config)# spanning-tree mode pvrst
```

2. Configure the ISL ports and place them into a port LAG:

```
RS G8124-E(config)# interface port 1-2
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 200
RS G8124-E(config-if)# exit
RS G8124-E(config)# vlag isl adminkey 200
```

Notes:

- In this case, a dynamic LAG is shown. A static LAG (portchannel) could be configured instead.

- ISL ports and VLAG ports must be members of the same VLANs.

3. Configure VLAG Tier ID. This is used to identify the VLAG switch in a multi-tier environment.

```
RS G8124-E(config)# vlag tier-id 10
```

4. Configure the ISL for the VLAG peer.

Make sure you configure the VLAG peer (VLAG Peer 2) using the same ISL aggregation type (dynamic or static), the same VLAN for VLAG and VLAG ISL ports, and the same STP mode and tier ID used on VLAG Peer 1.

Configuring the VLAG

To configure the VLAG:

1. Configure the VLAN for VLAG 1 ports. Make sure members include the ISL and VLAG 1 ports. Once the VLAN is ready, the ISL ports are automatically added to it.

```
RS G8124-E(config)# vlan 100
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 8
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 100
RS G8124-E(config-if)# exit
```

Note: In MSTP mode, VLANs are automatically mapped to CIST.

2. Place the VLAG 1 port(s) in a port LAG:

```
RS G8124-E(config)# interface port 8
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 1000
RS G8124-E(config-if)# exit
```

3. Assign the LAG to the VLAG:

```
RS G8124-E(config)# vlag adminkey 1000 enable
```

4. Continue by configuring all required VLAGs on VLAG Peer 1, and then repeat the configuration for VLAG Peer 2.

For each corresponding VLAG on the peer, the port LAG type (dynamic or static), VLAN, STP mode, and ID must be the same as on VLAG Peer 1.

5. Enable VLAG globally.

```
RS G8124-E(config)# vlag enable
```

6. Verify the completed configuration:

```
# show vlag information
```

VLAG Configuration - VLANs Mapped to MSTI

Follow the steps in this section to configure VLAG in environments where the STP mode is MSTP and no previous VLAG was configured.

Configuring the ISL

The ISL connecting the VLAG peers is shared by all their VLAGs. The ISL needs to be configured only once on each VLAG peer. Ensure you have the same region name, revision and VLAN-to-STG mapping on both VLAG switches.

1. Configure STP:

```
RS G8124-E(config)# spanning-tree mode mst
```

2. Configure the ISL ports and place them into a portchannel (dynamic or static):

```
RS G8124-E(config)# interface port 1-2
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 200
RS G8124-E(config-if)# exit
RS G8124-E(config)# vlag isl adminkey 200
```

Notes:

- In this case, a dynamic LAG is shown. A static LAG (portchannel) could be configured instead.

- ISL ports and VLAG ports must be members of the same VLANs.

3. Configure the VLAG Tier ID. This is used to identify the VLAG switch in a multi-tier environment.

```
RS G8124-E(config)# vlag tier-id 10
```

4. Configure the ISL for the VLAG peer.

   Make sure you configure the VLAG peer (VLAG Peer 2) using the same ISL aggregation type (dynamic or static), the same VLAN for VLAG and VLAG ISL ports, and the same STP mode and tier ID used on VLAG Peer 1.

Configuring the VLAG

To configure the VLAG:

1. Configure the VLAN for VLAG 1 ports. Once the VLAN s ready, the ISL ports are automatically added to it.

```
RS G8124-E(config)# vlan 100
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 8
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# exit
```

2. Map the VLAN to an MSTI.

```
RS G8124-E(config)# spanning-tree mst configuration
RS G8124-E(config-mst)# instance 1 vlan 100
```

3. Place the VLAG 1 port(s) in a LAG (static or dynamic) and assign it to the VLAG:

```
RS G8124-E(config)# interface port 8
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 1000
RS G8124-E(config-if)# exit
RS G8124-E(config)# vlag adminkey 1000 enable
```

4. Enable VLAG:

```
RS G8124-E(config)# vlag enable
```

5. Continue by configuring all required VLAGs on VLAG Peer 1, and then follow the steps for configuring VLAG Peer 2.

   For each corresponding VLAG on the peer, the port LAG type (dynamic or static), the port's VLAN, and STP mode and ID must be the same as on VLAG Peer 1.

6. Verify the completed configuration:

```
RS G8124-E# show vlag information
```

VLAGs with VRRP

Note: In a multi-layer environment, configure VRRP separately for each layer. We recommend that you configure VRRP only on the tier with uplinks. See "Configuring VLAGs in Multiple Layers" on page 182.

VRRP (see Chapter 31, "Virtual Router Redundancy Protocol") can be used in conjunction with VLAGs and LACP-capable devices to provide seamless redundancy.

Figure 18. Active-Active Configuration using VRRP and VLAGs
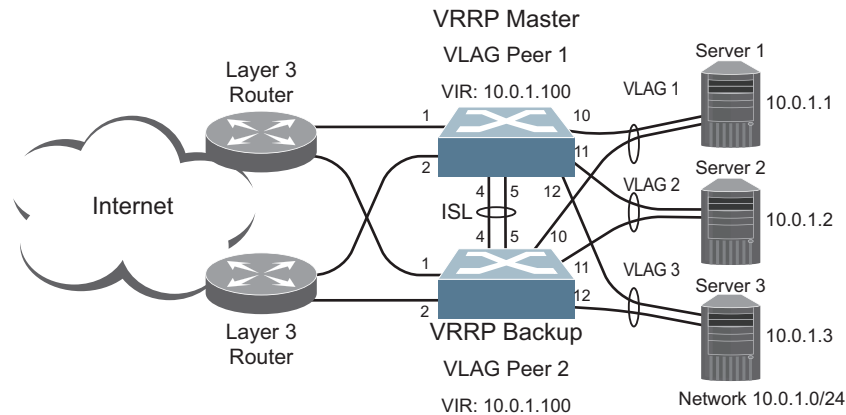
Task 1: Configure VLAG Peer 1

Note: Before enabling VLAG, you must configure the VLAG tier ID and ISL portchannel.

1. Configure VLAG tier ID

```
RS G8124-E(config)# vlag tier-id 10
```

2. Configure appropriate routing.

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area 1 area-id 0.0.0.1
RS G8124-E(config-router-ospf)# enable
RS G8124-E(config-router-ospf)# exit
```

Although OSPF is used in this example, static routing could also be deployed. For more information, see Chapter 27, "Open Shortest Path First" or Chapter 19, "Basic IP Routing."

3. Configure a server-facing interface.

```
RS G8124-E(config)# interface ip 3
RS G8124-E(config-ip-if)# ip address 10.0.1.10 255.255.255.0
RS G8124-E(config-ip-if)# vlan 100
RS G8124-E(config-ip-if)# exit
```

4. Turn on VRRP and configure the Virtual Interface Router.

```
RS G8124-E(config)# router vrrp
RS G8124-E(config-vrrp)# enable
RS G8124-E(config-vrrp)# virtual-router 1 virtual-router-id 1
RS G8124-E(config-vrrp)# virtual-router 1 interface 3
RS G8124-E(config-vrrp)# virtual-router 1 address 10.0.1.100
RS G8124-E(config-vrrp)# virtual-router 1 enable
```

5. Set the priority of Virtual Router 1 to 101, so that it becomes the Master.

```
RS G8124-E(config-vrrp)# virtual-router 1 priority 101
RS G8124-E(config-vrrp)# exit
```

6. Configure the ISL ports and place them into a port LAG:

```
RS G8124-E(config)# interface port 4-5
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 2000
RS G8124-E(config-if)# exit
```

Note: In this case, a dynamic LAG is shown. A static LAG (portchannel) could be configured instead.

7. Configure the upstream ports.

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# switchport access vlan 10
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 2
RS G8124-E(config-if)# switchport access vlan 20
RS G8124-E(config-if)# exit
```

8. Configure the server ports.

```
RS G8124-E(config)# interface port 10
RS G8124-E(config-if)# switchport access vlan 100
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 11
RS G8124-E(config-if)# switchport access vlan 100
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 12
RS G8124-E(config-if)# switchport access vlan 100
RS G8124-E(config-if)# exit
```

9. Configure all VLANs including VLANs for the VLAGs.

```
RS G8124-E(config)# vlan 10
RS G8124-E(config-vlan)# exit

RS G8124-E(config)# vlan 20
RS G8124-E(config-vlan)# exit

RS G8124-E(config)# vlan 100
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 4-5
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 100
RS G8124-E(config-if)# exit
```

10. Configure Internet-facing interfaces.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip address 172.1.1.10 255.255.255.0
RS G8124-E(config-ip-if)# vlan 10
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# ip ospf area 1
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip address 172.1.3.10 255.255.255.0
RS G8124-E(config-ip-if)# vlan 20
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# ip ospf area 1
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
```

11. Place the VLAG port(s) in their port LAGs.

```
RS G8124-E(config)# interface port 10
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 1000
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 11
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 1100
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 12
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 1200
RS G8124-E(config-if)# exit
```

12. Assign the LAGs to the VLAGs:

```
RS G8124-E(config)# vlag adminkey 1000 enable
RS G8124-E(config)# vlag adminkey 1100 enable
RS G8124-E(config)# vlag adminkey 1200 enable
```

13. Globally enable VLAG

```
RS G8124-E(config)# vlag enable
```

14. Verify the completed configuration:

```
# show vlag information
```

Task 2: Configure VLAG Peer 2

The VLAG peer (VLAG Peer 2) must be configured using the same ISL aggregation type (dynamic or static), the same VLAN for VLAG and VLAG ISL ports, and the same STP mode and Tier ID used on VLAG Switch 1.

For each corresponding VLAG on the peer, the port LAG type (dynamic or static), VLAN, and STP mode and ID must be the same as on VLAG Switch 1.

1. Configure VLAG tier ID and enable VLAG globally.

```
RS G8124-E(config)# vlag tier-id 10
RS G8124-E(config)# vlag enable
```

2. Configure appropriate routing.

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area 1 area-id 0.0.0.1
RS G8124-E(config-router-ospf)# enable
RS G8124-E(config-router-ospf)# exit
```

Although OSPF is used in this example, static routing could also be deployed.

3. Configure a server-facing interface.

```
RS G8124-E(config)# interface ip 3
RS G8124-E(config-ip-if)# ip address 10.0.1.11 255.255.255.0
RS G8124-E(config-ip-if)# vlan 100
RS G8124-E(config-ip-if)# exit
```

4. Turn on VRRP and configure the Virtual Interface Router.

```
RS G8124-E(config)# router vrrp
RS G8124-E(config-vrrp)# enable
RS G8124-E(config-vrrp)# virtual-router 1 virtual-router-id 1
RS G8124-E(config-vrrp)# virtual-router 1 interface 3
RS G8124-E(config-vrrp)# virtual-router 1 address 10.0.1.100
RS G8124-E(config-vrrp)# virtual-router 1 enable
```

5. Configure the ISL ports and place them into a port LAG:

```
RS G8124-E(config)# interface port 4-5
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 2000
RS G8124-E(config-if)# exit
```

6. Configure the upstream ports.

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# switchport access vlan 30
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 2
RS G8124-E(config-if)# switchport access vlan 40
RS G8124-E(config-if)# exit
```

7. Configure the server ports.

```
RS G8124-E(config)# interface port 10
RS G8124-E(config-if)# switchport access vlan 100
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 11
RS G8124-E(config-if)# switchport access vlan 100
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 12
RS G8124-E(config-if)# switchport access vlan 100
RS G8124-E(config-if)# exit
```

8. Configure all VLANs including VLANs for the VLAGs.

```
RS G8124-E(config)# vlan 30
RS G8124-E(config-vlan)# exit

RS G8124-E(config)# vlan 40
RS G8124-E(config-vlan)# exit

RS G8124-E(config)# vlan 100
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 4-5
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 100
RS G8124-E(config-if)# exit
```

9. Configure Internet-facing interfaces.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip address 172.1.2.11 255.255.255.0
RS G8124-E(config-ip-if)# vlan 30
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# ip ospf area 1
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip address 172.1.4.12 255.255.255.0
RS G8124-E(config-ip-if)# vlan 40
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# ip ospf area 1
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
```

10. Place the VLAG port(s) in their port LAGs.

```
RS G8124-E(config)# interface port 10
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 1000
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 11
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 1100
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 12
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# lacp key 1200
RS G8124-E(config-if)# exit
```

11. Assign the LAGs to the VLAGs:

```
RS G8124-E(config)# vlag adminkey 1000 enable
RS G8124-E(config)# vlag adminkey 1100 enable
RS G8124-E(config)# vlag adminkey 1200 enable
```

12. Verify the completed configuration:

```
# show vlag information
```

Configuring VLAGs in Multiple Layers

Figure 19. VLAG in Multiple Layers
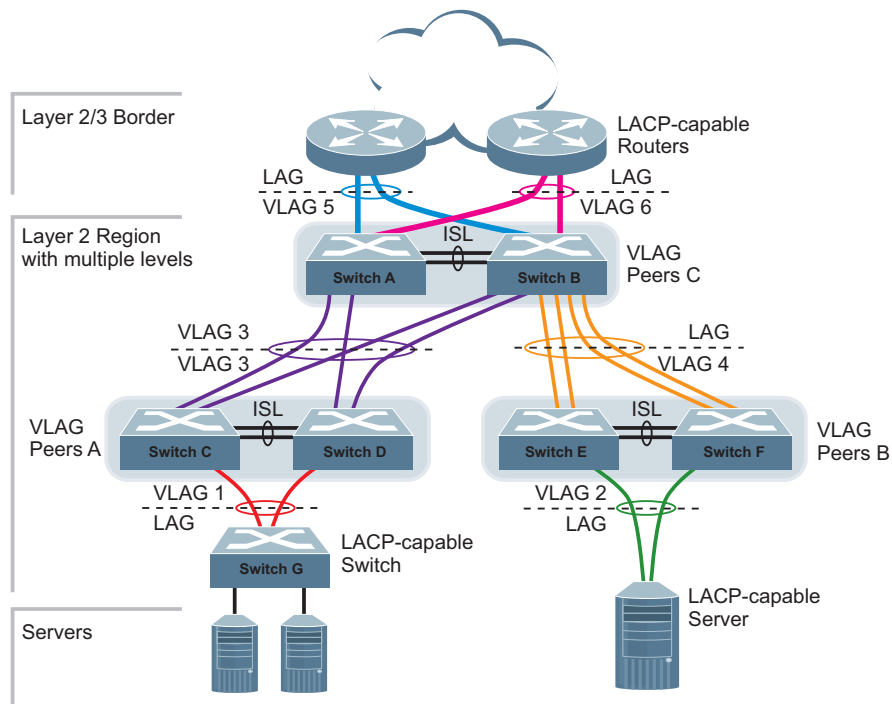
Figure 19 shows an example of VLAG being used in a multi-layer environment. Following are the configuration steps for the topology.

Task 1: Configure Layer 2/3 border switches.

Configure ports on border switch as follows:

```
RS G8124-E(config)# interface port 1,2
RS G8124-E(config-if)# lacp key 100
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# exit
```

Repeat the previous steps for the second border switch.

Task 2: Configure switches in the Layer 2 region.

Consider the following:

- ISL ports on switches A and B - ports 1, 2
- Ports connecting to Layer 2/3 - ports 5, 6
- Ports on switches A and B connecting to switches C and D: ports 10, 11
- Ports on switch B connecting to switch E: ports 15, 16
- Ports on switch B connecting to switch F: ports 17, 18

1. Configure VLAG tier ID and enable VLAG globally.

```
RS G8124-E(config)# vlag tier-id 10
RS G8124-E(config)# vlag enable
```

2. Configure ISL ports on Switch A.

```
RS G8124-E(config)# interface port 1,2
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# lacp key 200
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# exit

RS G8124-E(config)# vlag isl adminkey 200
RS G8124-E(config-vlan)# exit
```

3. Configure port on Switch A connecting to Layer 2/3 router 1.

```
RS G8124-E(config)# vlan 10
 VLAN number 10 with name "VLAN 10" created
 VLAN 10 was assigned to STG 10
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 1,2,5
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 10
RS G8124-E(config-if)# exit

RS G8124-E(config)# interface port 5
RS G8124-E(config-if)# lacp key 400
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# exit

RS G8124-E(config)# vlag adminkey 400 enable
```

Repeat the previous steps on Switch B for ports connecting to Layer 2/3 router 1.

4. Configure port on Switch A connecting to Layer 2/3 router 2.

```
RS G8124-E(config)# vlan 20
 VLAN number 20 with name "VLAN 20" created
 VLAN 20 was assigned to STG 20
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 1,2,6
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 20
RS G8124-E(config-if)# exit

RS G8124-E(config)# interface port 6
RS G8124-E(config-if)# lacp key 500
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# exit

RS G8124-E(config)# vlag adminkey 500 enable
```

Repeat these commands on Switch B for ports connecting to Layer 2/3 router 2.

5. Configure ports on Switch A connecting to downstream VLAG switches C and D.

```
RS G8124-E(config)# vlan 20
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 10,11
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 20
RS G8124-E(config-if)# lacp key 600
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# exit

RS G8124-E(config)# vlag adminkey 600 enable
```

Repeat these commands on Switch B for ports connecting to downstream VLAG switch C and D.

6. Configure ports on Switch B connecting to downstream switches E and F.

```
RS G8124-E(config)# vlan 30
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 15-18
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 30
RS G8124-E(config-if)# lacp key 700
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# exit
```

7. Configure ISL between switches C and D, and between E and F as shown in Step 1.

8. Configure the Switch G as shown in Step 2.

VLAG with PIM

Note: This section is applicable only to G8124-E.

Protocol Independent Multicast (PIM) is designed for efficiently routing multicast traffic across one or more IPv4 domains. PIM is used by multicast source stations, client receivers, and intermediary routers and switches, to build and maintain efficient multicast routing trees. PIM is protocol independent; It collects routing information using the existing unicast routing functions underlying the IPv4 network, but does not rely on any particular unicast protocol. For PIM to function, a Layer 3 routing protocol (such as BGP, OSPF, RIP, or static routes) must first be configured on the switch.

Lenovo Networking OS supports PIM in Sparse Mode (PIM-SM) and Dense Mode (PIM-DM). For more details on PIM, see Chapter 28, "Protocol Independent Multicast."

PIM, when configured in a VLAG topology, provides efficient multicast routing along with redundancy and failover. When the multicast source is located in the core L3 network, only the primary VLAG switch forwards multicast data packets to avoid duplicate packets reaching the access layer switch. The secondary VLAG switch is available as backup and forwards packets only when the primary VLAG switch is not available and during failover. When the multicast source is located in the L2 domain, behind the VLAG ports, either the primary or the secondary switch will forward the data traffic to the receiver, based on the shortest path detected by PIM.

See Figure 17 on page 172 for a basic VLAG topology. For PIM to function in a VLAG topology, the following are required:

- IGMP (v1 or v2) must be configured on the VLAG switches.

- A Layer 3 routing protocol (such as BGP, OSPF, RIP, or static routes) must be globally enabled and on VLAG-associated IP interfaces for multicast routing.

- The VLAG switches must be connected to upstream multicast routers.

- The Rendezvous Point (RP) and/or the Bootstrap router (BSR) must be configured on the upstream router.

- The multicast sources must be connected to the upstream router.

- Flooding must be disabled on the VLAG switches or in the VLAN associated with the VLAG ports.

- ISL ports must be members of VLANs that have VLAG ports as members.

For PIM configuration steps and commands, see "PIM Configuration Examples" on page 417.

Traffic Forwarding

In a VLAG with PIM topology, traffic forwarded by the upstream router is managed as follows when the multicast source is located in the core L3 network and the receiver is located in the L2 network:

- If the primary and secondary VLAG ports are up, the primary switch forwards traffic to the receiver. The secondary switch blocks the traffic. Multicast entries are created on both the VLAG switches: primary VLAG switch with forward state; secondary VLAG switch with pruned state.

- If the primary VLAG port fails, the secondary VLAG switch forwards traffic to the receiver. Multicast entries are created on both the VLAG switches: primary VLAG switch with forward state; secondary VLAG switch with VLAG pruned state.

- If the secondary VLAG port fails, the primary VLAG switch forwards traffic to the receiver. Multicast entries are created on both the VLAG switches: primary VLAG switch with forward state; secondary VLAG switch with pruned state.

- If the primary VLAG switch is down, the secondary VLAG switch forwards traffic to the receiver. When the primary VLAG switch boots up again, it becomes the secondary VLAG switch and blocks traffic to the receiver. The VLAG switch that was secondary initially becomes the primary and continues forwarding traffic to the receiver.

- If the secondary VLAG switch is down, the primary VLAG switch forwards traffic to the receiver. When the secondary VLAG switch is up, it blocks traffic. The primary switch forwards traffic to the receiver.

- If the uplink to the primary VLAG switch is down, the secondary VLAG switch forwards traffic to the receiver and to the primary VLAG switch over the ISL. The primary VLAG switch blocks traffic to the receiver so the receiver does not get double traffic. Both the VLAG switches will have multicast entries in forward state.

- If the uplink to the secondary VLAG switch is down, the primary VLAG switch forwards traffic to the receiver and to the secondary VLAG switch over the ISL. The secondary VLAG switch blocks traffic to the receiver so the receiver does not get double traffic. Both the VLAG switches will have multicast entries in the forward state.

When the multicast source is connected to VLAG ports (layer 2 domain), traffic forwarded by the VLAG routers is managed as follows:

- IPMC traffic from the access switch can be hashed to any of the VLAG switches. Consequently, both the primary and secondary VLAG switches must synchronize the (S,G) entries for faster failover.

- The Rendezvous Point sends (S,G) entries to either the primary or secondary VLAG switch, depending on which provides the shortest path to the source. However, Register-Stop messages are only sent to the primary VLAG switch. Based on the shortest path, one of the VLAG switches will forward traffic for a particular (S,G) entry to the receiver.

- For the VLAG multi-tier topology, an additional L3 backup path to ISL is supported. On the L3 backup interface, both L3 routing protocols and PIM must be enabled.

Health Check

In a VLAG with PIM topology, you must configure health check. See "Health Check" on page 186.

When health check is configured, and the ISL is down, the primary VLAG switch forwards traffic to the receiver. The secondary VLAG switch ports will be errdisable state and will block traffic to the receiver.

VLAG with IGMPv3

Consider the following when using VLAG with IGMPv3:

- To maintain synchronization of the groups and sources states on the VLAG peers, traffic received on vLAG trunks is processed by the primary and then forwarded to the secondary. The secondary forwards the packets to the peer when it receives traffic on a vLAG and processes them when received from the peer.

- Traffic received on non-vLAG ports but on an ISL VLAN is first processed by the receiving vLAG switch and then forwarded to the peer, which will synchronize them on the ISL trunk.

- The querier enabled option is supported. After querier election, when one of the peers is elected as querier, it will generate queries when required.

- When the secondary comes up after a reboot or ISL and health-check fail, the information on the primary will be considered valid and synchronized with the secondary.

- PIM and IGMPv3 are not supported to work together on VLAG.

- When configuring VLAG with IGMPv3, the same version of IGMP needs to be set on both VLAG switches.

- The number of IGMPv3 groups with VLAG depends on the number of included and excluded sources.

- The total number of supported groups is the same as the number of groups the switch can learn with VLAG disabled.

All other scenarios will have a behavior similar to IGMPv2 on VLAG.

Chapter 11. Quality of Service

Quality of Service features allow you to allocate network resources to mission-critical applications at the expense of applications that are less sensitive to such factors as time delays or network congestion. You can configure your network to prioritize specific types of traffic, ensuring that each type receives the appropriate Quality of Service (QoS) level.

The following topics are discussed in this section:

QoS Overview

QoS helps you allocate guaranteed bandwidth to the critical applications, and limit bandwidth for less critical applications. Applications such as video and voice must have a certain amount of bandwidth to work correctly; using QoS, you can provide that bandwidth when necessary. Also, you can put a high priority on applications that are sensitive to timing out or that cannot tolerate delay, by assigning their traffic to a high-priority queue.

By assigning QoS levels to traffic flows on your network, you can ensure that network resources are allocated where they are needed most. QoS features allow you to prioritize network traffic, thereby providing better service for selected applications.

Figure 20 shows the basic QoS model used by the switch.

Figure 20. QoS Model
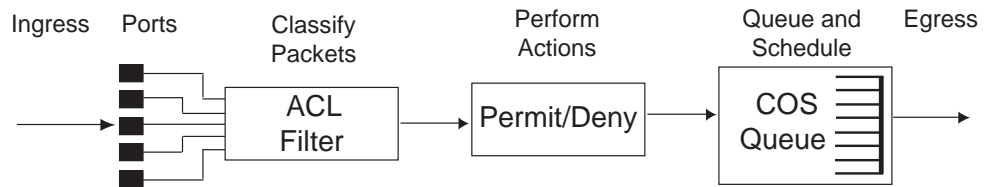
The basic QoS model works as follows:

- Classify traffic:
    - Read DSCP value.
    - Read 802.1p priority value.
    - Match ACL filter parameters.
- Perform actions:
    - Define bandwidth and burst parameters
    - Select actions to perform on in-profile and out-of-profile traffic
    - Deny packets
    - Permit packets
    - Mark DSCP or 802.1p Priority
    - Set COS queue (with or without re-marking)
- Queue and schedule traffic:
    - Place packets in one of the COS queues.
    - Schedule transmission based on the COS queue.

Using ACL Filters

Access Control Lists (ACLs) are filters that allow you to classify and segment traffic, so you can provide different levels of service to different traffic types. Each filter defines the conditions that must match for inclusion in the filter, and also the actions that are performed when a match is made.

Lenovo Networking OS 8.3 supports up to 127 ACLs when the switch is operating in the Balanced deployment mode (see "Available Profiles" on page 206). ACL menus and commands are not available in the Routing deployment mode.

The G8124-E allows you to classify packets based on various parameters. For example:

- Ethernet: source MAC, destination MAC, VLAN number/mask, Ethernet type, priority.
- IPv4: Source IP address/mask, destination address/mask, type of service, IP protocol number.
- TCP/UPD: Source port, destination port, TCP flag.
- Packet format

For ACL details, see Chapter 6, "Access Control Lists."

Summary of ACL Actions

Actions determine how the traffic is treated. The G8124-E QoS actions include the following:

- Pass or Drop
- Re-mark a new DiffServ Code Point (DSCP)
- Re-mark the 802.1p field
- Set the COS queue

ACL Metering and Re-Marking

You can define a profile for the aggregate traffic flowing through the G8124-E by configuring a QoS meter (if desired) and assigning ACLs to ports. When you add ACLs to a port, make sure they are ordered correctly in terms of precedence.

Actions taken by an ACL are called *In-Profile* actions. You can configure additional In-Profile and Out-of-Profile actions on a port. Data traffic can be metered, and re-marked to ensure that the traffic flow provides certain levels of service in terms of bandwidth for different types of network traffic.

Metering

QoS metering provides different levels of service to data streams through user-configurable parameters. A meter is used to measure the traffic stream against a traffic profile, which you create. Thus, creating meters yields In-Profile and Out-of-Profile traffic for each ACL, as follows:

● **In-Profile**–If there is no meter configured or if the packet conforms to the meter, the packet is classified as In-Profile.

● **Out-of-Profile**–If a meter is configured and the packet does not conform to the meter (exceeds the committed rate or maximum burst rate of the meter), the packet is classified as Out-of-Profile.

Using meters, you set a Committed Rate in Kbps (multiples of 64 Mbps). All traffic within this Committed Rate is In-Profile. Additionally, you set a Maximum Burst Size that specifies an allowed data burst larger than the Committed Rate for a brief period. These parameters define the In-Profile traffic.

Meters keep the sorted packets within certain parameters. You can configure a meter on an ACL, and perform actions on metered traffic, such as packet re-marking.

Re-Marking

Re-marking allows for the treatment of packets to be reset based on new network specifications or desired levels of service. You can configure the ACL to re-mark a packet as follows:

- Change the DSCP value of a packet, used to specify the service level traffic receives.

- Change the 802.1p priority of a packet.

The following example includes steps to configure a meter and out-of-profile DSCP remarking:

1. Create an ACL.

```
RS G8124-E(config)# access-control list 2 tcp-udp destination-port 80
```

2. Configure the meter with a committed rate.

```
RS G8124-E(config)# access-control list 2 meter committed-rate 10000000
RS G8124-E(config)# access-control list 2 meter enable
```

3. Set out-of-profile action to re-mark DSCP value based on the global DSCP mapping. The DSCP remarking value is global for all ACLs that have DSCP re-marking enabled.

```
RS G8124-E(config)# access-control list out-profile dscp 5 mkdn-dscp 0
```

4. Enable DSCP re-marking for the ACL.

```
RS G8124-E(config)# access-control list 2 re-mark out-profile dscp enable
```

Using DSCP Values to Provide QoS

The switch uses the Differentiated Services (DiffServ) architecture to provide QoS functions. DiffServ is described in IETF RFCs 2474 and 2475.

The six most significant bits in the TOS byte of the IP header are defined as DiffServ Code Points (DSCP). Packets are marked with a certain value depending on the type of treatment the packet must receive in the network device. DSCP is a measure of the Quality of Service (QoS) level of the packet.

The switch can classify traffic by reading the DiffServ Code Point (DSCP) or IEEE 802.1p priority value, or by using filters to match specific criteria. When network traffic attributes match those specified in a traffic pattern, the policy instructs the switch to perform specified actions on each packet that passes through it. The packets are assigned to different Class of Service (COS) queues and scheduled for transmission.

Differentiated Services Concepts

To differentiate between traffic flows, packets can be classified by their DSCP value. The Differentiated Services (DS) field in the IP header is an octet, and the first six bits, called the DS Code Point (DSCP), can provide QoS functions. Each packet carries its own QoS state in the DSCP. There are 64 possible DSCP values (0-63).

Figure 21. Layer 3 IPv4 packet

Version Length	ToS	Length	ID	Offset	TTL	Proto	FCS	SIP	DIP	Data

Differentiated Services Code Point (DSCP)						unused	
7	6	5	4	3	2	1	0
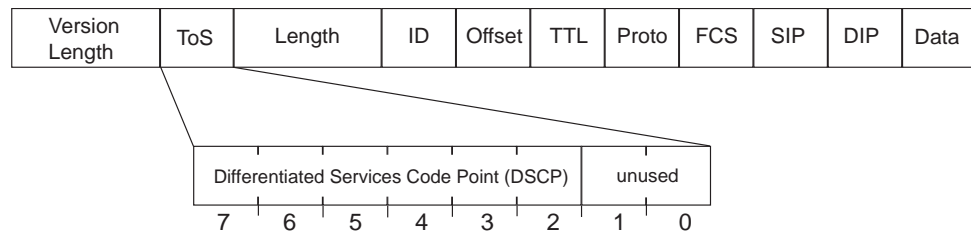
The switch can perform the following actions to the DSCP:

- Read the DSCP value of ingress packets.
- Re-mark the DSCP value to a new value
- Map the DSCP value to a Class of Service queue (COSq).

The switch can use the DSCP value to direct traffic prioritization.

With DiffServ, you can establish policies to direct traffic. A policy is a traffic-controlling mechanism that monitors the characteristics of the traffic, (for example, its source, destination, and protocol) and performs a controlling action on the traffic when certain characteristics are matched.

Trusted/Untrusted Ports

By default, all ports on the G8124-E are trusted. To configure untrusted ports, re-mark the DSCP value of the incoming packet to a lower DSCP value using the following commands:

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# dscp-marking
RS G8124-E(config-if)# exit
RS G8124-E(config)# qos dscp dscp-mapping <DSCP value (0-63)> <new value>
RS G8124-E(config)# qos dscp re-marking
```

Per Hop Behavior

The DSCP value determines the Per Hop Behavior (PHB) of each packet. The PHB is the forwarding treatment given to packets at each hop. QoS policies are built by applying a set of rules to packets, based on the DSCP value, as they hop through the network.

The default settings are based on the following standard PHBs, as defined in the IEEE standards:

● Expedited Forwarding (EF)—This PHB has the highest egress priority and lowest drop precedence level. EF traffic is forwarded ahead of all other traffic. EF PHB is described in RFC 2598.

● Assured Forwarding (AF)—This PHB contains four service levels, each with a different drop precedence, as shown in the following table. Routers use drop precedence to determine which packets to discard last when the network becomes congested. AF PHB is described in RFC 2597.

Drop Precedence	Class 1	Class 2	Class 3	Class 4
Low	AF11 (DSCP 10)	AF21 (DSCP 18)	AF31 (DSCP 26)	AF41 (DSCP 34)
Medium	AF12 (DSCP 12)	AF22 (DSCP 20)	AF32 (DSCP 28)	AF42 (DSCP 36)
High	AF13 (DSCP 14)	AF23 (DSCP 22)	AF33 (DSCP 30)	AF43 (DSCP 38)

● Class Selector (CS)—This PHB has eight priority classes, with CS7 representing the highest priority, and CS0 representing the lowest priority, as shown in the following table. CS PHB is described in RFC 2474.

Priority	Class Selector	DSCP
Highest	CS7	56
	CS6	48
	CS5	40
	CS4	32
	CS3	24
	CS2	16
	CS1	8
Lowest	CS0	0

QoS Levels

Table 16 shows the default service levels provided by the switch, listed from highest to lowest importance:

Table 16. *Default QoS Service Levels*

Service Level	Default PHB	802.1p Priority
Critical	CS7	7
Network Control	CS6	6
Premium	EF, CS5	5
Platinum	AF41, AF42, AF43, CS4	4
Gold	AF31, AF32, AF33, CS3	3
Silver	AF21, AF22, AF23, CS2	2
Bronze	AF11, AF12, AF13, CS1	1
Standard	DF, CS0	0

DSCP Re-Marking and Mapping

The switch can use the DSCP value of ingress packets to re-mark the DSCP to a new value, and to set an 802.1p priority value. Use the following command to view the default settings.

```
RS G8124-E# show qos dscp
Current DSCP Remarking Configuration: OFF

  DSCP     New DSCP  New 802.1p Prio
--------  --------  ---------------
    0         0          0
    1         1          0
    2         2          0
    3         3          0
    4         4          0
    5         5          0
    6         6          0
    7         7          0
    8         8          1
    9         9          0
   10        10          1
...
   54        54          0
   55        55          0
   56        56          7
   57        57          0
   58        58          0
   59        59          0
   60        60          0
   61        61          0
   62        62          0
   63        63          0
```

Use the following command to turn on DSCP re-marking globally:

```
RS G8124-E(config)# qos dscp re-marking
```

Then you must enable DSCP re-marking on any port that you wish to perform this function (Interface Port mode).

Note: If an ACL meter is configured for DSCP re-marking, the meter function takes precedence over QoS re-marking.

DSCP Re-Marking Configuration Examples

Example 1

The following example includes the basic steps for re-marking DSCP value and mapping DSCP value to 802.1p.

1. Turn DSCP re-marking on globally, and define the DSCP-DSCP-802.1p mapping. You can use the default mapping.

```
RS G8124-E(config)# qos dscp re-marking
RS G8124-E(config)# qos dscp dscp-mapping <DSCP value (0-63)> <new value>
RS G8124-E(config)# qos dscp dot1p-mapping <DSCP value (0-63)> <802.1p value>
```

2. Enable DSCP re-marking on a port.

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# qos dscp re-marking
RS G8124-E(config-if)# exit
```

Example 2

The following example assigns strict priority to VoIP traffic and a lower priority to all other traffic.

1. Create an ACL to re-mark DSCP value and COS queue for all VoIP packets.

```
RS G8124-E(config)# access-control list 2 tcp-udp source-port 5060 0xffff
RS G8124-E(config)# access-control list 2 meter committed-rate 10000000
RS G8124-E(config)# access-control list 2 meter enable
RS G8124-E(config)# access-control list 2 re-mark in-profile dscp 56
RS G8124-E(config)# access-control list 2 re-mark dot1p 7
RS G8124-E(config)# access-control list 2 action permit
```

2. Create an ACL to set a low priority to all other traffic.

```
RS G8124-E(config)# access-control list 3 action set-priority 1
RS G8124-E(config)# access-control list 3 action permit
```

3. Apply the ACLs to a port and enable DSCP marking.

```
RS G8124-E(config)# interface port 5
RS G8124-E(config-if)# access-control list 2
RS G8124-E(config-if)# access-control list 3 ethernet source-mac-address
00:00:00:00:00:00 00:00:00:00:00:00
RS G8124-E(config-if)# dscp-marking
RS G8124-E(config-if)# exit
```

4. Enable DSCP re-marking globally.

```
RS G8124-E(config)# qos dscp re-marking
```

5. Assign the DSCP re-mark value.

```
RS G8124-E(config)# qos dscp dscp-mapping 40 9
RS G8124-E(config)# qos dscp dscp-mapping 46 9
```

6. Assign strict priority to VoIP COS queue.

```
RS G8124-E(config)# qos transmit-queue weight-cos 7 0
```

7. Map priority value to COS queue for non-VoIP traffic.

```
RS G8124-E(config)# qos transmit-queue mapping 1 1
```

8. Assign weight to the non-VoIP COS queue.

```
RS G8124-E(config)# qos transmit-queue weight-cos 1 2
```

Using 802.1p Priority to Provide QoS

The G8124-E provides Quality of Service functions based on the priority bits in a packet's VLAN header. (The priority bits are defined by the 802.1p standard within the IEEE 802.1Q VLAN header.) The 802.1p bits, if present in the packet, specify the priority to be given to packets during forwarding. Packets with a numerically higher (non-zero) priority are given forwarding preference over packets with lower priority value.

The IEEE 802.1p standard uses eight levels of priority (0-7). Priority 7 is assigned to highest priority network traffic, such as OSPF or RIP routing table updates, priorities 5-6 are assigned to delay-sensitive applications such as voice and video, and lower priorities are assigned to standard applications. A value of 0 (zero) indicates a "best effort" traffic prioritization, and this is the default when traffic priority has not been configured on your network. The switch can filter packets based on the 802.1p values.

Figure 22. Layer 2 802.1q/802.1p VLAN tagged packet

Preamble	SFD	DMAC	SMAC	Tag	E Type	Data	FCS

Priority		VLAN Identifier (VID)
7 6 5	4	3 2 1 0 7 6 5 4 3 2 1 0
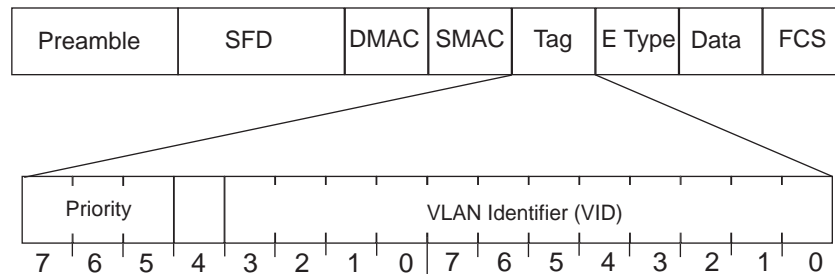
Ingress packets receive a priority value, as follows:

- **Tagged packets**—switch reads the 802.1p priority in the VLAN tag.

- **Untagged packets**—switch tags the packet and assigns an 802.1p priority value, based on the port's default 802.1p priority.

Egress packets are placed in a COS queue based on the priority value, and scheduled for transmission based on the COS queue number. Higher COS queue numbers provide forwarding precedence

To configure a port's default 802.1p priority value, use the following commands:

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# dot1p <802.1p value (0-7)>
RS G8124-E(config-if)# exit
```

Queuing and Scheduling

The G8124-E has 8 output Class of Service (COS) queues per port. If CEE is enabled, this is changed to 3 queues per port and ETS is then used to configure the scheduling in a manner different than what is described in this section. Each packet's 802.1p priority determines its COS queue, except when an ACL action sets the COS queue of the packet.

Note: When vNIC operations are enabled, the total number of COS queues available is 4.

You can configure the following attributes for COS queues:

● Map 802.1p priority value to a COS queue

● Define the scheduling weight of each COS queue

You can map 802.1p priority value to a COS queue, as follows:

```
RS G8124-E(config)# qos transmit-queue mapping <802.1p priority value (0-7)>
    <COS queue (0-7)>
```

To set the COS queue scheduling weight, use the following command:

```
RS G8124-E(config)# qos transmit-queue weight-cos <COSq number>
    <COSq weight (0-15)>
```

To achieve a balanced bandwidth allocation among the various priority groups, packets are scheduled according to a weighted deficit round-robin (WDRR) algorithm. WDRR is aware of packet sizes, which can vary significantly in a CEE environment, making WDRR more suitable than a regular weighted round-robin (WRR) method, which selects groups based only on packet counts. In addition, forwarding precedence is determined by COS queue number, with higher-numberd queues given higher precedence.

Note: For setting traffic ratios, COSq weight values are internally rounded up to 2, 4, 8, or 16.

Part 4: Advanced Switching Features

Chapter 12. Deployment Profiles

The Lenovo Networking OS software for the RackSwitch G8124-E can be configured to operate in different modes for different deployment scenarios. Each deployment profile sets different capacity levels for basic switch resources, such as the number of IP routes and ARP entries, to optimize the switch for different types of networks.

This chapter covers the following topics

- "Available Profiles" on page 206
- "Selecting Profiles" on page 207
- "Automatic Configuration Changes" on page 208

Available Profiles

The following deployment profiles are currently available on the G8124-E:

- Default Profile—This profile is recommended for general network usage. Switch resources are balanced to provide moderate capacity for IP routes, ARP entries, ACLs, and VMAPs.

- Aggregation Profile—This special deployment profile emphasizes ARP and routing. ACLs are unsupported in this mode.

- IPv6 Profile—For extended IPv6 applications.

- Routing Profile—This is a special deployment profile. It is recommended when additional IP routes are required on the switch. To provide the additional IP routes, the number of ARP entries is reduced, and the ACL, VMAP, and IGMP snooping features are unsupported.

- High-Frequency Traders (HFT) Profile—This special deployment profile emphasizes IPv4 multicast route entries and provides a moderate number of unicast routes. ARP entries and dynamic routes are reduced, and IPv6, VMready, DCBX/FCoE features are not supported in this mode.

The properties of each mode are compared in the following table.

Table 17. *Deployment Mode Comparison*

Switch Feature	Capacity, by Mode				
	Default	**Aggregation**	**IPv6**	**Routing**	**HFT**
ACLs	127	*Unsupported*	127	*Unsupported*	127
ARP entries	2,048	8,000	2,048	*1,000*	512
Dynamic routes	2,684	*2,000*	2,807	9,000	1024
IGMP Snooping	Available		*Unsupported*	*Unsupported*	2900
VM Policy Bandwidth Control	Available	*Unsupported*	*Unsupported*	*Unsupported*	*Unsupported*
VMAPs	127	*Unsupported*	*Unsupported*	*Unsupported*	*Unsupported*
VMready	Available	Available	*Unsupported*	*Unsupported*	*Unsupported*

Note: Throughout this guide, where feature capacities are listed, values reflect those of the Default profile only, unless otherwise noted.

Selecting Profiles

To change the deployment profile, the new profile must first be selected, and the switch must then be rebooted to use the new profile.

Note: Before changing profiles, it is recommended that you save the active switch configuration to a backup file so that it may be restored later if desired.

The following ISCLI commands are used to change the deployment profile:

```
RS G8124-E(config)# boot profile {default|aggregation|ipv6|routing}
                                        (Select deployment profile)
RS G8124-E(config)# exit               (To privileged EXEC mode)
RS G8124-E# reload                     (Reboot the switch)
```

To view the currently selected deployment profile, use the following ISCLI privileged EXEC command:

```
RS G8124-E# show boot
```

Note: When using a specialized profile, menus and commands are unavailable for features that are not supported under the profile. Such menus and commands will be available again only when a supporting profile is used.

Note: Deployment profiles other than those listed in this section should be used only under the direction of your support personnel.

Automatic Configuration Changes

Lenovo N/OS 8.3 configuration blocks and backup files generated under a specific deployment profile are generally compatible among all other deployment profiles. However, if commands or capacities configured under a prior profile are not available using the current profile, the switch will ignore the unsupported commands. Mutually supported commands will be processed normally between profiles.

All configuration commands from the prior profile are initially retained when changing profiles, even though some may be ignored when the switch starts with new profile. This allows the administrator to change back to the prior deployment profile with the prior configuration intact if desired. However, once the administrator saves the configuration under the new profile, all unsupported commands are immediately cleared. For example, when using the Routing profile, because ACLs are unsupported in that mode, their settings will be excluded when the configuration is saved. Then, if returning to the Default profile, it will be necessary to reconfigure the desired ACLs, or to use the backup configuration.

Chapter 13. Virtualization

Virtualization allows resources to be allocated in a fluid manner based on the logical needs of the data center, rather than on the strict, physical nature of components. The following virtualization features are included in Lenovo Networking OS 8.3 on the RackSwitch G8124-E (G8124-E):

- Virtual Local Area Networks (VLANs)

  VLANs are commonly used to split groups of networks into manageable broadcast domains, create logical segmentation of workgroups, and to enforce security policies among logical network segments.

  For details on this feature, see Chapter 7, "VLANs."

- Link aggregation

  A port LAG pools multiple physical switch ports into a single, high-bandwidth logical link to other devices. In addition to aggregating capacity, LAGs provides link redundancy.

  For details on this feature, see Chapter 8, "Ports and Link Aggregation."

- Virtual Link Aggregation (VLAGs)

  With VLAGs, two switches can act as a single logical device for the purpose of establishing port aggregation. Active LAG links from one device can lead to both VLAG peer switches, providing enhanced redundancy, including active-active VRRP configuration.

  For details on this feature, see Chapter 10, "Virtual Link Aggregation Groups."

- Virtual Network Interface Card (vNIC) support

  Some NICs, such as the Emulex Virtual Fabric Adapter, can virtualize NIC resources, presenting multiple virtual NICs to the server's OS or hypervisor. Each vNIC appears as a regular, independent NIC with some portion of the physical NIC's overall bandwidth. Lenovo N/OS 8.3 supports up to four vNICs over each server-side switch port.

  For details on this feature, see Chapter 14, "Virtual NICs".

- VMready

  The switch's VMready software makes it *virtualization aware*. Servers that run hypervisor software with multiple instances of one or more operating systems can present each as an independent *virtual machine* (VM). With VMready, the switch automatically discovers virtual machines (VMs) connected to switch.

  For details on this feature, see Chapter 15, "VMready."

- Edge Virtual Bridging (EVB)

  An IEEE 802.1Qbg standard that simplifies network management by providing a standards-based protocol that defines how virtual Ethernet bridges exchange configuration information. EVB bridges the gap between physical and virtual network resources by allowing networks to become virtual machine (VM)-aware.

  For details on this feature, see Chapter 21, "Edge Virtual Bridging."

N/OS virtualization features provide a highly-flexible framework for allocating and managing switch resources.

Chapter 14. Virtual NICs

A Network Interface Controller (NIC) is a component within a server that allows the server to be connected to a network. The NIC provides the physical point of connection, as well as internal software for encoding and decoding network packets.

Virtualizing the NIC helps o resolve issues caused by limited NIC slot availability. By virtualizing a 10Gbps NIC, its resources can be divided into multiple logical instances known as virtual NICs (vNICs), Each vNIC appear as a regular, independent NIC to the server operating system or a hypervisor, with each vNIC using some portion of the physical NIC's overall bandwidth.

Figure 23. Virtualizing the NIC for Multiple Virtual Pipes on Each Link
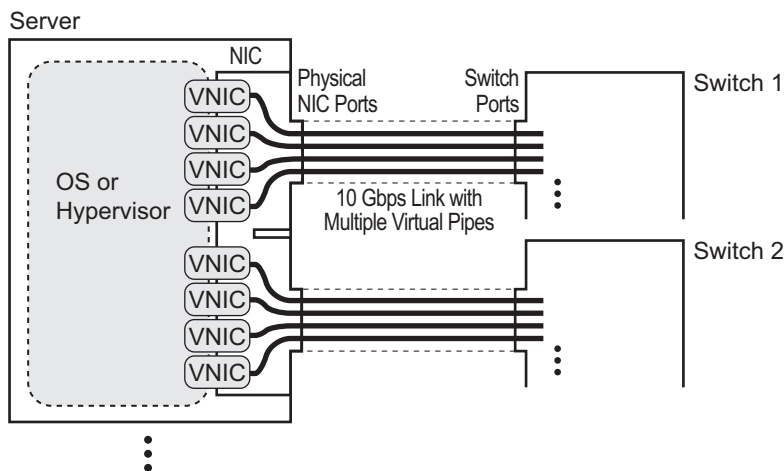
A G8124-E with Lenovo Networking OS 8.3 supports the Emulex Virtual Fabric Adapter (VFA) to provide the following vNIC features:

- Up to four vNICs are supported on each server port.

- vNICs can be grouped together, along with regular server ports, uplink ports, or LAGs, to define vNIC groups for enforcing communication boundaries.

- In the case of a failure on the uplink ports associated with a vNIC group, the switch can signal affected vNICs for failover while permitting other vNICs to continue operation.

- Each vNIC can be independently allocated a symmetric percentage of the 10Gbps bandwidth on the link (from NIC to switch, and from switch to NIC).

- The G8124-E can be used as the single point of vNIC configuration as long as the Emulex NIC is working in Lenovo Virtual Fabric mode.

The following restrictions apply to vNICs:

- vNICs are not supported simultaneously with VM groups (see Chapter 15, "VMready") on the same switch ports.

- vNICs are not supported simultaneously with DCBX (see "Data Center Bridging Capability Exchange" on page 269) or FCoE (see "Fibre Channel over Ethernet" on page 248) on the same switch.

By default, vNICs are disabled. As described in the following sections, the administrator must first define server ports prior to configuring and enabling vNICs as discussed in the rest of this section.

Defining Server Ports

vNICs are supported only on ports connected to servers. Before you configure vNICs on a port, the port must first be defined as a server port using the following command:

```
RS G8124-E(config)# system server-ports port <port alias or number>
```

Ports that are not defined as server ports are considered uplink ports and do not support vNICs.

Enabling the vNIC Feature

The vNIC feature can be globally enabled using the following command:

```
RS G8124-E(config)# vnic enable
```

Note: When the vNIC feature is enabled, the maximum number of QOS Class of Service queues available is four.

vNIC IDs

Lenovo N/OS 8.3 supports up to four vNICs attached to each server port. Each vNIC is provided its own independent virtual pipe on the port.

vNIC IDs on the Switch

On the switch, each vNIC is identified by its port and vNIC number as follows:

<port number or alias>.*<vNIC pipe number (1-4)>*

For example:

1.1, 1.2, 1.3, and 1.4 represent the vNICs on port 1.

2.1, 2.2, 2.3, and 2.4 represent the vNICs on port 2, etc.

These vNIC IDs are used when adding vNICs to vNIC groups, and are shown in some configuration and information displays.

vNIC Interface Names on the Server

When running in virtualization mode, the Emulex Virtual Fabric Adapter presents eight vNICs to the OS or hypervisor (four for each of the two physical NIC ports). Each vNIC is identified in the OS or hypervisor with a different vNIC function number (0-7). vNIC function numbers correlate to vNIC IDs on the switch as follows:

Table 18. *vNIC ID Correlation*

PCIe Function ID	NIC Port	vNIC Pipe	vNIC ID
0	0	1	$x.1$
2	0	2	$x.2$
4	0	3	$x.3$
6	0	4	$x.4$
1	1	1	$x.1$
3	1	2	$x.2$
5	1	3	$x.3$
7	1	4	$x.4$

In this, the x in the vNIC ID represents the switch port to which the NIC port is connected.

vNIC Bandwidth Metering

N/OS 8.3 supports bandwidth metering for vNIC traffic. By default, each of the four vNICs on any given port is allowed an equal share (25%) of NIC capacity when enabled. However, you may configure the percentage of available switch port bandwidth permitted to each vNIC.

vNIC bandwidth can be configured as a value from 1 to 100, with each unit representing 1% (or 100Mbps) of the 10Gbps link. By default, each vNICs enabled on a port is assigned 25 units (equal to 25% of the link, or 2.5Gbps). When traffic from the switch to the vNIC reaches its assigned bandwidth limit, the switch will drop packets egressing to the affected vNIC. Likewise, if traffic from the vNIC to the switch reaches its limit, the NIC will drop egress of any further packets. When traffic falls to less than the configured thresholds, traffic resumes at its allowed rate.

Note: Bandwidth metering drops excess packets when configured limits are reached. Consider using the ETS feature in applications where packet loss is not desirable (see "Enhanced Transmission Selection" on page 262).

To change the bandwidth allocation, use the following commands:

```
RS G8124-E(config)# vnic port <port alias or number> index <vNIC number (1-4)>
RS G8124-E(vnic-config)# bandwidth <allocated percentage>
```

Note: vNICs that are disabled are automatically allocated a bandwidth value of 0.

A combined maximum of 100 units can be allocated among vNIC pipes enabled for any specific port (bandwidth values for disabled pipes are not counted). If more than 100 units are assigned to enabled pipes, an error will be reported when attempting to apply the configuration.

The bandwidth metering configuration is synchronized between the switch and vNICs. Once configured on the switch, there is no need to manually configure vNIC bandwidth metering limits on the NIC as long as it is in Lenovo Virtual Fabric mode.

vNIC Groups

vNICs can be grouped together, along with uplink ports and LAGs, as well as other ports that were defined as server ports but not connected to vNICs. Each vNIC group is essentially a separate virtual network within the switch. Elements within a vNIC group have a common logical function and can communicate with each other, while elements in different vNIC groups are separated.

N/OS 8.3 supports up to 32 independent vNIC groups.

To enforce group boundaries, each vNIC group is assigned its own unique VLAN. The vNIC group VLAN ID is placed on all vNIC group packets as an "outer" tag. As shown in Figure 24, the outer vNIC group VLAN ID is placed on the packet in addition to any regular VLAN tag assigned by the network, server, or hypervisor. The outer vNIC group VLAN is used only between the G8124-E and the NIC.

Figure 24. Outer and Inner VLAN Tags
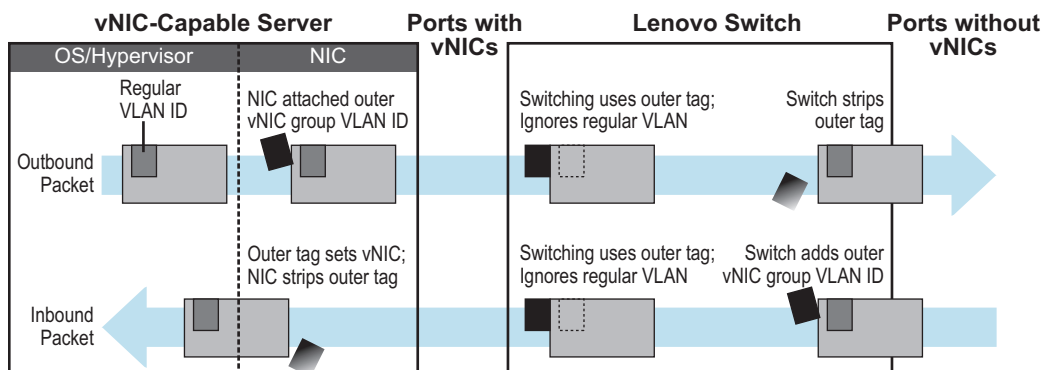
Within the G8124-E, all Layer 2 switching for packets within a vNIC group is based on the outer vNIC group VLAN. The G8124-E does not consider the regular, inner VLAN ID (if any) for any VLAN-specific operation.

The outer vNIC group VLAN is removed by the NIC before the packet reaches the server OS or hypervisor, or by the switch before the packet egresses any switch port which does not need it for vNIC processing.

The VLAN configured for the vNIC group will be automatically assigned to member vNICs, ports, and LAGs must not be manually configured for those elements.

Note: Once a VLAN is assigned to a vNIC group, that VLAN is used only for vNIC purposes and is no longer available for configuration. Likewise, any VLAN configured for regular purposes cannot be configured as a vNIC group VLAN.

Other vNIC group rules are as follows:

- vNIC groups may have one or more vNIC members. However, any given vNIC can be a member of only one vNIC group.

- All vNICs on a given port must belong to different vNIC groups.

- All members of a vNIC group must have the same vNIC pipe index. For instance, 1.1 and 2.1 share the same ".1" vNIC pipe index, but 3.2 uses the ".2" vNIC pipe index and cannot be placed in the same vNIC group.

- Uplink ports which are part of a LAG may not be individually added to a vNIC group. Only one individual uplink port or one static LAG (consisting of multiple uplink ports) may be added to any given vNIC group.

- When a port is added to a vNIC group, flow control is disabled automatically. If the port is removed from the vNIC group, the flow-control setting remains disabled.

- For any switch ports or port LAG connected to regular (non-vNIC) devices:
  - These elements can be placed in only one vNIC group (they cannot be members of multiple vNIC groups).
  - Once added to a vNIC group, the PVID for the element is automatically set to use the vNIC group VLAN number, and PVID tagging on the element is automatically disabled.
  - By default, STP is disabled on non-server ports or LAGs added to a vNIC group. STP cannot be re-enabled on the port.

- Because regular, inner VLAN IDs are ignored by the switch for traffic in vNIC groups, following rules and restrictions apply:
  - The inner VLAN tag may specify any VLAN ID in the full, supported range (1 to 4095) and may even duplicate outer vNIC group VLAN IDs.
  - Per-VLAN IGMP snooping is not supported in vNIC groups.
  - The inner VLAN tag is not processed in any way in vNIC groups: The inner tag cannot be stripped or added on egress port, is not used to restrict multicast traffic, is not matched against ACL filters, and does not influence Layer 3 switching.
  - For vNIC ports on the switch, because the outer vNIC group VLAN is transparent to the OS/hypervisor and upstream devices, configure VLAN tagging as normally required (on or off) for the those devices, ignoring any outer tag.

- Virtual machines (VMs) and other VEs associated with vNICs are automatically detected by the switch when VMready is enabled (see Chapter 15, "VMready"). However, vNIC groups are isolated from other switch elements. VEs in vNIC groups cannot be assigned to VM groups.

vNIC Teaming Failover

For NIC failover in a non-virtualized environment, when a service group's uplink ports fail or are disconnected, the switch disables the affected group's server ports, causing the server to failover to the backup NIC and switch.

However, in a virtualized environment, disabling the affected server ports would disrupt all vNIC pipes on those ports, not just those that have lost their uplinks (see Figure 25).

Figure 25. Regular Failover in a Virtualized Environment
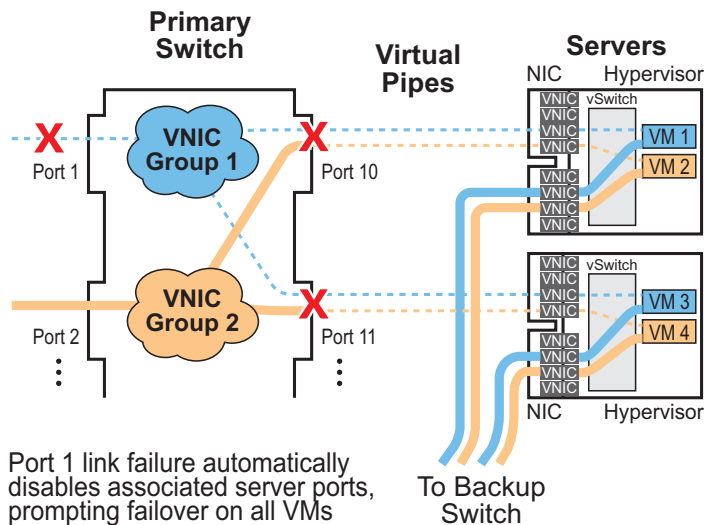
To avoid disrupting vNICs that have not lost their uplinks, N/OS 8.3 and the Emulex Virtual Fabric Adapter provide vNIC-aware failover. When a vNIC group's uplink ports fail, the switch cooperates with the affected NIC to prompt failover only on the appropriate vNICs. This allows the vNICs that are not affected by the failure to continue without disruption (see Figure 26 on page 218).

Figure 26. vNIC Failover Solution

Upon Port 1 link failure, the switch informs the server hypervisor for failover on affected VNICs only
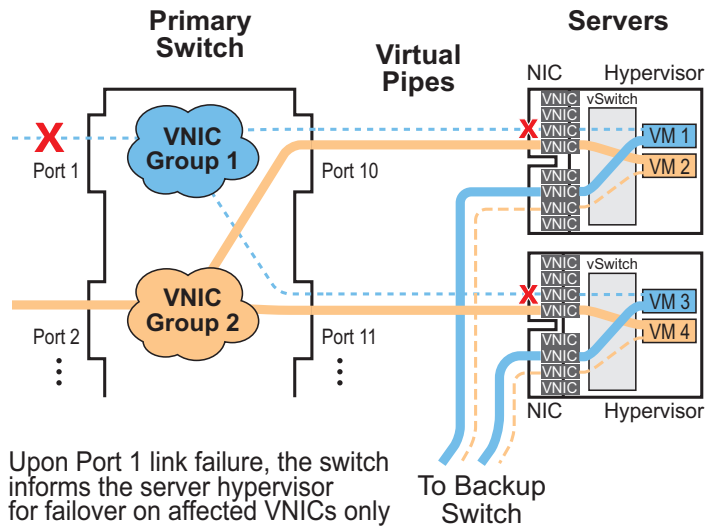
By default, vNIC Teaming Failover is disabled on each vNIC group, but can be enabled or disabled independently for each vNIC group using the following commands:

```
RS G8124-E(config)# vnic vnicgroup <group number>
RS G8124-E(vnic-group-config)# failover
```

vNIC Configuration Example

Basic vNIC Configuration

Consider the following example configuration:

Figure 27. Multiple vNIC Groups
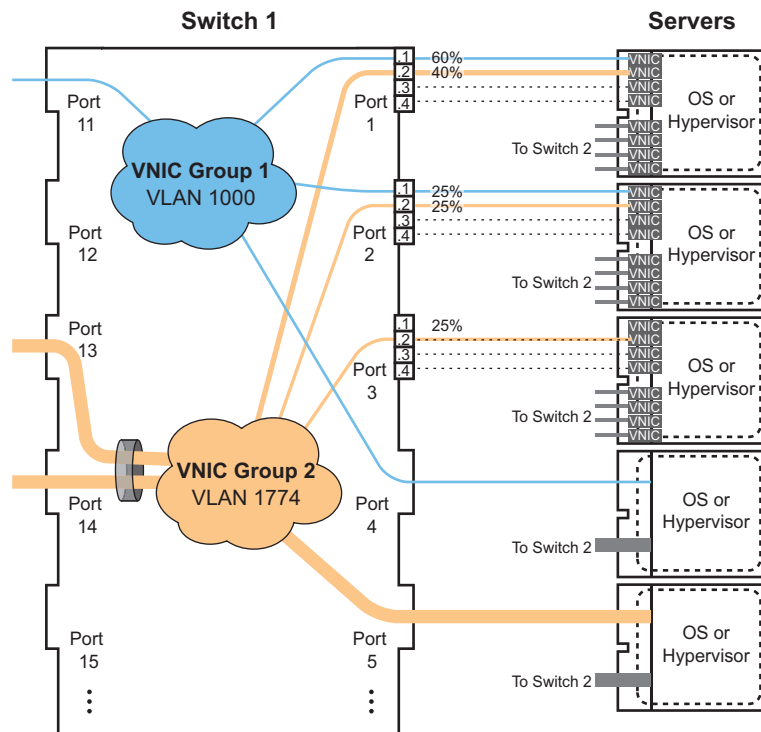
Figure 27 has the following vNIC network characteristics:

- vNIC group 1 has an outer tag for VLAN 1000. The group is comprised of vNIC pipes 1.1 and 2.1, switch server port 4 (a non-vNIC port), and uplink port 11.

- vNIC group 2 has an outer tag for VLAN 1774. The group is comprised of vNIC pipes 1.2, 2.2 and 3.2, switch server port 5, and an uplink LAG of ports 13 and 14.

- vNIC failover is enabled for both vNIC groups.

- vNIC bandwidth on port 1 is set to 60% for vNIC 1 and 40% for vNIC 2.

- Other enabled vNICs (2.1, 2.2, and 3.2) are permitted the default bandwidth of 25% (2.5Gbsp) on their respective ports.

- All remaining vNICs are disabled (by default) and are automatically allocated 0 bandwidth.

1. Define the server ports.

```
RS G8124-E(config)# system server-ports port 1-5
```

2. Configure the external LAG to be used with vNIC group 2.

```
RS G8124-E(config)# portchannel 1 port 13,14
RS G8124-E(config)# portchannel 1 enable
```

3. Enable the vNIC feature on the switch.

```
RS G8124-E(config)# vnic enable
```

4. Configure the virtual pipes for the vNICs attached to each server port:

```
RS G8124-E(config)# vnic port 1 index 1 (Select vNIC 1 on the port)
RS G8124-E(vnic-config)# enable        (Enable the vNIC pipe)
RS G8124-E(vnic-config)# bandwidth 60 (Allow 60% egress bandwidth)
RS G8124-E(vnic-config)# exit
RS G8124-E(config)# vnic port 1 index 2(Select vNIC 2 on the port)
RS G8124-E(vnic-config)# enable        (Enable the vNIC pipe)
RS G8124-E(vnic-config)# bandwidth 40 (Allow 40% egress bandwidth)
RS G8124-E(vnic-config)# exit

RS G8124-E(config)# vnic port 2 index 1 (Select vNIC 1 on the port)
RS G8124-E(vnic-config)# enable        (Enable the vNIC pipe)
RS G8124-E(vnic-config)# exit
RS G8124-E(config)# vnic port 2 index 2 (Select vNIC 2 on the port)
RS G8124-E(vnic-config)# enable        (Enable the vNIC pipe)
RS G8124-E(vnic-config)# exit
```

As a configuration shortcut, vNICs do not have to be explicitly enabled in this step. When a vNIC is added to the vNIC group (in the next step), the switch will prompt you to confirm automatically enabling the vNIC if it is not yet enabled (shown for 3.2).

Note: vNICs are not supported simultaneously on the same switch ports as VMready.

5. Add ports, LAGs, and virtual pipes to their vNIC groups.

```
RS G8124-E(config)# vnic vnicgroup 1    (Select vNIC group)
RS G8124-E(vnic-group-config)# vlan 1000(Specify the VLAN)
RS G8124-E(vnic-group-config)# member 1.1(Add vNIC pipes to the group)
RS G8124-E(vnic-group-config)# member 2.1
RS G8124-E(vnic-group-config)# port 4     (Add non-vNIC port to the group)
RS G8124-E(vnic-group-config)# port 11 (Add uplink port to the group
RS G8124-E(vnic-group-config)# failover(Enable vNIC failover for the group)
RS G8124-E(vnic-group-config)# enable   (Enable the vNIC group)
RS G8124-E(vnic-group-config)# exit

RS G8124-E(config)# vnic vnicgroup 2
RS G8124-E(vnic-group-config)# vlan 1774
RS G8124-E(vnic-group-config)# member 1.2
RS G8124-E(vnic-group-config)# member 2.2
RS G8124-E(vnic-group-config)# member 3.2
vNIC 3.2 is not enabled.
Confirm enabling vNIC3.2 [y/n]: y
RS G8124-E(vnic-group-config)# port 5
RS G8124-E(vnic-group-config)# trunk 1
RS G8124-E(vnic-group-config)# failover
RS G8124-E(vnic-group-config)# enable
RS G8124-E(vnic-group-config)# exit
```

Once VLAN 1000 and 1774 are configured for vNIC groups, they will not be available for configuration in the regular VLAN menus
(RS G8124-E(config)# vlan <VLAN number>).

Note: vNICs are not supported simultaneously on the same switch ports as VMready.

6. Save the configuration.

vNICs for iSCSI on Emulex Endeavor 2

The N/OS vNIC feature works with standard network applications like iSCSI as previously described. However, the Emulex Endeavor 2 NIC expects iSCSI traffic to occur only on a single vNIC pipe. When using the Emulex Endeavor 2, only vNIC pipe 2 may participate in ISCSI.

To configure the switch for this solution, place iSCSI traffic in its own vNIC group, comprised of the uplink port leading to the iSCSI target, and the related *<port>.2* vNIC pipes connected to the participating servers. For example:

1. Define the server ports.

```
RS G8124-E(config)# system server-ports port 1-3
```

2. Enable the vNIC feature on the switch.

```
RS G8124-E # vnic enable
```

3. Configure the virtual pipes for the iSCSI vNICs attached to each server port:

```
RS G8124-E(config)# vnic port 1 index 2   (Select vNIC 2 on the server port)
RS G8124-E(vnic_config)# enable            (Enable the vNIC pipe)
RS G8124-E(vnic_config)# exit
RS G8124-E(config)# vnic port 2 index 2    (Select vNIC 2 on the server port)
RS G8124-E(vnic_config)# enable            (Enable the vNIC pipe)
RS G8124-E(vnic_config)# exit
RS G8124-E(config)# vnic port 3 index 2    (Select vNIC 2 on the server port)
RS G8124-E(vnic_config)# enable            (Enable the vNIC pipe)
RS G8124-E(vnic_config)# exit
```

Note: vNICs are not supported simultaneously on the same switch ports as VMready, or on the same switch as DCBX or FCoE.

4. Add ports and virtual pipes to a vNIC group.

```
RS G8124-E(config)# vnic vnicgroup 1    (Select vNIC group)
RS G8124-E(vnic-group-config)# vlan 1000 (Specify the VLAN)
RS G8124-E(vnic-group-config)# member 1.2 (Add iSCSI vNIC pipes to the group)
RS G8124-E(vnic-group-config)# member 2.2
RS G8124-E(vnic-group-config)# member 3.2
RS G8124-E(vnic-group-config)# port 4   (Add the uplink port to the group)
RS G8124-E(vnic-group-config)# enable   (Enable the vNIC group)
RS G8124-E(vnic-group-config)# exit
```

5. Save the configuration.

Chapter 16, "FCoE and CEE"

Chapter 15. VMready

Virtualization is used to allocate server resources based on logical needs, rather than on strict physical structure. With appropriate hardware and software support, servers can be virtualized to host multiple instances of operating systems, known as virtual machines (VMs). Each VM has its own presence on the network and runs its own service applications.

Software known as a *hypervisor* manages the various virtual entities (VEs) that reside on the host server: VMs, virtual switches, and so on. Depending on the virtualization solution, a virtualization management server may be used to configure and manage multiple hypervisors across the network. With some solutions, VMs can even migrate between host hypervisors, moving to different physical hosts while maintaining their virtual identity and services.

The Lenovo Networking OS 8.3 VMready feature supports up to 2048 VEs in a virtualized data center environment. The switch automatically discovers the VEs attached to switch ports, and distinguishes between regular VMs, Service Console Interfaces, and Kernel/Management Interfaces in a VMware® environment.

VEs may be placed into VM groups on the switch to define communication boundaries: VEs in the same VM group may communicate with each other, while VEs in different groups may not. VM groups also allow for configuring group-level settings such as virtualization policies and ACLs.

The administrator can also pre-provision VEs by adding their MAC addresses (or their IPv4 address or VM name in a VMware environment) to a VM group. When a VE with a pre-provisioned MAC address becomes connected to the switch, the switch will automatically apply the appropriate group membership configuration.

The G8124-E with VMready also detects the migration of VEs across different hypervisors. As VEs move, the G8124-E NMotion™ feature automatically moves the appropriate network configuration as well. NMotion gives the switch the ability to maintain assigned group membership and associated policies, even when a VE moves to a different port on the switch.

VMready also works with VMware Virtual Center (vCenter) management software. Connecting with a vCenter allows the G8124-E to collect information about more distant VEs, synchronize switch and VE configuration, and extend migration properties.

Note: VM groups and policies, VE pre-provisioning, and VE migration features are not supported simultaneously on the same ports as vNICs (see Chapter 14, "Virtual NICs").

VE Capacity

When VMready is enabled, the switch will automatically discover VEs that reside in hypervisors directly connected on the switch ports. Lenovo N/OS 8.3 supports up to 2048 VEs. Once this limit is reached, the switch will reject additional VEs.

Note: In rare situations, the switch may reject new VEs prior to reaching the supported limit. This can occur when the internal hash corresponding to the new VE is already in use. If this occurs, change the MAC address of the VE and retry the operation. The MAC address can usually be changed from the virtualization management server console (such as the VMware Virtual Center).

Defining Server Ports

Before you configure VMready features, you must first define whether ports are connected to servers or are used as uplink ports. Use the following ISCLI configuration command to define a port as a server port:

```
RS G8124-E(config)# system server-ports port <port alias or number>
```

Ports that are not defined as server ports are automatically considered uplink ports.

VM Group Types

VEs, as well as switch server ports, switch uplink ports, static LAGs, and LACP LAGs, can be placed into VM groups on the switch to define virtual communication boundaries. Elements in a given VM group are permitted to communicate with each other, while those in different groups are not. The elements within a VM group automatically share certain group-level settings.

N/OS 8.3 supports up to 1024 VM groups. There are two different types:

● Local VM groups are maintained locally on the switch. Their configuration is not synchronized with hypervisors. Of the 2048 VEs supported on the switch, up to 500 VEs may be used in local groups.
● Distributed VM groups are automatically synchronized with a virtualization management server (see "Assigning a vCenter" on page 234).

Each VM group type is covered in detail in the following sections.

Note: VM groups are not supported simultaneously on the same ports as vNICs (see Chapter 14, "Virtual NICs").

Local VM Groups

The configuration for local VM groups is maintained on the switch (locally) and is not directly synchronized with hypervisors. Local VM groups may include only local elements: local switch ports and LAGs, and only those VEs connected to one of the switch ports or pre-provisioned on the switch. Of the 2048 VEs supported on the switch, up to 500 VEs may be used in local groups.

Local VM groups support limited VE migration: as VMs and other VEs move to different hypervisors connected to different ports on the switch, the configuration of their group identity and features moves with them. However, VE migration to and from more distant hypervisors (those not connected to the G8124-E, may require manual configuration when using local VM groups.

Configuring a Local VM Group

Use the following ISCLI configuration commands to assign group properties and membership:

```
RS G8124-E(config)# virt vmgroup <VM group number> ?
  key <LACP LAG key>                          (Add LACP LAG to group)
  port <port alias or number>                 (Add port member to group)
  portchannel <LAG number>                    (Add static LAG to group)
  profile <profile name>                      (Not used for local groups)
  stg <Spanning Tree group>                   (Add STG to group)
  tag                                         (Set VLAN tagging on ports)
  validate <advanced|basic>                   (Validate mode for the group)
  vlan <VLAN number>                          (Specify the group VLAN)
  vm <MAC>|<index>|<UUID>|<IPv4 address>|<name>(Add VM member to group)
  vmap <VMAP number> [intports|extports](Specify VMAP number)
```

The following rules apply to the local VM group configuration commands:

- `key`: Add LACP LAGs to the group.
- `port`: Add switch server ports or switch uplink ports to the group. Note that VM groups and vNICs (see Chapter 14, "Virtual NICs") are not supported simultaneously on the same port.
- `portchannel`: Add static port LAGs to the group.
- `profile`: The profile options are not applicable to local VM groups. Only distributed VM groups may use VM profiles (see "VM Profiles" on page 227).
- `stg`: The group may be assigned to a Spanning-Tree group for broadcast loop control (see Chapter 9, "Spanning Tree Protocols").
- `tag`: Enable VLAN tagging for the VM group. If the VM group contains ports which also exist in other VM groups, enable tagging in both VM groups.
- `validate`: Set validation mode for the group.
- `vlan`: Each VM group must have a unique VLAN number. This is required for local VM groups. If one is not explicitly configured, the switch will automatically assign the next unconfigured VLAN when a VE or port is added to the VM group.
- `vmap`: Each VM group may optionally be assigned a VLAN-based ACL (see "VLAN Maps" on page 238).
- `vm`: Add VMs.

  VMs and other VEs are primarily specified by MAC address. They can also be specified by UUID, IP address, or by the index number as shown in various VMready information output (see "VMready Information Displays" on page 241).

Use the `no` variant of the commands to remove or disable VM group configuration settings:

```
RS G8124-E(config)# no virt vmgroup <VM group number> [?]
```

Note: Local VM groups are not supported simultaneously on the same ports as vNICs (see Chapter 14, "Virtual NICs").

Distributed VM Groups

Distributed VM groups allow configuration profiles to be synchronized between the G8124-E and associated hypervisors and VEs. This allows VE configuration to be centralized, and provides for more reliable VE migration across hypervisors.

Using distributed VM groups requires a virtualization management server. The management server acts as a central point of access to configure and maintain multiple hypervisors and their VEs (VMs, virtual switches, and so on).

The G8124-E must connect to a virtualization management server before distributed VM groups can be used. The switch uses this connection to collect configuration information about associated VEs, and can also automatically push configuration profiles to the virtualization management server, which in turn configures the hypervisors and VEs. See "Virtualization Management Servers" on page 234 for more information.

Note: Distributed VM groups are not supported simultaneously on the same ports as vNICs (see Chapter 14, "Virtual NICs").

VM Profiles

VM profiles are required for configuring distributed VM groups. They are not used with local VM groups. A VM profile defines the VLAN and virtual switch bandwidth shaping characteristics for the distributed VM group. The switch distributes these settings to the virtualization management server, which in turn distributes them to the appropriate hypervisors for VE members associated with the group.

Creating VM profiles is a two part process. First, the VM profile is created as shown in the following command on the switch:

```
RS G8124-E(config)# virt vmprofile <profile name>
```

Next, the profile must be edited and configured using the following configuration commands:

```
RS G8124-E(config)# virt vmprofile edit <profile name> ?
eshaping <average bandwidth> <burst size> <peak>
shaping <average bandwidth> <burst size> <peak>
vlan <VLAN number>
```

For virtual switch bandwidth shaping parameters, average and peak bandwidth are specified in kilobits per second (a value of 1000 represents 1 Mbps). Burst size is specified in kilobytes (a value of 1000 represents 1 MB). Eshaping (egress shaping) is used for distributed virtual switch.

Note: The bandwidth shaping parameters in the VM profile are used by the hypervisor virtual switch software. To set bandwidth policies for individual VEs, see "VM Policy Bandwidth Control" on page 239.

Once configured, the VM profile may be assigned to a distributed VM group as shown in the following section.

Initializing a Distributed VM Group

Note: A VM profile is required before a distributed VM group may be configured. See "VM Profiles" on page 227 for details.

Once a VM profile is available, a distributed VM group may be initialized using the following configuration command:

```
RS G8124-E(config)# virt vmgroup <VM group number> profile <VM profile name>
```

Only one VM profile can be assigned to a given distributed VM group. To change the VM profile, the old one must first be removed using the following ISCLI configuration command:

```
RS G8124-E(config)# no virt vmgroup <VM group number> profile
```

Note: The VM profile can be added only to an empty VM group (one that has no VLAN, VMs, or port members). Any VM group number currently configured for a local VM group (see "Local VM Groups" on page 225) cannot be converted and must be deleted before it can be used for a distributed VM group.

Assigning Members

VMs, ports, and LAGs may be added to the distributed VM group only after the VM profile is assigned. Group members are added, pre-provisioned, or removed from distributed VM groups in the same manner as with local VM groups ("Local VM Groups" on page 225), with the following exceptions:

- VMs: VMs and other VEs are not required to be local. Any VE known by the virtualization management server can be part of a distributed VM group.

- The VM group vlan option (see page 226) cannot be used with distributed VM groups. For distributed VM groups, the VLAN is assigned in the VM profile.

Synchronizing the Configuration

When the configuration for a distributed VM group is modified, the switch updates the assigned virtualization management server. The management server then distributes changes to the appropriate hypervisors.

For VM membership changes, hypervisors modify their internal virtual switch port groups, adding or removing server port memberships to enforce the boundaries defined by the distributed VM groups. Virtual switch port groups created in this fashion can be identified in the virtual management server by the name of the VM profile, formatted as follows:

Lenovo_<*VM profile name*>

(or)

Lenovo_<*VM profile name*> <*index number*>   *(for vDS)*

Adding a server host interface to a distributed VM group does not create a new port group on the virtual switch or move the host. Instead, because the host interface already has its own virtual switch port group on the hypervisor, the VM profile settings are applied to its existing port group.

Note: When applying the distributed VM group configuration, the virtualization management server and associated hypervisors must take appropriate actions. If a hypervisor is unable to make requested changes, an error message will be displayed on the switch. Be sure to evaluate all error message and take the appropriate actions for the expected changes to apply.

Removing Member VEs

Removing a VE from a distributed VM group on the switch will have the following effects on the hypervisor:

- The VE will be moved to the Lenovo_Default port group in VLAN 0 (zero).

- Traffic shaping will be disabled for the VE.

- All other properties will be reset to default values inherited from the virtual switch.

VMcheck

The G8124-E primarily identifies virtual machines by their MAC addresses. An untrusted server or a VM could identify itself by a trusted MAC address leading to MAC spoofing attacks. Sometimes, MAC addresses get transferred to another VM, or they get duplicated.

The VMcheck solution addresses these security concerns by validating the MAC addresses assigned to VMs. The switch periodically sends hello messages on server ports. These messages include the switch identifier and port number. The hypervisor listens to these messages on physical NICs and stores the information, which can be retrieved using the VMware Infrastructure Application Programming Interface (VI API). This information is used to validate VM MAC addresses. Two modes of validation are available: Basic and Advanced.

Use the following command to select the validation mode or to disable validation:

```
RS G8124-E(config)# [no] virt vmgroup <VM group number> validate
{basic|advanced}
```

Basic Validation

This mode provides port-based validation by identifying the port used by a hypervisor. It is suitable for environments in which MAC reassignment or duplication cannot occur.

The switch, using the hello message information, identifies a hypervisor port. If the hypervisor port is found in the hello message information, it is deemed to be a trusted port. Basic validation should be enabled when:

- A VM is added to a VM group, and the MAC address of the VM interface is in the Layer 2 table of the switch.

- A VM interface that belongs to a VM group experiences a "source miss" i.e. is not able to learn new MAC address.

- A trusted port goes down. Port validation must be performed to ensure that the port does not get connected to an untrusted source when it comes back up.

Use the following command to set the action to be performed if the switch is unable to validate the VM MAC address:

```
RS G8124-E(config)# virt vmcheck action basic {log|link}

log - generates a log
link - disables the port
```

Advanced Validation

This mode provides VM-based validation by mapping a switch port to a VM MAC address. It is suitable for environments in which spoofing, MAC reassignment, or MAC duplication is possible.

When the switch receives frames from a VM, it first validates the VM interface based on the VM MAC address, VM Universally Unique Identifier (UUID), Switch port, and Switch ID available in the hello message information. Only if all the four parameters are matched, the VM MAC address is considered valid.

In advanced validation mode, if the VM MAC address validation fails, an ACL can be automatically created to drop the traffic received from the VM MAC address on the switch port. Use the following command to specify the number of ACLs to be automatically created for dropping traffic:

```
RS G8124-E(config)# virt vmcheck acls max <1-127>
```

Use the following command to set the action to be performed if the switch is unable to validate the VM MAC address:

```
RS G8124-E(config)# virt vmcheck action advanced {log|link|acl}
```

Following are the other VMcheck commands:

Table 19. *VMcheck Commands*

Command	Description
RS G8124-E(config)# virt vmware hello {ena| hport *<port number>*|haddr|htimer}	Hello messages setting: enable/add port/advertise this IP address in the hello messages instead of the default management IP address/set the timer to send the hello messages
RS G8124-E(config)# no virt vmware hello {enable|hport *<port number>*}	Disable hello messages/remove port
RS G8124-E(config)# [no] virt vmcheck trust *<port number or range>*	Mark a port as trusted; Use the no form of the command to mark port as untrusted
RS G8124-E# no virt vmcheck acls	ACLs cannot be used for VMcheck

Virtual Distributed Switch

A virtual Distributed Switch (vDS) allows the hypervisor's NIC to be attached to the vDS instead of its own virtual switch. The vDS connects to the vCenter and spans across multiple hypervisors in a datacenter. The administrator can manage virtual machine networking for the entire data center from a single interface. The vDS enables centralized provisioning and administration of virtual machine networking in the data center using the VMware vCenter server.

When a member is added to a distributed VM group, a distributed port group is created on the vDS. The member is then added to the distributed port group.

Distributed port groups on a vDS are available to all hypervisors that are connected to the vDS. Members of a single distributed port group can communicate with each other.

Note: vDS works with ESX 4.0 or higher versions.

To add a vDS, use the command:

```
RS G8124-E# virt vmware dvswitch add <datacenter name> <dvSwitch name>
[<dvSwitch-version>]
```

Prerequisites

Before adding a vDS on the G8124-E, ensure the following:
- VMware vCenter is fully installed and configured and includes a "bladevm" administration account and a valid SSL certificate.
- A virtual distributed switch instance has been created on the vCenter. The vDS version must be higher or the same as the hypervisor version on the hosts.
- At least two hypervisors are configured.

Guidelines

Before migrating VMs to a vDS, consider the following:
- At any one time, a VM NIC can be associated with only one virtual switch: to the hypervisor's virtual switch, or to the vDS.

- Management connection to the server must be ensured during the migration. The connection is via the Service Console or the Kernel/Management Interface.

- The vDS configuration and migration can be viewed in vCenter at the following locations:
  - vDS: `Home> Inventory > Networking`
  - vDS Hosts: `Home > Inventory > Networking > vDS > Hosts`

Note: These changes will not be displayed in the running configuration on the G8124-E.

Migrating to vDS

You can migrate VMs to the vDS using vCenter. The migration may also be accomplished using the operational commands on the G8124-E available in the following CLI menus:

For VMware vDS operations:

```
RS G8124-E# virt vmware dvswitch ?
  add       Add a dvSwitch to a DataCenter
  addhost   Add a host to a dvSwitch
  adduplnk  Add a physical NIC to dvSwitch uplink ports
  del       Remove a dvSwitch from a DataCenter
  remhost   Remove a host from a dvSwitch
  remuplnk  Remove a physical NIC from dvSwitch uplink ports
```

For VMware distributed port group operations:

```
RS G8124-E# virt vmware dpg ?
  add     Add a port group to a dvSwitch
  del     Delete a port group from a dvSwitch
  update  Update a port group on a dvSwitch
  vmac    Change a VM NIC's port group
```

Virtualization Management Servers

The G8124-E can connect with a virtualization management server to collect configuration information about associated VEs. The switch can also automatically push VM group configuration profiles to the virtualization management server, which in turn configures the hypervisors and VEs, providing enhanced VE mobility.

One virtual management server must be assigned on the switch before distributed VM groups may be used. N/OS 8.3 currently supports only the VMware Virtual Center (vCenter).

Note: Although VM groups and policies are not supported simultaneously on the same ports as vNICs (see Chapter 14, "Virtual NICs"), vCenter synchronization can provide additional information about VEs on vNIC and non-vNIC ports.

Assigning a vCenter

Assigning a vCenter to the switch requires the following:

- The vCenter must have a valid IPv4 address which is accessible to the switch (IPv6 addressing is not supported for the vCenter).
- A user account must be configured on the vCenter to provide access for the switch. The account must have (at a minimum) the following vCenter user privileges:
    - Network
    - Host Network > Configuration
    - Virtual Machine > Modify Device Settings

Once vCenter requirements are met, the following configuration command can be used on the G8124-E to associate the vCenter with the switch:

```
 RS G8124-E(config)# virt vmware vcspec <vCenter IPv4 address> <username> [noauth]
```

This command specifies the IPv4 address and account username that the switch will use for vCenter access. Once entered, the administrator will be prompted to enter the password for the specified vCenter account.

The noauth option causes to the switch to ignores SSL certificate authentication. This is required when no authoritative SSL certificate is installed on the vCenter.

Note: By default, the vCenter includes only a self-signed SSL certificate. If using the default certificate, the noauth option is required.

Once the vCenter configuration has been applied on the switch, the G8124-E will connect to the vCenter to collect VE information.

vCenter Scans

Once the vCenter is assigned, the switch will periodically scan the vCenter to collect basic information about all the VEs in the datacenter, and more detailed information about the local VEs that the switch has discovered attached to its own ports.

The switch completes a vCenter scan approximately every two minutes. Any major changes made through the vCenter may take up to two minutes to be reflected on the switch. However, you can force an immediate scan of the vCenter by using one of the following ISCLI privileged EXEC commands:

```
RS G8124-E# virt vmware scan              (Scan the vCenter)
  -or-
RS G8124-E# show virt vm -v -r            (Scan vCenter and display result)
```

Deleting the vCenter

To detach the vCenter from the switch, use the following configuration command:

```
RS G8124-E(config)# no virt vmware vcspec
```

Note: Without a valid vCenter assigned on the switch, any VE configuration changes must be manually synchronized.

Deleting the assigned vCenter prevents synchronizing the configuration between the G8124-E and VEs. VEs already operating in distributed VM groups will continue to function as configured, but any changes made to any VM profile or distributed VM group on the switch will affect only switch operation; changes on the switch will not be reflected in the vCenter or on the VEs. Likewise, any changes made to VE configuration on the vCenter will no longer be reflected on the switch.

Exporting Profiles

VM profiles for discovered VEs in distributed VM groups are automatically synchronized with the virtual management server and the appropriate hypervisors. However, VM profiles can also be manually exported to specific hosts before individual VEs are defined on them.

By exporting VM profiles to a specific host, virtual machine port groups will be available to the host's internal virtual switches so that new VMs may be configured to use them.

VM migration requires that the target hypervisor includes all the virtual switch port groups to which the VM connects on the source hypervisor. The VM profile export feature can be used to distribute the associated port groups to all the potential hosts for a given VM.

A VM profile can be exported to a host using the following ISCLI privileged EXEC command:

```
RS G8124-E# virt vmware export <VM profile name> <host list> <virtual switch name>
```

The host list can include one or more target hosts, specified by host name, IPv4 address, or UUID, with each list item separated by a space.

Once executed, the requisite port group will be created on the specified virtual switch. If the specified virtual switch does not exist on the target host, the port group will not be created.

VMware Operational Commands

The G8124-E may be used as a central point of configuration for VMware virtual switches and port groups using the following ISCLI privileged EXEC commands:

```
RS G8124-E# virt vmware ?
dpg        Distributed port group operations
dvswitch   VMWare dvSwitch operations
export     Create or update a vm profile on one host
pg         Add a port group to a host
scan       Perform a VM Agent scan operation now
updpg      Update a port group on a host
vmacpg     Change a vnic's port group
vsw        Add a vswitch to a host
```

Pre-Provisioning VEs

VEs may be manually added to VM groups in advance of being detected on the switch ports. By pre-provisioning the MAC address of VEs that are not yet active, the switch will be able to later recognize the VE when it becomes active on a switch port, and immediately assign the proper VM group properties without further configuration.

Undiscovered VEs are added to or removed from VM groups using the following configuration commands:

```
RS G8124-E(config)# [no] virt vmgroup <VM group number> vm <VE MAC address>
```

For the pre-provisioning of undiscovered VEs, a MAC address is required. Other identifying properties, such as IPv4 address or VM name permitted for known VEs, cannot be used for pre-provisioning.

Note: Because VM groups are isolated from vNIC groups (see "vNIC Groups" on page 215), pre-provisioned VEs that appear on vNIC ports will not be added to the specified VM group upon discovery.

VLAN Maps

A VLAN map (VMAP) is a type of Access Control List (ACL) that is applied to a VLAN or VM group rather than to a switch port as with regular ACLs (see Chapter 6, "Access Control Lists"). In a virtualized environment, VMAPs allow you to create traffic filtering and metering policies that are associated with a VM group VLAN, allowing filters to follow VMs as they migrate between hypervisors.

Note: VLAN maps for VM groups are not supported simultaneously on the same ports as vNICs (see Chapter 14, "Virtual NICs").

N/OS 8.3 supports up to 127 VMAPs when the switch is operating in the Balanced deployment mode (see "Available Profiles" on page 206). VMAP menus and commands are not available in the Routing deployment modeIndividual VMAP filters are configured in the same fashion as regular ACLs, except that VLANs cannot be specified as a filtering criteria (unnecessary, since VMAPs are assigned to a specific VLAN or associated with a VM group VLAN).

VMAPs are configured using the following ISCLI configuration command path:

```
RS G8124-E(config)# access-control vmap <VMAP ID> ?
  action        Set filter action
  ethernet      Ethernet header options
  ipv4          IP version 4 header options
  meter         ACL metering configuration
  mirror        Mirror options
  re-mark       ACL re-mark configuration
  statistics    Enable access control list statistics
  tcp-udp       TCP and UDP filtering options
```

Once a VMAP filter is created, it can be assigned or removed using the following commands:

● For regular VLANs, use config-vlan mode:

```
RS G8124-E(config)# vlan <VLAN ID>
RS G8124-E(config-vlan)# [no] vmap <VMAP ID> [serverports|
non-serverports]
```

● For a VM group, use the global configuration mode:

```
RS G8124-E(config)# [no] virt vmgroup <ID> vmap <VMAP ID>
[serverports|non-serverports]
```

Note: Each VMAP can be assigned to only one VLAN or VM group. However, each VLAN or VM group may have multiple VMAPs assigned to it.

The optional serverports or non-serverports parameter can be specified to apply the action (to add or remove the VMAP) for either the switch server ports (serverports) or switch uplink ports (non-serverports). If omitted, the operation will be applied to all ports in the associated VLAN or VM group.

Note: VMAPs have a lower priority than port-based ACLs. If both an ACL and a VMAP match a particular packet, both filter actions will be applied as long as there is no conflict. In the event of a conflict, the port ACL will take priority, though switch statistics will count matches for both the ACL and VMAP.

VM Policy Bandwidth Control

Note: VM policy bandwidth control is supported only when the switch is operating with the Default deployment profile (see "Available Profiles" on page 206). If using the Routing profile, VM policy bandwidth control commands will not be available.

In a virtualized environment where VEs can migrate between hypervisors and thus move among different ports on the switch, traffic bandwidth policies must be attached to VEs, rather than to a specific switch port.

VM Policy Bandwidth Control allows the administrator to specify the amount of data the switch will permit to flow from a particular VE, without defining a complicated matrix of ACLs or VMAPs for all port combinations where a VE may appear.

VM Policy Bandwidth Control Commands

VM Policy Bandwidth Control can be configured using the following configuration commands:

```
RS G8124-E(config)# virt vmpolicy vmbwidth <VM MAC>|<index>|<UUID>| <IPv4
address>|<name>?
txrate <committed rate> <burst> [<ACL number>]  (Set the VM transmit bandwidth –
                                                  ingress for switch)
bwctrl                                           (Enable bandwidth control)
```

Bandwidth allocation can be defined for transmit (TX) traffic or receive (RX) traffic. Because bandwidth allocation is specified from the perspective of the VE, and the switch command for TX Rate Control (txrate) sets the data rate to be sent from the VM to the switch.

The *committed rate* is specified in multiples of 64 kbps, from 64 to 40,000,000. The maximum *burst* rate is specified as 32, 64, 128, 256, 1024, 2048, or 4096 kb. If both the committed rate and burst are set to 0, bandwidth control will be disabled.

When txrate is specified, the switch automatically selects an available ACL for internal use with bandwidth control. Optionally, if automatic ACL selection is not desired, a specific ACL may be selected. If there are no unassigned ACLs available, txrate cannot be configured.

Bandwidth Policies vs. Bandwidth Shaping

VM Profile Bandwidth Shaping differs from VM Policy Bandwidth Control.

VM Profile Bandwidth Shaping (see "VM Profiles" on page 227) is configured per VM group and is enforced on the server by a virtual switch in the hypervisor. Shaping is unidirectional and limits traffic transmitted from the virtual switch to the G8124-E. Shaping is performed prior to transmit VM Policy Bandwidth Control. If the egress traffic for a virtual switch port group exceeds shaping parameters, the traffic is dropped by the virtual switch in the hypervisor. Shaping uses server CPU resources, but prevents extra traffic from consuming bandwidth between the server and the G8124-E. Shaping is not supported simultaneously on the same ports as vNICs.

VM Policy Bandwidth Control is configured per VE, and can be set independently for transmit traffic. Bandwidth policies are enforced by the G8124-E. VE traffic that exceeds configured levels is dropped by the switch upon ingress. Setting `txrate` uses ACL resources on the switch.

Bandwidth shaping and bandwidth policies can be used separately or in concert.

VMready Information Displays

The G8124-E can be used to display a variety of VMready information.

Note: Some displays depict information collected from scans of a VMware vCenter and may not be available without a valid vCenter. If a vCenter is assigned (see), scan information might not be available for up to two minutes after the switch boots or when VMready is first enabled. Also, any major changes made through the vCenter may take up to two minutes to be reflected on the switch unless you force an immediate vCenter scan (see .

Local VE Information

A concise list of local VEs and pre-provisioned VEs is available with the following ISCLI privileged EXEC command:

```
RS G8124-E# show virt vm

IP Address      VMAC Address     Index Port     VM Group (Profile)
Check status
---------------- ----------------- ----- -------- ------------------
------------
 0.0.0.0        00:50:56:55:47:0c 5     17.3
~0.0.0.0        00:50:56:b3:1e:7b 2     17.3     1       test
~0.0.0.0        00:50:56:b3:1f:16 1     17.3     1       test
~0.0.0.0        00:50:56:b3:2c:b9 4     18       2
~0.0.0.0        00:50:56:b3:5f:32 3     18       1       test
~0.0.0.0        00:50:56:b3:69:5a 0     19.3     1       test
VMReady ports:  17-21

Number of entries: 6
~ indicates inactive VMs
0.0.0.0 indicates IP address not yet available
```

Note: The Index numbers shown in the VE information displays can be used to specify a particular VE in configuration commands.

If a vCenter is available, more verbose information can be obtained using the following ISCLI privileged EXEC command option:

```
RS G8124-E# show virt vm -v

Index MAC Address, Name (VM or Host), Port, Group Vswitch,
      IP Address @Host (VMs only) VLAN Port Group
----- ------------ ------------------ ----- ----- ----------
 0 00:50:56:ba:1b:23 New Virtual Machine ST 1ST 1 100 vSwitch2
       10.10.10.101 @10.241.5.49 100 Lenovo_vlan100

 2 00:50:56:ba:25:8a VmForGaborII 26 vSwitch1
       10.10.10.101 @10.241.5.49 0 IBM_Default

 3 00:50:56:ba:1b:00 New Virtual Machine 2626 vSwitch1
       0.0.0.0 @10.241.5.49 100 VM Network 2

--
 3 of 3 entries printed
 0.0.0.0 indicates IP Address is not available
 Use the "-v -r" options to refresh data before displaying results

EVB Virtual Station Interface Information:
Total number of VM Association entries :
```

To view additional detail regarding any specific VE, see "vCenter Switchport Mapping Details" on page 244).

vCenter Hypervisor Hosts

If a vCenter is available, the following ISCLI privileged EXEC command displays the name and UUID of all VMware hosts, providing an essential overview of the data center:

```
RS G8124-E# show virt vmware hosts
UUID                                     Name(s), IP Address
----------------------------------------------------------------
00a42681-d0e5-5910-a0bf-bd23bd3f7800  172.16.41.30
002e063c-153c-dd11-8b32-a78dd1909a00  172.16.46.10
00f1fe30-143c-dd11-84f2-a8ba2cd7ae00  172.16.44.50
0018938e-143c-dd11-9f7a-d8defa4b8300  172.16.46.20
...
```

Using the following command, the administrator can view more detailed vCenter host information, including a list of virtual switches and their port groups, as well as details for all associated VEs:

```
RS G8124-E# show virt vmware showhost {<UUID>|<IPv4 address>|<host name>}
Vswitches available on the host:
            vSwitch0
Port Groups and their Vswitches on the host:
            Lenovo_Default                  vSwitch0
            VM Network              vSwitch0
            Service Console         vSwitch0
            VMkernel                vSwitch0
-----------------------------------------------------------------------
MAC Address        00:50:56:9c:21:2f
Port               4
Type               Virtual Machine
VM vCenter Name    halibut
VM OS hostname     localhost.localdomain
VM IP Address      172.16.46.15
VM UUID            001c41f3-ccd8-94bb-1b94-6b94b03b9200
Current VM Host    172.16.46.10
Vswitch            vSwitch0
Port Group         Lenovo_Default
VLAN ID            0
...
```

vCenter VEs

If a vCenter is available, the following ISCLI privileged EXEC command displays a list of all known VEs:

```
RS G8124-E# show virt vmware vms
UUID                              Name(s), IP Address
---------------------------------------------------------------------
001cdf1d-863a-fa5e-58c0-d197ed3e3300  30vm1
001c1fba-5483-863f-de04-4953b5caa700  VM90
001c0441-c9ed-184c-7030-d6a6bc9b4d00  VM91
001cc06e-393b-a36b-2da9-c71098d9a700  vm_new
001c6384-f764-983c-83e3-e94fc78f2c00  sturgeon
001c7434-6bf9-52bd-c48c-a410da0c2300  VM70
001cad78-8a3c-9cbe-35f6-59ca5f392500  VM60
001cf762-a577-f42a-c6ea-090216c11800  30VM6
001c41f3-ccd8-94bb-1b94-6b94b03b9200  halibut, localhost.localdomain,
                                      172.16.46.15
001cf17b-5581-ea80-c22c-3236b89ee900  30vm5
001c4312-a145-bf44-7edd-49b7a2fc3800  vm3
001caf40-a40a-de6f-7b44-9c496f123b00  30VM7
```

vCenter VE Details

If a vCenter is available, the following ISCLI privileged EXEC command displays detailed information about a specific VE:

```
RS G8124-E# show virt vmware showvm {<VM UUID>|<VM IPv4 address>|<VM name>}
---------------------------------------------------------------------
MAC Address        00:50:56:9c:21:2f
Port               4
Type               Virtual Machine
VM vCenter Name    halibut
VM OS hostname     localhost.localdomain
VM IP Address      172.16.46.15
VM UUID            001c41f3-ccd8-94bb-1b94-6b94b03b9200
Current VM Host    172.16.46.10
Vswitch            vSwitch0
Port Group         Lenovo_Default
VLAN ID            0
```

vCenter Switchport Mapping Details

If a vCenter is available, the following ISCLI privileged EXEC command displays detailed information about VE switchport mapping:

```
RS G8124-E# show virt vmware switchport-mapping
ST 5 ==> 10.241.32.133 vmnic5
ST 5 ==> 10.241.32.133 vmnic5
23 ==> 10.241.32.131 vmnic3
```

VMready Configuration Example

This example has the following characteristics:

- A VMware vCenter is fully installed and configured prior to VMready configuration and includes a "bladevm" administration account and a valid SSL certificate.
- The distributed VM group model is used.
- The VM profile named "Finance" is configured for VLAN 30, and specifies NIC-to-switch bandwidth shaping for 1Mbps average bandwidth, 2MB bursts, and 3Mbps maximum bandwidth.
- The VM group includes four discovered VMs on switch server ports 1 and 2, and one static LAG (previously configured) that includes switch uplink ports 3 and 4.

1. Define the server ports.

```
RS G8124-E(config)# system server-ports port 1-2
```

2. Enable the VMready feature.

```
RS G8124-E(config)# virt enable
```

3. Specify the VMware vCenter IPv4 address.

```
RS G8124-E(config)# virt vmware vmware vcspec 172.16.100.1 bladevm
```

When prompted, enter the user password that the switch must use for access to the vCenter.

4. Create the VM profile.

```
RS G8124-E(config)# virt vmprofile Finance
RS G8124-E(config)# virt vmprofile edit Finance vlan 30
RS G8124-E(config)# virt vmprofile edit Finance shaping 1000 2000 3000
```

5. Define the VM group.

```
RS G8124-E(config)# virt vmgroup 1 profile Finance
RS G8124-E(config)# virt vmgroup 1 vm arctic
RS G8124-E(config)# virt vmgroup 1 vm monster
RS G8124-E(config)# virt vmgroup 1 vm sierra
RS G8124-E(config)# virt vmgroup 1 vm 00:50:56:4f:f2:00
RS G8124-E(config)# virt vmgroup 1 portchannel 1
```

When VMs are added, the server ports on which they appear are automatically added to the VM group. In this example, there is no need to manually add ports 1 and 2.

Note: VM groups and vNICs (see Chapter 14, "Virtual NICs") are not supported simultaneously on the same switch ports.

6. If necessary, enable VLAN tagging for the VM group:

```
RS G8124-E(config)# virt vmgroup 1 tag
```

Note: If the VM group contains ports that also exist in other VM groups, make sure tagging is enabled in both VM groups. In this example configuration, no ports exist in more than one VM group.

7. Save the configuration.

Chapter 16. FCoE and CEE

This chapter provides conceptual background and configuration examples for using Converged Enhanced Ethernet (CEE) features of the RackSwitch G8124-E, with an emphasis on Fibre Channel over Ethernet (FCoE) solutions. The following topics are addressed in this chapter:

- "Fibre Channel over Ethernet" on page 248

  Fibre Channel over Ethernet (FCoE) allows Fibre Channel traffic to be transported over Ethernet links. This provides an evolutionary approach toward network consolidation, allowing Fibre Channel equipment and tools to be retained, while leveraging cheap, ubiquitous Ethernet networks for growth.

- "Converged Enhanced Ethernet" on page 250

  Converged Enhanced Ethernet (CEE) refers to a set of IEEE standards developed primarily to enable FCoE, requiring enhancing the existing Ethernet standards to make them lossless on a per-priority traffic basis, and providing a mechanism to carry converged (LAN/SAN/IPC) traffic on a single physical link. CEE features can also be utilized in traditional LAN (non-FCoE) networks to provide lossless guarantees on a per-priority basis, and to provide efficient bandwidth allocation.

- "FCoE Initialization Protocol Snooping" on page 253

  Using FCoE Initialization Protocol (FIP) snooping, the G8124-E examines the FIP frames exchanged between ENodes and FCFs. This information is used to dynamically determine the ACLs required to block certain types of undesired or unvalidated traffic on FCoE links.

- "Priority-Based Flow Control" on page 259

  Priority-Based Flow Control (PFC) extends 802.3x standard flow control to allow the switch to pause traffic based on the 802.1p priority value in each packet's VLAN tag. PFC is vital for FCoE environments, where SAN traffic must remain lossless and must be paused during congestion, while LAN traffic on the same links is delivered with "best effort" characteristics.

- "Enhanced Transmission Selection" on page 262

  Enhanced Transmission Selection (ETS) provides a method for allocating link bandwidth based on the 802.1p priority value in each packet's VLAN tag. Using ETS, different types of traffic (such as LAN, SAN, and management) that are sensitive to different handling criteria can be configured either for specific bandwidth characteristics, low-latency, or best-effort transmission, despite sharing converged links as in an FCoE environment.

- "Data Center Bridging Capability Exchange" on page 269

  Data Center Bridging Capability Exchange Protocol (DCBX) allows neighboring network devices to exchange information about their capabilities. This is used between CEE-capable devices for the purpose of discovering their peers, negotiating peer configurations, and detecting misconfigurations.

Fibre Channel over Ethernet

Fibre Channel over Ethernet (FCoE) is an effort to converge two of the different physical networks in today's data centers. It allows Fibre Channel traffic (such as that commonly used in Storage Area Networks, or SANs) to be transported without loss over 10Gb Ethernet links (typically used for high-speed Local Area Networks, or LANs). This provides an evolutionary approach toward network consolidation, allowing Fibre Channel equipment and tools to be retained, while leveraging cheap, ubiquitous Ethernet networks for growth.

With server virtualization, servers capable of hosting both Fibre Channel and Ethernet applications will provide advantages in server efficiency, particularly as FCoE-enabled network adapters provide consolidated SAN and LAN traffic capabilities.

The RackSwitch G8124-E with Lenovo Networking OS 8.3 software is compliant with the INCITS T11.3, FC-BB-5 FCoE specification.

Note: The G8124-E supports up to 384 FCoE login sessions.

The FCoE Topology

In an end-to-end Fibre Channel network, switches and end devices generally establish trusted, point-to-point links. Fibre Channel switches validate end devices, enforce zoning configurations and device addressing, and prevent certain types of errors and attacks on the network.

In a converged multi-hop FCoE network where Fibre Channel devices are bridged to Ethernet devices, the direct point-to-point QoS capabilities normally provided by the Fibre Channel fabric may be lost in the transition between the different network types. The G8124-E provides a solution to overcome this.

Figure 28. A Mixed Fibre Channel and FCoE Network
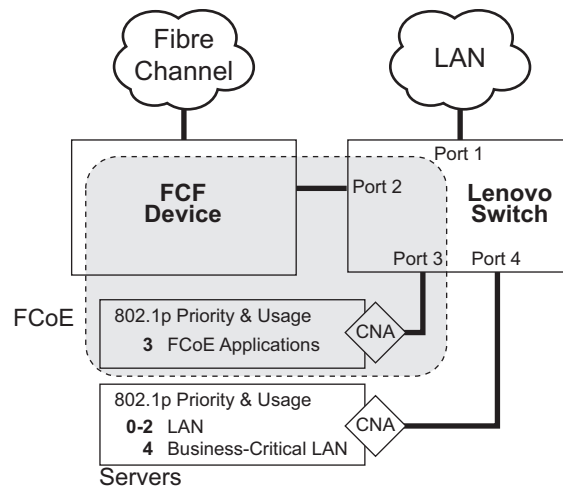
In Figure 28 on page 248, the Fibre Channel network is connected to the FCoE network through an FCoE Forwarder (FCF). The FCF acts as a Fibre Channel gateway to and from the FCoE network.

For the FCoE portion of the network, the FCF is connected to the FCoE-enabled G8124-E, which is connected to a server (running Fibre Channel applications) through an FCoE-enabled Converged Network Adapter (CNA) known in Fibre Channel as Ethernet Nodes (ENodes).

Note: The figure also shows a non-FCoE LAN server connected to the G8124-E using a CNA. This allows the LAN server to take advantage of some CEE features that are useful even outside of an FCoE environment.

To block undesired or unvalidated traffic on FCoE links that exists outside the regular Fibre Channel topology, Ethernet ports used in FCoE are configured with Access Control Lists (ACLs) that are narrowly tailored to permit expected FCoE traffic to and from confirmed FCFs and ENodes, and deny all other FCoE or FIP traffic. This ensures that all FCoE traffic to an from the ENode passes through the FCF.

Because manual ACL configuration is an administratively complex task, the G8124-E can automatically and dynamically configure the ACLs required for use with FCoE. Using FCoE Initialization Protocol (FIP) snooping (see "FCoE Initialization Protocol Snooping" on page 253), the G8124-E examines the FIP frames normally exchanged between the FCF and ENodes to determine information about connected FCoE devices. This information is used to automatically determine the appropriate ACLs required to block certain types of undesired or unvalidated FCoE traffic.

Automatic FCoE-related ACLs are independent from ACLs used for typical Ethernet purposes.

FCoE Requirements

The following are required for implementing FCoE using the G8124-E with N/OS 8.3 software:

- The G8124-E must be connected to the Fibre Channel network using an external FCF such as a Lenovo Rackswitch G8264CS or a Cisco Nexus 5000 Series Switch.

- For each G8124-E port participating in FCoE, the connected server must use the supported Converged Network Adapter (CNA) and have the FCoE license enabled (if applicable) on the CNA.

- CEE must be turned on (see "Turning CEE On or Off" on page 250). When CEE is on, the DCBX, PFC, and ETS features are enabled and configured with default FCoE settings. These features may be reconfigured, but must remain enabled for FCoE to function.

- FIP snooping must be turned on (see "FCoE Initialization Protocol Snooping" on page 253). When FIP snooping is turned on, the feature is enabled on all ports by default. The administrator can disable FIP snooping on individual ports that do not require FCoE, but FIP snooping must remain enabled on all FCoE ports for FCoE to function.

Note: FCoE and vNICs (see Chapter 14, "Virtual NICs") are not supported simultaneously on the same G8124-E.

Converged Enhanced Ethernet

Converged Enhanced Ethernet (CEE) refers to a set of IEEE standards designed to allow different physical networks with different data handling requirements to be converged together, simplifying management, increasing efficiency and utilization, and leveraging legacy investments without sacrificing evolutionary growth.

CEE standards were developed primarily to enable Fibre Channel traffic to be carried over Ethernet networks. This required enhancing the existing Ethernet standards to make them lossless on a per-priority traffic basis, and to provide a mechanism to carry converged (LAN/SAN/IPC) traffic on a single physical link. Although CEE standards were designed with FCoE in mind, they are not limited to FCoE installations. CEE features can be utilized in traditional LAN (non-FCoE) networks to provide lossless guarantees on a per-priority basis, and to provide efficient bandwidth allocation based on application needs.

Turning CEE On or Off

By default on the G8124-E, CEE is turned off. To turn CEE on or off, use the following CLI commands:

```
RS G8124-E(config)# [no] cee enable
```

⚠️

CAUTION:
Turning CEE on will automatically change some 802.1p QoS and 802.3x standard flow control settings on the G8124-E. Read the following material carefully to determine whether you will need to take action to reconfigure expected settings.

It is recommended that you backup your configuration prior to turning CEE on. Viewing the file will allow you to manually re-create the equivalent configuration once CEE is turned on, and will also allow you to recover your prior configuration if you need to turn CEE off.

Effects on Link Layer Discovery Protocol

When CEE is turned on, Link Layer Discovery Protocol (LLDP) is automatically turned on and enabled for receiving and transmitting DCBX information. LLDP cannot be turned off while CEE is turned on.

Effects on 802.1p Quality of Service

While CEE is off (the default), the G8124-E allows 802.1p priority values to be used for Quality of Service (QoS) configuration (see page 200). 802.1p QoS default settings are shown in Table 20, but can be changed by the administrator.

When CEE is turned on, 802.1p QoS is replaced by ETS (see "Enhanced Transmission Selection" on page 262). As a result, while CEE is turned on, using 802.1p QoS configuration commands causes an error message to appear saying that the `cee global ets` command must be used to configure 802.1p-based QoS.

In addition, when CEE is turned on, prior 802.1p QoS settings are replaced with new defaults designed for use with ETS priority groups (PGIDs) as shown in Table 20:

Table 20. *CEE Effects on 802.1p Defaults*

802.1p QoS Configuration With CEE Off (default)				ETS Configuration With CEE On		
Priority-	COSq	Weight		Priority	COSq	PGID
0	0	1		0	2	2
1	1	2		1	2	2
2	2	3		2	2	2
3	3	4		3	3	3
4	4	5		4	4	4
5	5	7		5	4	4
6	6	15		6	4	4
7	7	0		7	7	15

When CEE is on, the default ETS configuration also allocates a portion of link bandwidth to each PGID as shown in Table 21:

Table 21. *Default ETS Bandwidth Allocation*

PGID	Typical Use	Bandwidth
2	LAN	10%
3	SAN	50%
4	Latency-sensitive LAN	40%

If the prior, non-CEE configuration used 802.1p priority values for different purposes, or does not expect bandwidth allocation as shown in Table 21 on page 251, when CEE is turned on, have the administrator reconfigure ETS settings as appropriate.

It is recommended that a configuration backup be made prior to turning CEE on or off. Viewing the configuration file will allow the administrator to manually re-create the equivalent configuration under the new CEE mode, and will also allow for the recovery of the prior configuration if necessary.

Effects on Flow Control

When CEE is off (the default), 802.3x standard flow control is enabled on all switch ports by default.

When CEE is turned on, standard flow control is disabled on all ports, and in its place, PFC (see "Priority-Based Flow Control" on page 259) is enabled on all ports for 802.1p priority value 3. This default is chosen because priority value 3 is commonly used to identify FCoE traffic in a CEE environment and must be guaranteed lossless behavior. PFC is disabled for all other priority values.

It is recommend that a configuration backup be made prior to turning CEE on or off. Viewing the configuration file will allow the administrator to manually re-create the equivalent configuration under the new CEE mode, and will also allow for the recovery of the prior configuration if necessary.

When CEE is on, PFC can be enabled only on priority value 3 and one other priority. If flow control is required on additional priorities on any given port, consider using standard flow control on that port, so that regardless of which priority traffic becomes congested, a flow control frame is generated.

FCoE Initialization Protocol Snooping

FCoE Initialization Protocol (FIP) snooping is an FCoE feature. To enforce point-to-point links for FCoE traffic outside the regular Fibre Channel topology, Ethernet ports used in FCoE can be automatically and dynamically configured with Access Control Lists (ACLs).

Using FIP snooping, the G8124-E examines the FIP frames normally exchanged between the FCF and ENodes to determine information about connected FCoE devices. This information is used to create narrowly tailored ACLs that permit expected FCoE traffic to and from confirmed Fibre Channel nodes, and deny all other undesirable FCoE or FIP traffic.

In case of LAGs, FIP traffic from a particular FCF can be received by any member port on which the FCF was detected.

FIP Snooping Requirements

The following are required for implementing the FIP snooping bridge feature using the G8124-E with N/OS 8.3 software:

- The G8124-E must be connected to the Fibre Channel network through an FCF such as a Lenovo Rackswitch G8264CS or a Cisco Nexus 5000 Series Switch.

- For each G8124-E port participating in FCoE, the connected server must use a FCoE-licensed Converged Network Adapter (CNA) and have the FCoE license enabled (if applicable) on the CNA.

- CEE must be turned on (see "Turning CEE On or Off" on page 250). When CEE is on, the DCBX, PFC, and ETS features are enabled and configured with default FCoE settings. These features may be reconfigured, but must remain enabled for FCoE to function.

- FIP snooping must be turned on (see "FCoE Initialization Protocol Snooping" on page 253). When FIP snooping is turned on, the feature is enabled on all ports by default. The administrator can disable FIP snooping on individual ports that do not require FCoE, but FIP snooping must remain enabled on all FCoE ports for FCoE to function.

Note: FCoE and vNICs (see Chapter 14, "Virtual NICs") are not supported simultaneously on the same G8124-E.

Port Trunking

Lenovo N/OS 8.3 supports port trunking for FCoE connections. The Link Aggregation (LAG) can be used for separate FCoE traffic, or for Ethernet and FCoE traffic. Ports directly connected to servers cannot be combined in a LAG group.

Uplink ports, connected to the FCF, can be grouped as static or dynamic trunks.

Normal trunk operations such as creating or enabling the trunk and adding or removing member ports can be performed. When a port is added to a trunk group, FCFs previously detected on the port will be deleted. The deleted FCF may be relearned later. However, this may cause flickering in the network traffic. We recommend that you make trunk group changes, if any, prior to live FCoE traffic.

Data Center Bridging (DCBX) is configured on a per-port basis. Each port in a trunk must have the same ETS, PFC, and DCBX configuration. When a port ceases to be the trunk group member, its configuration does not change.

Note: If the ports chosen to make a LAG or a trunk do not have the same configuration of PFC, ETS, or DCBX, the switch will throw an error.

Global FIP Snooping Settings

By default, the FIP snooping feature is turned off for the G8124-E. The following commands are used to turn the feature on or off:

```
RS G8124-E(config)# [no] fcoe fips enable
```

Note: FIP snooping requires CEE to be turned on (see "Turning CEE On or Off" on page 250).

When FIP snooping is on, port participation may be configured on a port-by-port basis (see the next sections).

When FIP snooping is off, all FCoE-related ACLs generated by the feature are removed from all switch ports.

FIP snooping configuration must be the same on all member ports in a LAG. If the configuration of a member port is changed, an error message, similar to the following, will be displayed.

```
"FAIL: Trunk X FIP Snooping port Yand port Z need to have the same fcf
mode config"
```

The configuration changes are applied to all member ports in a LAG.

FIP Snooping for Specific Ports

When FIP snooping is globally turned on (see the previous section), ports may be individually configured for participation in FIP snooping and automatic ACL generation. By default, FIP snooping is enabled for each port. To change the setting for any specific port, use the following CLI commands:

```
RS G8124-E(config)# [no] fcoe fips port <port alias, number, list, or range> enable
```

When FIP snooping is enabled on a port, FCoE-related ACLs will be automatically configured.

When FIP snooping is disabled on a port, all FCoE-related ACLs on the port are removed, and the switch will enforce no FCoE-related rules for traffic on the port.

Port FCF and ENode Detection

When FIP snooping is enabled on a port, the port is placed in FCF auto-detect mode by default. In this mode, the port assumes connection to an ENode unless FIP packets show the port is connected to an FCF.

Ports can also be specifically configured as to whether automatic FCF detection will be used, or whether the port is connected to an FCF or ENode:

```
RS G8124-E(config)# fcoe fips port <port alias, number, list, or range> fcf-mode
{auto|on|off}
```

When FCF mode is on, the port is assumed to be connected to a trusted FCF, and only ACLs appropriate to FCFs will be installed on the port. When off, the port is assumed to be connected to an ENode, and only ACLs appropriate to ENodes will be installed. When the mode is changed (either through manual configuration or as a result of automatic detection), the appropriate ACLs are automatically added, removed, or changed to reflect the new FCF or ENode connection.

FCoE Connection Timeout

FCoE-related ACLs are added, changed, and removed as FCoE device connection and disconnection are discovered. In addition, the administrator can enable or disable automatic removal of ACLs for FCFs and other FCoE connections that timeout (fail or are disconnected) without FIP notification.

By default, automatic removal of ACLs upon timeout is enabled. To change this function, use the following CLI command:

```
RS G8124-E(config)# [no] fcoe fips timeout-acl
```

FCoE ACL Rules

When FIP Snooping is enabled on a port, the switch automatically installs the appropriate ACLs to enforce the following rules for FCoE traffic:

- Ensure that FIP frames from ENodes may only be addressed to FCFs.
- Flag important FIP packets for switch processing.
- Ensure no end device uses an FCF MAC address as its source.
- Each FCoE port is assumed to be connected to an ENode and include ENode-specific ACLs installed, until the port is either detected or configured to be connected to an FCF.
- Ports that are configured to have FIP snooping disabled will not have any FIP or FCoE related ACLs installed.
- Prevent transmission of all FCoE frames from an ENode prior to its successful completion of login (FLOGI) to the FCF.
- After successful completion of FLOGI, ensure that the ENode uses only those FCoE source addresses assigned to it by FCF.
- After successful completion of FLOGI, ensure that all ENode FCoE source addresses originate from or are destined to the appropriate ENode port.
- After successful completion of each FLOGI, ensure that FCoE frames may only be addressed to the FCFs that accept them.

Initially, a basic set of FCoE-related ACLs will be installed on all ports where FIP snooping is enabled. As the switch encounters FIP frames and learns about FCFs and ENodes that are attached or disconnect, ACLs are dynamically installed or expanded to provide appropriate security.

When an FCoE connection logs out, or times out (if ACL timeout is enabled), the related ACLs will be automatically removed.

FCoE-related ACLs are independent of manually configured ACLs used for regular Ethernet purposes (see Chapter 6, "Access Control Lists"). FCoE ACLs generally have a higher priority over standard ACLs, and do not inhibit non-FCoE and non-FIP traffic.

FCoE VLANs

FCoE packets to any FCF will be confined to the VLAN advertised by the FCF (typically VLAN 1002). The appropriate VLAN must be configured on the switch with member FCF ports and must be supported by the participating CNAs. You must manually configure the tag settings and the appropriate VLANs for the Enode ports.

Viewing FIP Snooping Information

ACLs automatically generated under FIP snooping are independent of regular, manually configure ACLs, and are not listed with regular ACLs in switch information and statistics output. Instead, FCoE ACLs are shown using the following CLI commands:

```
RS G8124-E# show fcoe fips information (Show all FIP-related information)
RS G8124-E# show fcoe fips port <ports> information    (Show FIP info for a selected port)
```

For example:

```
RS G8124-E# show fcoe fips port 21 information

FIP Snooping on port 21:
This port has been configured to automatically detect FCF.
 It has currently detected to have 0 FCF connecting to it.

FIPS ACLs configured on this port:
SMAC 00:05:73:ce:96:67, action deny.
DMAC 00:05:73:ce:96:67, ethertype 0x8914, action permit.
SMAC 0e:fc:00:44:04:04, DMAC 00:05:73:ce:96:67, ethertype 0x8906, vlan
1002, action permit.
DMAC 01:10:18:01:00:01, Ethertype 0x8914, action permit.
DMAC 01:10:18:01:00:02, Ethertype 0x8914, action permit.
Ethertype 0x8914, action deny.
Ethertype 0x8906, action deny.
SMAC 0e:fc:00:00:00:00, SMAC mask ff:ff:ff:00:00:00, action deny.
```

For each ACL, the required traffic criteria are listed, along with the action taken (permit or deny) for matching traffic. ACLs are listed in order of precedence and evaluated in the order shown.

The administrator can also view other FCoE information:

```
RS G8124-E# show fcoe fips fcf          (Show all detected FCFs)
RS G8124-E# show fcoe fips fcoe         (Show all FCoE connections)
```

Operational Commands

The administrator may use the operational commands to delete FIP-related entries from the switch.

To delete a specific FCF entry and all associated ACLs from the switch, use the following command:

```
RS G8124-E# no fcoe fips fcf <FCF MAC address> [<VLAN number>]
```

FIP Snooping Configuration

In this example, as shown in Figure 28 on page 248, FCoE devices are connected to port 2 for the FCF device, and port 3 for an ENode. FIP snooping can be configured on these ports using the following ISCLI commands:

1. Enable VLAN tagging on the FCoE ports:

```
RS G8124-E(config)# interface port 2,3 ( Select FCoE ports)
RS G8124-E(config-if)# switchport mode trunk( Enable VLAN tagging)
RS G8124-E(config-if)# exit            (Exit port configuration mode)
```

Note: If you are using Emulex CNA BE 2 - FCoE mode, you must enable PVID tagging on the Enode ports.

2. Place FCoE ports into a VLAN supported by the FCF and CNAs (typically VLAN 1002):

```
RS G8124-E(config)# vlan 1002                     (Select a VLAN)
RS G8124-E(config-vlan)# exit                     (Exit VLAN configuration mode)
RS G8124-E(config)# interface port 2,3            (Add FCoE ports to the VLAN)
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan add 1002
RS G8124-E(config-if)# exit
```

Note: Placing ports into the VLAN (Step 2) *after* tagging is enabled (Step 1) helps to ensure that their port VLAN ID (PVID) is not accidentally changed.

3. Turn CEE on.

```
RS G8124-E(config)# cee enable
```

Note: Turning CEE on will automatically change some 802.1p QoS and 802.3x standard flow control settings and menus (see "Turning CEE On or Off" on page 250).

4. Turn global FIP snooping on:

```
RS G8124-E(config)# fcoe fips enable
```

5. If using an Emulex CNA, disable automatic VLAN creation.

```
RS G8124-E(config)# no fcoe fips automatic-vlan
```

6. Enable FIP snooping on FCoE ports, and set the desired FCF mode:

```
RS G8124-E(config)# fcoe fips port 2 enable        (Enable FIPS on port 2)
RS G8124-E(config)# fcoe fips port 2 fcf-mode on   (Set as FCF connection)
RS G8124-E(config)# fcoe fips port 3 enable        (Enable FIPS on port 3)
RS G8124-E(config)# fcoe fips port 3 fcf-mode off  (Set as ENode connection)
```

Note: By default, FIP snooping is enabled on all ports and the FCF mode set for automatic detection. The configuration in this step is unnecessary, if default settings have not been changed, and is shown merely as a manual configuration example.

7. Save the configuration.

Priority-Based Flow Control

Priority-based Flow Control (PFC) is defined in IEEE 802.1Qbb. PFC extends the IEEE 802.3x standard flow control mechanism. Under standard flow control, when a port becomes busy, the switch manages congestion by pausing all the traffic on the port, regardless of the traffic type. PFC provides more granular flow control, allowing the switch to pause specified types of traffic on the port, while other traffic on the port continues.

PFC pauses traffic based on 802.1p priority values in the VLAN tag. The administrator can assign different priority values to different types of traffic and then enable PFC for up to two specific priority values: priority value 3, and one other. The configuration can be applied globally for all ports on the switch. Then, when traffic congestion occurs on a port (caused when ingress traffic exceeds internal buffer thresholds), only traffic with priority values where PFC is enabled is paused. Traffic with priority values where PFC is disabled proceeds without interruption but may be subject to loss if port ingress buffers become full.

Although PFC is useful for a variety of applications, it is required for FCoE implementation where storage (SAN) and networking (LAN) traffic are converged on the same Ethernet links. Typical LAN traffic tolerates Ethernet packet loss that can occur from congestion or other factors, but SAN traffic must be lossless and requires flow control.

For FCoE, standard flow control would pause both SAN and LAN traffic during congestion. While this approach would limit SAN traffic loss, it could degrade the performance of some LAN applications that expect to handle congestion by dropping traffic. PFC resolves these FCoE flow control issues. Different types of SAN and LAN traffic can be assigned different IEEE 802.1p priority values. PFC can then be enabled for priority values that represent SAN and LAN traffic that must be paused during congestion, and disabled for priority values that represent LAN traffic that is more loss-tolerant.

PFC requires CEE to be turned on (). When CEE is turned on, PFC is enabled on priority value 3 by default. Optionally, the administrator can also enable PFC on one other priority value, providing lossless handling for another traffic type, such as for a business-critical LAN application.

Note: For any given port, only one flow control method can be implemented at any given time: either PFC or standard IEEE 802.3x flow control.

Global Configuration

PFC requires CEE to be turned on ("Turning CEE On or Off" on page 250). When CEE is turned on, standard flow control is disabled on all ports, and PFC is enabled on all ports for 802.1p priority value 3. While CEE is turned on, PFC cannot be disabled for priority value 3. This default is chosen because priority value 3 is commonly used to identify FCoE traffic in a CEE environment and must be guaranteed lossless behavior. PFC is disabled for all other priority values by default, but can be enabled for one additional priority value.

- Global PFC configuration is preferable in networks that implement end-to-end CEE devices. For example, if all ports are involved with FCoE and can use the same SAN and LAN priority value configuration with the same PFC settings, global configuration is easy and efficient.

- Global PFC configuration can also be used in some mixed environments where traffic with PFC-enabled priority values occurs only on ports connected to CEE devices, and not on any ports connected to non-CEE devices. In such cases, PFC can be configured globally on specific priority values even though not all ports make use them.

- PFC is not restricted to CEE and FCoE networks. In any LAN where traffic is separated into different priorities, PFC can be enabled on priority values for loss-sensitive traffic.

- If you want to enable PFC on a priority, add the priority to a priority group with the same number. For example, if you want to enable PFC on priority 0, you must map priority 0 to PG 0.

Note: When using global PFC configuration in conjunction with the ETS feature (see "Enhanced Transmission Selection" on page 262), ensure that only pause-tolerant traffic (such as lossless FCoE traffic) is assigned priority values where PFC is enabled. Pausing other types of traffic can have adverse effects on LAN applications that expect uninterrupted traffic flow and tolerate dropping packets during congestion.

PFC Configuration Example

Note: DCBX may be configured to permit sharing or learning PFC configuration with or from external devices. This example assumes that PFC configuration is being performed manually. See "Data Center Bridging Capability Exchange" on page 269 for more information on DCBX. Even if the G8124-E learns the PFC configuration from a DCBX peer, the PFC configuration must be performed manually.

This example is consistent with the network shown in Figure 28 on page 248. In this example, the following topology is used.

Table 22. *PFC Configuration*

Switch Port	802.1p Priority	Usage	PFC Setting
1-4	0-2	LAN	Disabled
	3	FCoE	**Enabled**
	4	Business-critical LAN	**Enabled**
	others	(not used)	Disabled

In this example, PFC is to facilitate lossless traffic handling for FCoE (priority value 3) and a business-critical LAN application (priority value 4).

Assuming that CEE is off (the G8124-E default), the example topology shown in the table above can be configured using the following commands:

1. Turn CEE on.

```
RS G8124-E(config)# cee enable
```

Note: Turning CEE on will automatically change some 802.1p QoS and 802.3x standard flow control settings and menus (see "Turning CEE On or Off" on page 250).

2. Enable PFC for the FCoE traffic.

Note: PFC is enabled on priority 3 by default. If using the defaults, the manual configuration commands shown in this step are not necessary.

```
RS G8124-E(config)# cee global pfc priority 3 enable(Enable on FCoE priority)
RS G8124-E(config)# cee global pfc priority 3 description "FCoE"
                                             (Optional description)
```

3. Enable PFC for the business-critical LAN application:

```
RS G8124-E(config)# cee global ets priority-group pgid 5 priority 5,6
pgid 4 bandwidth 20 pgid 5 bandwidth 20
RS G8124-E(config)# cee global pfc priority 4 enable( Enable on LAN priority)
RS G8124-E(config)# cee global pfc priority 4 description "Critical LAN"
                                             (Optional description)
```

4. Save the configuration.

Enhanced Transmission Selection

Enhanced Transmission Selection (ETS) is defined in IEEE 802.1Qaz. ETS provides a method for allocating port bandwidth based on 802.1p priority values in the VLAN tag. Using ETS, different amounts of link bandwidth can specified for different traffic types (such as for LAN, SAN, and management).

ETS is an essential component in a CEE environment that carries different types of traffic, each of which is sensitive to different handling criteria, such as Storage Area Networks (SANs) that are sensitive to packet loss, and LAN applications that may be latency-sensitive. In a single converged link, such as when implementing FCoE, ETS allows SAN and LAN traffic to coexist without imposing contrary handling requirements upon each other.

The ETS feature requires CEE to be turned on (see).

802.1p Priority Values

Under the 802.1p standard, there are eight available priority values, with values numbered 0 through 7, which can be placed in the priority field of the 802.1Q VLAN tag:

16 bits	3 bits	1	12 bits
Tag Protocol ID (0x8100)	Priority	CFI	VLAN ID

0   15   16   32
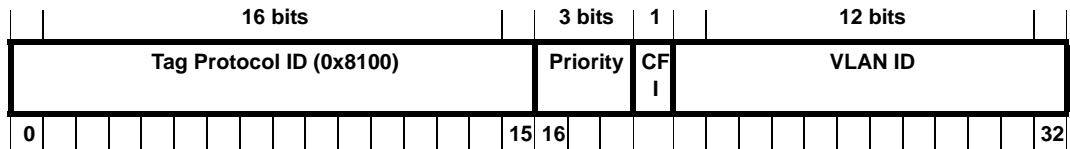
Servers and other network devices may be configured to assign different priority values to packets belonging to different traffic types (such as SAN and LAN).

ETS uses the assigned 802.1p priority values to identify different traffic types. The various priority values are assigned to priority groups (PGID), and each priority group is assigned a portion of available link bandwidth.

Priorities values within in any specific ETS priority group are expected to have similar traffic handling requirements with respect to latency and loss.

802.1p priority values may be assigned by the administrator for a variety of purposes. However, when CEE is turned on, the G8124-E sets the initial default values for ETS configuration as follows:

Figure 29. Default ETS Priority Groups

Typical Traffic Type	802.1p Priority	PGID	Bandwidth Allocation
LAN	0		
LAN	1	2	10%
LAN	2		
SAN	3	3	50%
Latency-Sensitive LAN	4		
Latency-Sensitive LAN	5	4	40%
Latency-Sensitive LAN	6		
Network Management	7	15	-

In the assignment model shown in Figure 29 on page 262, priorities values 0 through 2 are assigned for regular Ethernet traffic, which has "best effort" transport characteristics.

Because CEE and ETS features are generally associated with FCoE, Priority 3 is typically used to identify FCoE (SAN) traffic.

Priorities 4-7 are typically used for latency sensitive traffic and other important business applications. For example, priority 4 and 5 are often used for video and voice applications such as IPTV, Video on Demand (VoD), and Voice over IP (VoIP). Priority 6 and 7 are often used for traffic characterized with a "must get there" requirement, with priority 7 used for network control which is requires guaranteed delivery to support configuration and maintenance of the network infrastructure.

Note: The default assignment of 802.1p priority values on the G8124-E changes depending on whether CEE is on or off. See "Turning CEE On or Off" on page 250 for details.

Priority Groups

For ETS use, each 801.2p priority value is assigned to a priority group which can then be allocated a specific portion of available link bandwidth. To configure a priority group, the following is required:

- CEE must be turned on ("Turning CEE On or Off" on page 250) for the ETS feature to function.
- A priority group must be assigned a priority group ID (PGID), one or more 802.1p priority values, and allocated link bandwidth greater than 9%.

PGID

Each priority group is identified with number (0 through 7, and 15) known as the PGID.

PGID 0 through 7 may each be assigned a portion of the switch's available bandwidth.

PGID 8 through 14 are reserved as per the 802.1Qaz ETS standard.

PGID 15 is a strict priority group. It is generally used for critical traffic, such as network management. Any traffic with priority values assigned to PGID 15 is permitted as much bandwidth as required, up to the maximum available on the switch. After serving PGID 15, any remaining link bandwidth is shared among the other groups, divided according to the configured bandwidth allocation settings.

Make sure all 802.1p priority values assigned to a particular PGID have similar traffic handling requirements. For example, PFC-enabled traffic must not be grouped with non-PFC traffic. Also, traffic of the same general type must be assigned to the same PGID. Splitting one type of traffic into multiple 802.1p priorities, and then assigning those priorities to different PGIDs may result in unexpected network behavior.

Each 802.1p priority value may be assigned to only one PGID. However, each PGID may include multiple priority values. Up to eight PGIDs may be configured at any given time. However, no more than three ETS Priority Groups may include priority values for which PFC is disabled.

Assigning Priority Values to a Priority Group

Each priority group may be configured from its corresponding ETS Priority Group, available using the following command:

```
RS G8124-E(config)# cee global ets priority-group pgid <group number (0-7, or
    15)> priority <priority list>
```

where *priority list* is one or more 802.1p priority values (with each separated by a comma). For example, to assign priority values 0 through 2:

```
RS G8124-E(config)# cee global ets priority-group pgid <group number (0-7, or
    15)> priority 0,1,2
```

Note: Within any specific PGID, the PFC settings (see "Priority-Based Flow Control" on page 259) must be the same (enabled or disabled) for all priority values within the group. PFC can be enabled only on priority value 3 and one other priority. If the PFC setting is inconsistent within a PGID, an error is reported when attempting to apply the configuration. Also, no more than three ETS Priority Groups may include priority values for which PFC is disabled.

When assigning priority values to a PGID, the specified priority value will be automatically removed from its old group and assigned to the new group when the configuration is applied.

Each priority value must be assigned to a PGID. Priority values may not be deleted or unassigned. To remove a priority value from a PGID, it must be moved to another PGID.

For PGIDs 0 through 7, bandwidth allocation can also be configured through the ETS Priority Group menu. See for "Allocating Bandwidth" on page 264 for details.

Deleting a Priority Group

A priority group is automatically deleted when it contains no associated priority values, and its bandwidth allocation is set to 0%.

Note: The total bandwidth allocated to PGID 0 through 7 must equal exactly 100%. Reducing the bandwidth allocation of any group will require increasing the allocation to one or more of the other groups (see "Allocating Bandwidth" on page 264).

Allocating Bandwidth

Allocated Bandwidth for PGID 0 Through 7

The administrator may allocate a portion of the switch's available bandwidth to PGIDs 0 through 7. Available bandwidth is defined as the amount of link bandwidth that remains after priorities within PGID 15 are serviced (see

), and assuming that all PGIDs are fully subscribed. If any PGID does not fully consume its allocated bandwidth, the unused portion is made available to the other priority groups.

Priority group bandwidth allocation can be configured using the following command:

```
RS G8124-E(config)# cee global ets priority-group pgid <priority group number>
   bandwidth <bandwidth allocation>pgid
```

where *bandwidth allocation* represents the percentage of link bandwidth, specified as a number between 10 and 100, in 1% increments, or 0.

The following bandwidth allocation rules apply:
● Bandwidth allocation must be 0% for any PGID that has no assigned 802.1p priority values.
● Any PGID assigned one or more priority values must have a bandwidth allocation greater than 9%.
● Total bandwidth allocation for groups 0 through 7 must equal exactly 100%. Increasing or reducing the bandwidth allocation of any PGID also requires adjusting the allocation of other PGIDs to compensate.
● The total bandwidth allocated to all PGIDs which include priority values where PFC is disabled must not exceed 50%.

If these conditions are not met, the switch will report an error when applying the configuration.

To achieve a balanced bandwidth allocation among the various priority groups, packets are scheduled according to a weighted deficit round-robin (WDRR) algorithm. WDRR is aware of packet sizes, which can vary significantly in a CEE environment, making WDRR more suitable than a regular weighted round-robin (WRR) method, which selects groups based only on packet counts.

Note: Actual bandwidth used by any specific PGID may vary from configured values by up to 10% of the available bandwidth in accordance with 802.1Qaz ETS standard. For example, a setting of 10% may be served anywhere from 0% to 20% of the available bandwidth at any given time.

Unlimited Bandwidth for PGID 15

PGID 15 is permitted unlimited bandwidth and is generally intended for critical traffic (such as switch management). Traffic in this group is given highest priority and is served before the traffic in any other priority group.

If PGID 15 has low traffic levels, most of the switch's bandwidth will be available to serve priority groups 0 through 7. However, if PGID 15 consumes a larger part of the switch's total bandwidth, the amount available to the other groups is reduced.

Note: Consider traffic load when assigning priority values to PGID 15. Heavy traffic in this group may restrict the bandwidth available to other groups.

Configuring ETS

Consider an example consistent with that used forPFC configuration (on page 261):1

Table 23. *ETS Configuration*

Priority	Usage	PGID	Bandwidth
0	LAN (best effort delivery)	2	10%
1	LAN (best effort delivery)		
2	LAN (best effort delivery)		
3	SAN (Fibre Channel over Ethernet, with PFC)	3	20%
4	Business Critical LAN (lossless Ethernet, with PFC)	4	30%
5	Latency-sensitive LAN	5	40%
6	Latency-sensitive LAN		
7	Network Management (strict)	15	unlimited

The example shown in Table 23 is only slightly different than the default configuration shown in Figure 29 on page 262. In this example, latency-sensitive LAN traffic (802.1p priority 5 through 6) are moved from priority group 4 to priority group 5. This leaves Business Critical LAN traffic (802.1p priority 4) in priority group 4 by itself. Also, a new group for network management traffic has been assigned. Finally, the bandwidth allocation for priority groups 3, 4, and 5 are revised.

Note: DCBX may be configured to permit sharing or learning PFC configuration with or from external devices. This example assumes that PFC configuration is being performed manually. See "Data Center Bridging Capability Exchange" on page 269 for more information on DCBX.

This example can be configured using the following commands:

1. Turn CEE on.

```
RS G8124-E(config)# cee enable
```

Note: Turning CEE on will automatically change some 802.1p QoS and 802.3x standard flow control settings and menus (see "Turning CEE On or Off" on page 250).

2. Configure each allocated priority group with a description (optional), list of 802.1p priority values, and bandwidth allocation:

```
RS G8124-E(config)# cee global ets priority-group pgid 2 priority 0,1,2
                                        (Select a group for regular LAN, and
                                        set for 802.1p priorities 0, 1, and 2)
RS G8124-E(config)# cee global ets priority-group pgid 2 description
   "Regular LAN"                        (Set a group description—optional)

RS G8124-E(config)# cee global ets priority-group pgid 3 priority 3
                                        (Select a group for SAN traffic, and
                                        set for 802.1p priority 3)
RS G8124-E(config)# cee global ets priority-group pgid 3 description
   "SAN"
                                        (Set a group description—optional)

RS G8124-E(config)# cee global ets priority-group pgid 4 priority 4
                                        (Select a group for latency traffic,
                                        and set for 802.1p priority 4)
RS G8124-E(config)# cee global ets priority-group pgid 4 description
   "Biz-Critical LAN"                   (Set a group description—optional)

RS G8124-E(config)# cee global ets priority-group pgid 5 priority 5,6
   pgid 4 bandwidth 20 pgid 5 bandwidth 20
```

3. Enable PFC on priority 4.

```
RS G8124-E(config)# cee global pfc priority 4 enable
RS G8124-E(config)# cee global ets priority-group pgid 2 bandwidth 10
pgid 3 bandwidth 20 pgid 4 bandwidth 30 pgid 5 bandwidth 40
```

4. Save the configuration.

To view the configuration, use the following command:

```
RS G8124-E(config)# show cee global ets

Current ETS Configuration:
Number of COSq: 8

Current Mapping of 802.1p Priority to Priority Groups:

Priority  PGID  COSq
--------  ----  ----
    0       2    2
    1       2    2
    2       2    2
    3       3    3
    4       4    4
    5       5    4
    6       5    4
    7      15    7

Current Bandwidth Allocation to Priority Groups:

PGID  PG%  Description
----  ---  -----------
  2   10   Regular LAN
  3   20   SAN
  4   30   Biz-Critical LAN
  5   40   Latency-sensitive LAN
 15    -   Network Management
```

```
RS G8124-E(config)# show cee global pfc

Current Global PFC Configuration:

PFC - ON

Priority   State   Description
--------   -----   -----------
   0        Dis
   1        Dis
   2        Dis
   3        Ena
   4        Ena
   5        Dis
   6        Dis
   7        Dis
-------------------------------------------------------------------------
State - indicates whether PFC is Enabled/Disabled on a particular
priority
```

Data Center Bridging Capability Exchange

Data Center Bridging Capability Exchange (DCBX) protocol is a vital element of CEE. DCBX allows peer CEE devices to exchange information about their advanced capabilities. Using DCBX, neighboring network devices discover their peers, negotiate peer configurations, and detect misconfigurations.

DCBX provides two main functions on the G8124-E:

- Peer information exchange

  The switch uses DCBX to exchange information with connected CEE devices. For normal operation of any FCoE implementation on the G8124-E, DCBX must remain enabled on all ports participating in FCoE.

- Peer configuration negotiation

  DCBX also allows CEE devices to negotiate with each other for the purpose of automatically configuring advanced CEE features such as PFC, ETS, and (for some CNAs) FIP. The administrator can determine which CEE feature settings on the switch are communicated to and matched by CEE neighbors, and also which CEE feature settings on the switch may be configured by neighbor requirements.

The DCBX feature requires CEE to be turned on (see).

DCBX Settings

When CEE is turned on, DCBX is enabled for peer information exchange on all ports. For configuration negotiation, the following default settings are configured:

- Application Protocol: FCoE and FIP snooping is set for traffic with 802.1p priority 3

- PFC: Enabled on 802.1p priority 3

- ETS

  - Priority group 2 includes priority values 0 through 2, with bandwidth allocation of 10%

  - Priority group 3 includes priority value 3, with bandwidth allocation of 50%

  - Priority group 4 includes priority values 4 through 7, with bandwidth allocation of 40%

Enabling and Disabling DCBX

When CEE is turned on, DCBX can be enabled and disabled on a per-port basis, using the following commands:

```
RS G8124-E(config)# [no] cee port <port alias or number> dcbx enable
```

Note: DCBX and vNICs (see Chapter 14, "Virtual NICs") are not supported simultaneously on the same G8124-E.

When DCBX is enabled on a port, Link Layer Detection Protocol (LLDP) is used to exchange DCBX parameters between CEE peers. Also, the interval for LLDP transmission time is set to one second for the first five initial LLDP transmissions, after which it is returned to the administratively configured value. The minimum delay between consecutive LLDP frames is also set to one second as a DCBX default.

Peer Configuration Negotiation

CEE peer configuration negotiation can be set on a per-port basis for a number of CEE features. For each supported feature, the administrator can configure two independent flags:

- The `advertise` flag

    When this flag is set for a particular feature, the switch settings will be transmit to the remote CEE peer. If the peer is capable of the feature, and willing to accept the G8124-E settings, it will be automatically reconfigured to match the switch.

- The `willing` flag

    Set this flag when required by the remote CEE peer for a particular feature as part of DCBX signaling and support. Although some devices may also expect this flag to indicate that the switch will accept overrides on feature settings, the G8124-E retains its configured settings. As a result, the administrator must configure the feature settings on the switch to match those expected by the remote CEE peer.

These flags are available for the following CEE features:

- Application Protocol

    DCBX exchanges information regarding FCoE and FIP snooping, including the 802.1p priority value used for FCoE traffic. The `advertise` flag is set or reset using the following command:

    ```
    RS G8124-E(config)# [no] cee port <port alias or number> dcbx app_proto
        advertise
    ```

    The `willing` flag is set or reset using the following command:

    ```
    RS G8124-E(config)# [no] cee port <port alias or number> dcbx app_proto
        willing
    ```

- PFC

    DCBX exchanges information regarding whether PFC is enabled or disabled on the port. The `advertise` flag is set or reset using the following command:

    ```
    RS G8124-E(config)# [no] cee port <port alias or number> dcbx pfc advertise
    ```

    The willing flag is set or reset using the following command:

    ```
    RS G8124-E(config)# [no] cee port <port alias or number> dcbx pfc willing
    ```

- ETS

  DCBX exchanges information regarding ETS priority groups, including their 802.1p priority members and bandwidth allocation percentages. The `advertise` flag is set or reset using the following command:

  ```
  RS G8124-E(config)# [no] cee port <port alias or number> dcbx ets advertise
  ```

  The `willing` flag is set or reset using the following command:

  ```
  RS G8124-E(config)# [no] cee port <port alias or number> dcbx ets willing
  ```

Configuring DCBX

Consider an example consistent Figure 28 on page 248 and used with the previous FCoE examples in this chapter:

- FCoE is used on ports 2 and 3.
- CEE features are also used with LANs on ports 1 and 4.
- All other ports are disabled or are connected to regular (non-CEE) LAN devices.

In this example, the G8124-E acts as the central point for CEE configuration. FCoE-related ports will be configured for advertising CEE capabilities, but not to accept external configuration. Other LAN ports that use CEE features will also be configured to advertise feature settings to remote peers, but not to accept external configuration. DCBX will be disabled on all non-CEE ports.

This example can be configured using the following commands:

1. Turn CEE on.

   ```
   RS G8124-E(config)# cee enable
   ```

   Note: Turning CEE on will automatically change some 802.1p QoS and 802.3x standard flow control settings and menus (see "Turning CEE On or Off" on page 250).

2. Enable desired DCBX configuration negotiation on FCoE ports:

   ```
   RS G8124-E(config)# cee port 2 dcbx enable
   RS G8124-E(config)# cee port 2 dcbx app_proto advertise
   RS G8124-E(config)# cee port 2 dcbx ets advertise
   RS G8124-E(config)# cee port 2 dcbx pfc advertise

   RS G8124-E(config)# cee port 3 dcbx enable
   RS G8124-E(config)# cee port 3 dcbx app_proto advertise
   RS G8124-E(config)# cee port 3 dcbx ets advertise
   RS G8124-E(config)# cee port 3 dcbx pfc advertise
   ```

3. Enable desired DCBX advertisements on other CEE ports:

```
RS G8124-E(config)# cee port 1 dcbx enable
RS G8124-E(config)# cee port 1 dcbx app_proto advertise
RS G8124-E(config)# cee port 1 dcbx ets advertise
RS G8124-E(config)# cee port 1 dcbx pfc advertise

RS G8124-E(config)# cee port 4 dcbx enable
RS G8124-E(config)# cee port 4 dcbx app_proto advertise
RS G8124-E(config)# cee port 4 dcbx ets advertise
RS G8124-E(config)# cee port 4 dcbx pfc advertise
```

4. Disable DCBX for each non-CEE port as appropriate:

```
RS G8124-E(config)# no cee port 5-24 dcbx enable
```

5. Save the configuration.

Chapter 17. Static Multicast ARP

The Microsoft Windows operating system includes Network Load Balancing (NLB) technology that helps to balance incoming IP traffic among multi-node clusters. In multicast mode, NLB uses a shared multicast MAC address with a unicast IP address. Since the address resolution protocol (ARP) can map an IP address to only one MAC address, port, and VLAN, the packet reaches only one of the servers (the one attached to the port on which the ARP was learnt).

To avoid the ARP resolution, you must create a static ARP entry with multicast MAC address. You must also specify the list of ports through which the multicast packet must be sent out from the gateway or Layer 2/Layer 3 node.

With these configurations, a packet with a unicast IPv4 destination address and multicast MAC address can be sent out as per the multicast MAC address configuration. NLB maps the unicast IP address and multicast MAC address as follows:

> Cluster multicast MAC address: 03-BF-W-X-Y-Z; where W.X.Y.Z is the cluster unicast IP address.

You must configure the static multicast ARP entry only at the Layer 2/Layer 3 or Router node, and not at the Layer 2-only node.

Lenovo Networking OS supports a maximum of 20 static multicast ARP entries. When the ARP table is full, an error message appears in the syslog.

Note: If you use the ACL profile, an ACL entry is consumed for each Static Multicast ARP entry that you configure. Hence, you can configure a maximum of 896 ACLs and multicast MAC entries together when using the ACL profile.The ACL entries have a higher priority. In the default profile, the number of static multicast ARP entries that you configure does not affect the total number of ACL entries.

Configuring Static Multicast ARP

To configure multicast MAC ARP, you must perform the following steps:

- Configure the static multicast forwarding database (FDB) entry: Since there is no port list specified for static multicast ARP, and the associated MAC address is multicast, you must specify a static multicast FDB entry for the cluster MAC address to limit the multicast domain. If there is no static multicast FDB entry defined for the cluster MAC address, traffic will not be forwarded. Use the following command:

```
RS G8124-E(config)# mac-address-table multicast <cluster MAC address> <port(s)>
```

- Configure the static multicast ARP entry: Multicast ARP static entries should be configured without specifying the list of ports to be used. Use the following command:

```
RS G8124-E(config)# ip arp <destination unicast IP address> <destination multicast MAC
                      address> vlan <cluster VLAN number>
```

Configuration Example

Consider the following example:

- Cluster unicast IP address: 10.10.10.42
- Cluster multicast MAC address: 03:bf:0A:0A:0A:2A
- Cluster VLAN: 42
- List of individual or port LAGs to which traffic should be forwarded: 54 and 56

Following are the steps to configure the static multicast ARP based on the given example:

1. Configure the static multicast FDB entry.

```
RS G8124-E(config)# mac-address-table multicast 03:bf:0A:0A:0A:2A 42
54,56
```

2. Configure the static multicast ARP entry:

```
RS G8124-E(config)# ip arp 10.10.10.42 03:bf:0A:0A:0A:2A vlan 42
```

You can verify the configuration using the following commands:

- Verify static multicast FDB entry:

```
RS G8124-E(config)# show mac-address-table multicast address
03:bf:0A:0A:0A:2A

 Multicast Address  VLAN  Port(s)
 -----------------  ----  ---------
 03:bf:0A:0A:0A:2A   42  54 56
```

● Verify static multicast ARP entry:

```
RS G8124-E(config)# show ip arp
-----------------------------------------------------------------
Mgmt ARP entries:

Total number of Mgmt arp entries : 2
    IP address    Flags    MAC address      VLAN  Age Port
  --------------- -----  ----------------- ------  --- ----
   10.100.121.132  P     08:17:f4:62:64:fe  4095      MGT
   10.100.121.1          00:22:00:ad:45:00  4095   1  MGT

Data ARP entries:

Current ARP configuration:
  rearp 5

Current static ARP:
  IP address         MAC address       Port   VLAN
  --------------- ----------------- ----- ----
   10.10.10.42     03:bf:0a:0a:0a:2a          42

Total number of arp entries : 2
    IP address    Flags    MAC address     VLAN  Age Port
  --------------- -----  ----------------- ------  --- ----
   10.10.10.1      P     08:17:f4:62:64:00   42
   10.10.10.42     P     03:bf:0a:0a:0a:2a   42
```

Limitations

● You must configure the ARP only in the Layer 2/Layer 3 node or the router node but not in the Layer 2-only node. Lenovo N/OS cannot validate if the node is Layer 2-only.

● The packet is always forwarded to all the ports as specified in the Multicast MAC address configuration. If VLAN membership changes for the ports, you must update this static multicast MAC entry. If not, the ports, whose membership has changed, will report discards.

● ACLs take precedence over static multicast ARP. If an ACL is configured to match and permit ingress of unicast traffic, the traffic will be forwarded based on the ACL rule, and the static multicast ARP will be ignored.

Chapter 18. Dynamic ARP Inspection

Address Resolution Protocol (ARP) provides IP communication within a Layer 2 broadcast domain by mapping an IP address to a MAC address. Network devices maintain this mapping in a cache that they consult when forwarding packets to other devices. If the ARP cache does not contain an entry for the destination device, the host broadcasts an ARP request for that device's address and stores the response in the cache.

Understanding ARP Spoofing Attacks

ARP spoofing (also referred to as ARP cache poisoning) is one way to initiate man-in-the-middle attacks. A malicious user could poison the ARP caches of connected systems (hosts, switches, routers) by sending forged ARP responses and could intercept traffic intended for other hosts on the LAN segment.

For example, in Figure 30, the attacker (Host C) can send an ARP Reply to Host A pretending to be Host B. As a result, Host A populates its ARP cache with a poisoned entry having IP address IB and MAC address MC. Host A will use the MAC address MC as the destination MAC address for traffic intended for Host B. Host C then intercepts that traffic. Because Host C knows the true MAC addresses associated with Host B, it forwards the intercepted traffic to that host by using the correct MAC address as the destination, keeping the appearance of regular behavior.

Figure 30. ARP Cache Poisoning
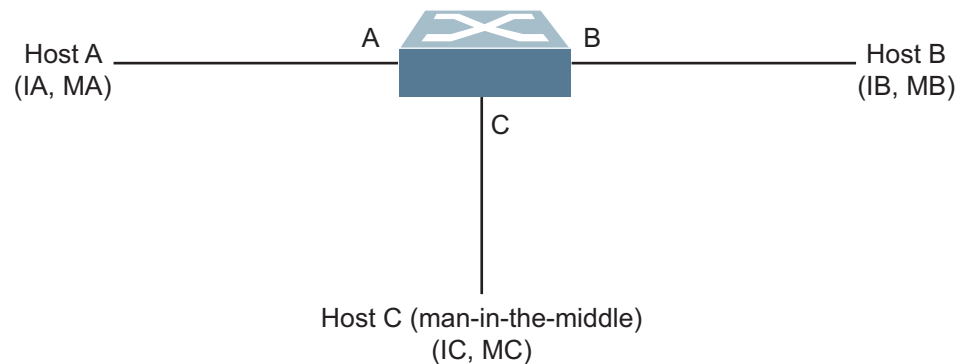
Understanding DAI

Dynamic ARP Inspection (DAI) is an addition to the feature DHCP Snooping.

Dynamic ARP Inspection is a security feature that lets the switch intercept and examine all ARP request and response packets in a subnet, discarding those packets with invalid IP to MAC address bindings. This capability protects the network from man-in-the-middle attacks.

A switch on which ARP Inspection is configured does the following:

● Intercepts all ARP requests and responses on untrusted ports.

- Verifies that each of these intercepted packets has a valid IP/MAC/VLAN/port binding before updating the local ARP cache or before forwarding the packet to the appropriate destination.
- Drops invalid ARP packets and sends a syslog message with details about each dropped packet.

DAI determines the validity of an ARP packet based on valid IP-to-MAC address bindings stored in a trusted database, the DHCP snooping binding database. This database is built by DHCP snooping if DHCP snooping is enabled on the VLANs and on the switch. As shown in Figure 31, if the ARP packet is received on a trusted interface, the switch forwards the packet without any checks. On untrusted interfaces, the switch forwards the packet only if it is valid.

For hosts with statically configured IP addresses, static DHCP snooping binding entries can be configured with a big lease time.

Figure 31. Dynamic ARP inspection at work
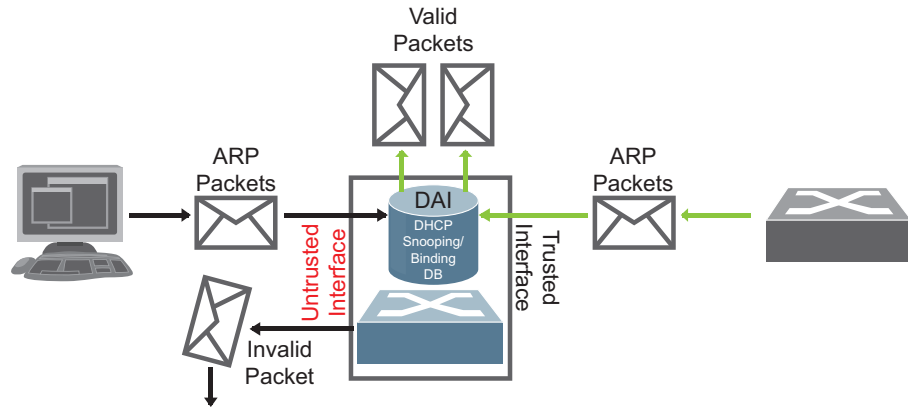
Interface Trust States and Network Security

DAI associates a trust state with each interface on the switch.

In a typical network configuration, you configure all switch ports connected to host ports as untrusted and configure all switch ports connected to switches as trusted. With this configuration, all ARP packets entering the network from a given switch bypass the security check.

The trust state configuration should be done carefully: configuring interfaces as untrusted when they should be trusted can result in a loss of connectivity.

In Figure 32, assume that both Switch A and Switch B are running DAI on the VLAN that includes Host 1 and Host 2. If Host 1 and Host 2 acquire their IP addresses from the DHCP server connected to Switch A, only Switch A has the DHCP IP-to-MAC binding of Host 1. Therefore, if the interface between Switch A and Switch B is untrusted, the ARP packets from Host 1 are dropped by Switch B. Connectivity between Host 1 and Host 2 is lost.

Figure 32. ARP Packet Validation on a VLAN Enabled for DAI

DHCP server

Port 1 | **Switch A**   **Switch B**

Port 2   Port 2

Port 3   Port 3

Host 1   Host 2
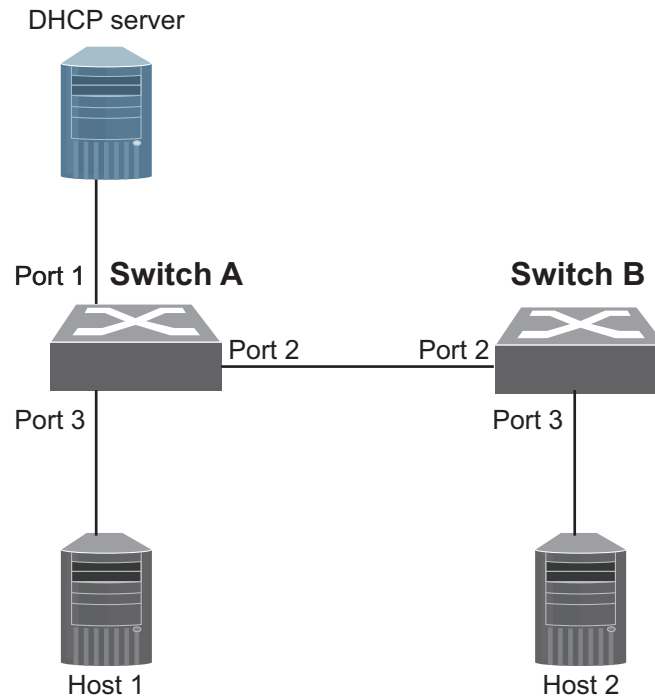
If Switch A is not running DAI, Host 1 can easily poison the ARP caches of Switch B and Host 2, if the link between the switches is configured as trusted. This condition can occur even though Switch B is running DAI.

The best option for the setup from Figure 32 is to have DAI running on both switches and to have the link between the switches configured as trusted.

In cases in which some switches in a VLAN run DAI and other switches do not, configure the interfaces connecting such switches as untrusted. However, to validate the bindings of packets from switches where DAI is not configured, configure static DHCP snooping binding entries on the switch running DAI. When you cannot determine such bindings, isolate switches running DAI at Layer 3 from switches not running DAI.

DAI ensures that hosts (on untrusted interfaces) connected to a switch running DAI do not poison the ARP caches of other hosts in the network. However, DAI does not prevent hosts in other portions of the network connected through a trusted interface from poisoning the caches of the hosts that are connected to a switch running DAI.

DAI Configuration Guidelines and Restrictions

When configuring DAI, follow these guidelines and restrictions:

● DAI is an ingress security feature; it does not perform any egress checking.

● DAI is not effective for hosts connected to switches that do not support DAI or that do not have this feature enabled. Because man-in-the-middle attacks are limited to a single Layer 2 broadcast domain, separate the domain with DAI checks from the one with no checking. This action secures the ARP caches of hosts in the domain enabled for DAI.

● DAI depends on the entries in the DHCP snooping binding database to verify IP-to-MAC address bindings in incoming ARP requests and ARP responses.

● For non-DHCP environments, for each static IP address add a static DHCP Snooping binding entry with the biggest lease time in order not to expire.

● Ports belonging to a port-channel must have the same trust state.

DAI Configuration Example

Following is the configuration for the example in .

```
SwitchA(config)# interface port 1-3
SwitchA(config-if)# switchport access vlan 2
SwitchA(config)# interface port 1-2
SwitchA(config-if)# ip arp inspection trust
SwitchA(config-if)# exit
SwitchA(config)# interface port 3
SwitchA(config-if)# no ip arp inspection trust
SwitchA(config-if)# exit
SwitchA(config)# ip arp inspection vlan 2

SwitchB(config)# interface port 2-3
SwitchB(config-if)# switchport access vlan 2
SwitchB(config)# interface port 2
SwitchB(config-if)# ip arp inspection trust
SwitchB(config-if)# exit
SwitchB(config)# interface port 3
SwitchB(config-if)# no ip arp inspection trust
SwitchB(config-if)# exit
SwitchB(config)# ip arp inspection vlan 2
```

The DHCP Snooping binding tables will be similar to the following:

```
Mac Address        IP Address      Lease(seconds) Type      VLAN
Interface
---------------------------------------------------------------------
-----
00:00:00:00:00:01  Host1_IP        1000           Dynamic   2    3

00:00:00:00:00:02  Host2_IP        2000           Dynamic   2    2
Total number of bindings: 2

SwitchB#show ip dhcp snooping binding
Mac Address        IP Address      Lease(seconds) Type      VLAN
Interface
---------------------------------------------------------------------
-----
00:00:00:00:00:02  Host2_IP        2000           Dynamic   2    3
Total number of bindings: 1
```

```
SwitchA# show ip dhcp snooping binding

                                  Output of show commands:

SwitchA# show ip arp inspection vlan

Vlan                              Configuration
----                              -------------
  2                               Enabled

SwitchA# show ip arp inspection interfaces

Interface                         Trust State
---------                         -----------
                                     1Trusted
                                     2Trusted
                                     3Untrusted
                                     4Untrusted
...

SwitchA# show ip arp inspection statistics

Vlan                              ForwardedDropped
----                              ---------------
  2                               100   200
```

When Host 1 tries to send an ARP with an IP address of 1.1.1.3 that is not present in the DHCP Binding table, the packet is dropped and an error message similar to the following is logged:

"Dec 16 21:00:10 192.168.49.50 NOTICE ARP-Inspection: Invalid ARP Request on port 3, VLAN 2 ([00:02:00:02:00:02/1.1.1.3/00:00:00:00:00:00/1.1.1.4])"

Part 5: IP Routing

This section discusses Layer 3 switching functions. In addition to switching traffic at near line rates, the application switch can perform multi-protocol routing. This section discusses basic routing and advanced routing protocols:

- Basic IP Routing
- Routed Ports
- Internet Protocol Version 6
- IPsec with IPv6
- Routing Information Protocol
- Internet Group Management Protocol
- Multicast Listener Discovery
- Border Gateway Protocol
- Open Shortest Path First
- Protocol Independent Multicast

Chapter 19. Basic IP Routing

This chapter provides configuration background and examples for using the G8124-E to perform IP routing functions. The following topics are addressed in this chapter:

- "IP Routing Benefits" on page 286
- "Routing Between IP Subnets" on page 286
- "Example of Subnet Routing" on page 287
- "ECMP Static Routes" on page 291
- "Dynamic Host Configuration Protocol" on page 293

IP Routing Benefits

The switch uses a combination of configurable IP switch interfaces and IP routing options. The switch IP routing capabilities provide the following benefits:

- Connects the server IP subnets to the rest of the backbone network.

- Provides the ability to route IP traffic between multiple Virtual Local Area Networks (VLANs) configured on the switch.

Routing Between IP Subnets

The physical layout of most corporate networks has evolved over time. Classic hub/router topologies have given way to faster switched topologies, particularly now that switches are increasingly intelligent. The G8124-E is intelligent and fast enough to perform routing functions at wire speed.

The combination of faster routing and switching in a single device allows you to build versatile topologies that account for legacy configurations.

For example, consider a corporate campus that has migrated from a router-centric topology to a faster, more powerful, switch-based topology. As is often the case, the legacy of network growth and redesign has left the system with a mix of illogically distributed subnets.

This is a situation that switching alone cannot cure. Instead, the router is flooded with cross-subnet communication. This compromises efficiency in two ways:

- Routers can be slower than switches. The cross-subnet side trip from the switch to the router and back again adds two hops for the data, slowing throughput considerably.

- Traffic to the router increases, increasing congestion.

Even if every end-station could be moved to better logical subnets (a daunting task), competition for access to common server pools on different subnets still burdens the routers.

This problem is solved by using switches with built-in IP routing capabilities. Cross-subnet LAN traffic can now be routed within the switches with wire speed switching performance. This eases the load on the router and saves the network administrators from reconfiguring every end-station with new IP addresses.

Example of Subnet Routing

Consider the role of the G8124-E in the following configuration example:

Figure 33. Switch-Based Routing Topology
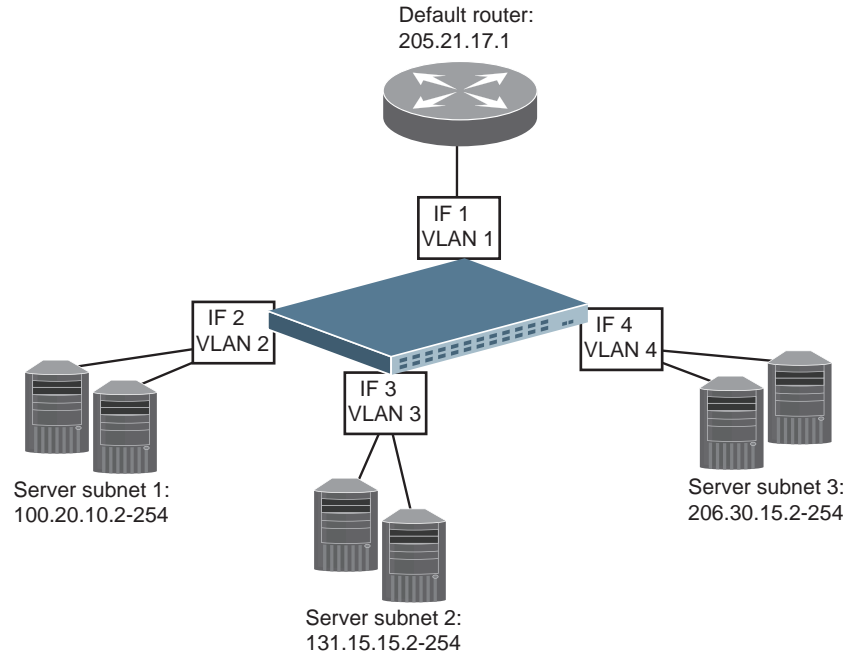
The switch connects the Gigabit Ethernet and Fast Ethernet LAGs from various switched subnets throughout one building. Common servers are placed on another subnet attached to the switch. A primary and backup router are attached to the switch on yet another subnet.

Without Layer 3 IP routing on the switch, cross-subnet communication is relayed to the default gateway (in this case, the router) for the next level of routing intelligence. The router fills in the necessary address information and sends the data back to the switch, which then relays the packet to the proper destination subnet using Layer 2 switching.

With Layer 3 IP routing in place on the switch, routing between different IP subnets can be accomplished entirely within the switch. This leaves the routers free to handle inbound and outbound traffic for this group of subnets.

Using VLANs to Segregate Broadcast Domains

If you want to control the broadcasts on your network, use VLANs to create distinct broadcast domains. Create one VLAN for each server subnet, and one for the router.

Configuration Example

This section describes the steps used to configure the example topology shown in .

1. Assign an IP address (or document the existing one) for each router and each server.

   The following IP addresses are used:

   Table 24. *Subnet Routing Example: IP Address Assignments*

Subnet	Devices	IP Addresses
1	Default router	205.21.17.1
2	Web servers	100.20.10.2-254
3	Database servers	131.15.15.2-254
4	Terminal Servers	206.30.15.2-254

2. Assign an IP interface for each subnet attached to the switch.

   Since there are four IP subnets connected to the switch, four IP interfaces are needed:

   Table 25. *Subnet Routing Example: IP Interface Assignments*

Interface	Devices	IP Interface Address
IF 1	Default router	205.21.17.3
IF 2	Web servers	100.20.10.1
IF 3	Database servers	131.15.15.1
IF 4	Terminal Servers	206.30.15.1

3. Determine which switch ports and IP interfaces belong to which VLANs.

   The following table adds port and VLAN information:

   Table 26. *Subnet Routing Example: Optional VLAN Ports*

Devices	IP Interface	Switch Ports	VLAN #
Default router	1	22	1
Web servers	2	1 and 2	2
Database servers	3	3 and 4	3
Terminal Servers	4	5 and 6	4

Note: To perform this configuration, you must be connected to the switch Industry Standard Command Line Interface (ISCLI) as the administrator.

4. Add the switch ports to their respective VLANs.

The VLANs shown in Table 26 are configured as follows:

```
RS G8124-E(config)# vlan 1
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 22  (Add ports to VLAN 1)
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 1
RS G8124-E(config-if)# exit

RS G8124-E(config)# vlan 2
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 1,2 (Add ports to VLAN 2)
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 2
RS G8124-E(config-if)# exit

RS G8124-E(config)# vlan 3
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 3,4 (Add ports to VLAN 3)
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 3
RS G8124-E(config-if)# exit


RS G8124-E(config)# vlan 4
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 5,6 (Add ports to VLAN 4)
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 4
RS G8124-E(config-if)# exit
```

Each time you add a port to a VLAN, you may get the following prompt:

```
Port 4 is an untagged port and its PVID is changed from 1 to 3.
```

5. Assign a VLAN to each IP interface.

   Now that the ports are separated into VLANs, the VLANs are assigned to the appropriate IP interface for each subnet. From Table 26 on page 288, the settings are made as follows:

```
RS G8124-E(config)# interface ip 1              (Select IP interface 1)
RS G8124-E(config-ip-if)# ip address 205.21.17.3
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# vlan 1              (Add VLAN 1)
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface ip 2              (Select IP interface 2)
RS G8124-E(config-ip-if)# ip address 100.20.10.1
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# vlan 2              (Add VLAN 2)
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 3              (Select IP interface 3)
RS G8124-E(config-ip-if)# ip address 131.15.15.1
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# vlan 3              (Add VLAN 3)
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 4              (Select IP interface 4)
RS G8124-E(config-ip-if)# ip address 206.30.15.1
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# vlan 4              (Add VLAN 4)
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

6. Configure the default gateway to the routers' addresses.

   The default gateway allows the switch to send outbound traffic to the router:

```
RS G8124-E(config)# ip gateway 1 address 205.21.17.1
RS G8124-E(config)# ip gateway 1 enable
```

7. Enable IP routing.

```
RS G8124-E(config)# ip routing
```

8. Verify the configuration.

```
RS G8124-E(config)# show vlan
RS G8124-E(config)# show interface information
RS G8124-E(config)# show interface ip
```

Examine the resulting information. If any settings are incorrect, make the appropriate changes.

ECMP Static Routes

Equal-Cost Multi-Path (ECMP) is a forwarding mechanism that routes packets along multiple paths of equal cost. ECMP provides equally-distributed link load sharing across the paths. The hashing algorithm used is based on the destination IP and source IP (DIPSIP) addresses or only on the source IP address (SIP). ECMP routes allow the switch to choose between several next hops toward a given destination. The switch performs periodic health checks (ping) on each ECMP gateway. If a gateway fails, it is removed from the routing table, and an SNMP trap is sent.

ECMP Route Hashing

You can configure the parameters used to perform ECMP route hashing, as follows:

- `sip`: Source IP address
- `dip`: Destination IP address
- `sport`: Source port
- `dport`: Destination port

Note: Source port (`sport`) and/or destination port (`dport`) options for the ECMP route hash (`ip route ecmphash`) are supported only when Layer 4 `tcpl4` and/or `udpl4` options are enabled. Conversely, when `tcp14` and/or `udpl4` are enabled, hashing options for `sport` and/or `dport` must also be enabled.

- `protocol`: Layer 3 protocol
- `tcpl4`: Layer 4 TCP port
- `udpl4`: Layer 4 UDP port

Note: The default ECMP has mechanism is based on sip, dip, tcpl4, udpl4, and dport.

The ECMP hash setting applies to all ECMP routes.

Configuring ECMP Static Routes

To configure ECMP static routes, add the same route multiple times, each with the same destination IP address, but with a different gateway IP address. These routes become ECMP routes.

1. Add a static route (IP address, subnet mask, gateway, and interface number).

```
RS G8124-E(config)# ip route 10.10.1.1 255.255.255.255 100.10.1.1 1
```

2. Add another static route with the same IP address and mask, but a different gateway address.

```
RS G8124-E(config)# ip route 10.10.1.1 255.255.255.255 200.20.2.2 1
```

3. Select an ECMP hashing method (optional).

```
RS G8124-E(config)# ip route ecmphash [sip|dip|protocol|tcpl4|udpl4|
sport|dport]
```

You may add up to 16 gateways for each static route.

Use the following commands to check the status of ECMP static routes:

```
RS G8124-E(config)# show ip route static
Current static routes:
  Destination     Mask            Gateway         If   ECMP
  --------------- --------------- --------------- ---- ------
  10.20.2.2       255.255.255.255 10.4.4.1              *
                                  10.5.5.1             *
                                  10.6.6.1             *
                                   ...
                                  10.35.35.1           *
ECMP health-check ping interval: 1
ECMP health-check retries number: 3
ECMP Hash Mechanism: dipsip
Gateway healthcheck: enabled
```

```
RS G8124-E(config)# show ip ecmp
Current ecmp static routes:
  Destination     Mask            Gateway         If   GW Status
  --------------- --------------- --------------- ---- -----------
  10.20.2.2       255.255.255.255 10.4.4.1              up
                                  10.5.5.1             up
                                  10.6.6.1             up
                                   ...
                                  10.34.34.1           up
                                  10.35.35.1           up
```

Dynamic Host Configuration Protocol

Dynamic Host Configuration Protocol (DHCP) is a transport protocol that provides a framework for automatically assigning IP addresses and configuration information to other IP hosts or clients in a large TCP/IP network. Without DHCP, the IP address must be entered manually for each network device. DHCP allows a network administrator to distribute IP addresses from a central point and automatically send a new IP address when a device is connected to a different place in the network.

The switch accepts gateway configuration parameters if they have not been configured manually. The switch ignores DHCP gateway parameters if the gateway is configured.

DHCP is an extension of another network IP management protocol, Bootstrap Protocol (BOOTP), with an additional capability of being able to allocate reusable network addresses and configuration parameters for client operation.

Built on the client/server model, DHCP allows hosts or clients on an IP network to obtain their configurations from a DHCP server, thereby reducing network administration. The most significant configuration the client receives from the server is its required IP address; (other optional parameters include the "generic" file name to be booted, the address of the default gateway, and so forth).

To enable DHCP on a switch management interface, use the following command:

```
RS G8124-E(config)# system dhcp mgta
```

To configure DHCP operation on data interfaces, use the following command:

```
RS G8124-E(config)# system bootp
```

DHCP is described in RFC 2131, and the DHCP relay agent supported on the G8124-E is described in RFC 1542. DHCP uses UDP as its transport protocol. The client sends messages to the server on port 67 and the server sends messages to the client on port 68.

DHCP defines the methods through which clients can be assigned an IP address for a finite lease period and allowing reassignment of the IP address to another client later. Additionally, DHCP provides the mechanism for a client to gather other IP configuration parameters it needs to operate in the TCP/IP network.

In the DHCP environment, the G8124-E acts as a relay agent. The DHCP relay feature enables the switch to forward a client request for an IP address to two BOOTP servers with IP addresses that have been configured on the switch.

When a switch receives a UDP broadcast on port 67 from a DHCP client requesting an IP address, the switch acts as a proxy for the client, replacing the client source IP (SIP) and destination IP (DIP) addresses. The request is then forwarded as a UDP Unicast MAC layer message to two BOOTP servers whose IP addresses are configured on the switch. The servers respond as a UDP Unicast message back to the switch, with the default gateway and IP address for the client. The destination IP address in the server response represents the interface address on the switch that received the client request. This interface address tells the switch on which VLAN to send the server response to the client.

To enable the G8124-E to be the BOOTP forwarder, you need to configure the DHCP/BOOTP server IP addresses on the switch. Generally, it is best to configure the switch IP interface on the client side to match the client's subnet, and configure VLANs to separate client and server subnets. The DHCP server knows from which IP subnet the newly allocated IP address will come.

In G8124-E implementation, there is no need for primary or secondary servers. The client request is forwarded to the BOOTP servers configured on the switch. The use of two servers provide failover redundancy. However, no health checking is supported.

Use the following commands to configure the switch as a DHCP relay agent:

```
RS G8124-E(config)# ip bootp-relay server 1 <IP address>
RS G8124-E(config)# ip bootp-relay server 2 <IP address>
RS G8124-E(config)# ip bootp-relay enable
RS G8124-E(config)# show ip bootp-relay
```

Additionally, DHCP Relay functionality can be assigned on a per interface basis. Use the following commands to enable the Relay functionality:

```
RS G8124-E(config)# interface ip <Interface number>
RS G8124-E(config-ip-if)# relay
```

Chapter 20. Routed Ports

By default, all ports on the RackSwitch G8124-E behave as switch ports, which are capable of performing Layer 2 switch functions, such as VLANs, STP, or bridging. Switch ports also provide a physical point of access for the switch IP interfaces, which can perform global Layer 3 functions, such as routing for BGP or OSPF.

However, G8124-E ports can also be configured as routed ports. Routed ports are configured with their own IP address belonging to a unique Layer 3 network, and behave similar to a port on a conventional router. Routed ports are typically used for connecting to a server or to a router.

When a switch port is configured as a routed port, it forwards Layer 3 traffic and no longer performs Layer 2 switching functions.

Overview

A routed port has the following characteristics:

- Does not participate in bridging.
- Does not belong to any user-configurable VLAN.
- Does not implement any Layer 2 functionality, such as Spanning Tree Protocol (STP).
- Is always in a forwarding state.
- Can participate in IPv4 routing.
- Can be configured with basic IP protocols, such as Internet Control Message Protocol (ICMP), and with Layer 3 protocols, such as Protocol-Independent Multicast (PIM), Routing Information Protocol (RIP), Open Shortest Path First (OSPF), and Border Gateway Protocol (BGP).
- Can be configured with Internet Group Management Protocol (IGMP) querier and snooping functionality.
- Layer 3 configuration is saved even when the interface is shutdown.
- MAC address learning is always enabled.
- Tagging and port VLAN ID (PVID) tagging is disabled.
- Flooding is disabled.
- Bridge Protocol Data Unit (BPDU)-guard is disabled.
- Link Aggregation Control Protocol (LACP) is disabled.
- Multicast threshold is disabled.
- Static Multicast MAC and static unicast MAC can be configured.

Notes:

- Ports on which LACP or portchannel is enabled cannot be changed to routed ports.
- Ports that have Static MAC addresses configured cannot be changed to routed ports.

When a switch port is configured as a routed port, the following configuration changes are automatically implemented:

- The port is removed from all the VLANs it belonged to.
- The port is added to an internal VLAN on which flooding is disabled. The ID of this internal VLAN could be 4094 or lower. The internal VLAN is assigned to Spanning Tree Group (STG) 1, if RSTP/PVRST is configured; or to Common Internal Spanning Tree (CIST), if MSTP is configured. You cannot change the VLAN number assigned to the routed port.
- STP is disabled and the port is set to a forwarding state.

  Note: The maximum number of VLANs you can configure on the RackSwitch G8124-E is 4095. This maximum number will be reduced by the number of routed ports you configure.

- All the Layer 2 configuration is lost.

When a routed port is changed back to a switch port, the following changes take place:

- All the IP configuration is lost.
- The ARP entry corresponding to the IP address is lost.
- The switch port is added to the default VLAN and STG. In case of MSTP, it is added to the CIST.
- STP is turned on.
- The switch port can participate in STG and VLAN flooding.
- Can participate in bridging.
- LACP port attributes are set to default.
- Multicast threshold remains disabled.
- BPDU guard remains disabled.
- IGMP configuration is lost.

Note: When you configure a routed port to back to a switch port, it does not restore the Layer 2 configuration it had before it was changed to a routed port.

Configuring a Routed Port

Note: Use only the ISCLI to configure routed ports. Configurations using the BBI are not supported. Configurations made using SNMP cannot be saved or applied.

Note: You cannot configure a management port to be a routed port.

Following are the basic steps for configuring a routed port:

1. Enter the interface configuration mode for the port.

```
RS G8124-E(config)# interface port <port number>
```

Note: You must enter only one port number. If you need to change multiple ports to routed ports, repeat the configuration steps for each port.

2. Enable routing.

```
RS G8124-E(config-if)# no switchport
```

3. Assign an IP address.

```
RS G8124-E(config-if)# ip address <IP address> <Subnet Mask> enable
```

4. (Optional) Enable a Layer 3 routing protocol.

```
RS G8124-E(config-if)# ip {<ospf>|<pim>|<rip>}
```

Note: Configure the Layer 3 routing protocol-related parameters in the interface configuration mode.

Configuring OSPF on Routed Ports

The following OSPF configuration commands are supported on routed ports:

```
RS G8124-E(config-if)# ip ospf ?

  area                 Set area index
  cost                 Set interface cost
  dead-interval        Set dead interval in seconds or milliseconds
  enable               Enable OSPF for this interface
  hello-interval       Set hello interval in seconds or milliseconds
  key                  Set authentication key
  message-digest-key   Set MD5 key ID
  passive-interface    Enable passive interface
  point-to-point       Enable point-to-point interface
  priority             Set interface router priority
  retransmit-interval  Set retransmit interval in seconds
  transit-delay        Set transit delay in seconds
```

See Chapter 27, "Open Shortest Path First," for details on the OSPF protocol and its configuration.

OSPFv3 cannot be configured on routed ports.

OSPF Configuration Example

The following example includes the basic steps for configuring OSPF on a routed port:

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area 0 enable
RS G8124-E(config-router-ospf)# enable
RS G8124-E(config-router-ospf)# exit
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# no switchport
wait...
RS G8124-E(config-if)# ip address 11.1.12.1 255.255.255.0 enable
wait...
RS G8124-E(config-if)# ip ospf area 0
RS G8124-E(config-if)# ip ospf enable
RS G8124-E(config-if)# exit
```

Configuring RIP on Routed Ports

The following RIP configuration commands are supported on routed ports:

```
RS G8124-E(config-if)# ip rip ?

  authentication     Set IP authentication
  default-action     Set default route action
  enable             Enable RIP interface
  listen             Enable listening to route updates
  metric             Set metric
  multicast-updates  Enable multicast updates
  poison             Enable poisoned reverse
  split-horizon      Enable split horizon
  supply             Enable supplying route updates
  triggered          Enable triggered updates
  version            RIP version
```

See Chapter 23, "Routing Information Protocol," for details on the RIP protocol and its configuration.

RIP Configuration Example

The following example includes steps for a basic RIP configuration on a routed port:

```
RS G8124-E(config)# router rip
RS G8124-E(config-router-rip)# enable
RS G8124-E(config-router-rip)# exit
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# no switchport
wait...
RS G8124-E(config-if)# ip address 11.1.12.1 255.255.255.0 enable
wait...
RS G8124-E(config-if)# ip rip enable
RS G8124-E(config-if)# exit
```

Configuring PIM on Routed Ports

The following PIM configuration commands are supported on routed ports:

```
RS G8124-E(config-if)# ip pim ?

  border-bit          Set interface as border interface
  cbsr-preference     Set preference for local interface as a candidate
                      bootstrap router
  component-id        Add interface to the component
  dr-priority         Set designated router priority for the router
                      interface
  enable              Enable PIM on this interface
  hello-holdtime      Set hello message holdtime for the interface
  hello-interval      Set the frequency of PIM hello messages on the
                      interface
  join-prune-interval Set frequency of PIM Join or Prune interval
  lan-delay           Set lan delay for the router interface
  lan-prune-delay     Enable lan delay advertisement on interface
  neighbor-addr       Neighbor address
  neighbor-filter     Enable neighbor filter
  override-interval   Set override interval for router interface
```

See for details on the PIM protocol and its configuration.

PIM Configuration Example

The following example includes the basic steps for configuring PIM on a routed port:

```
RS G8124-E(config)# ip pim enable
RS G8124-E(config)# interface port 26
RS G8124-E(config-if)# no switchport
wait...
RS G8124-E(config-if)# ip address 26.26.26.1 255.255.255.0 enable
wait...
RS G8124-E(config-if)# ip pim enable
RS G8124-E(config-if)# exit

RS G8124-E(config)# ip pim component 1
RS G8124-E(config-ip-pim-component)# rp-candidate rp-address 224.0.0.0
240.0.0.0 26.26.26.1
RS G8124-E(config-ip-pim-component)# rp-candidate holdtime 200
RS G8124-E(config-ip-pim-component)# exit

RS G8124-E(config)# interface port 26
RS G8124-E(config-if)# ip pim cbsr-preference 200
RS G8124-E(config-if)# exit
```

Verify the configuration using the following command:

```
RS G8124-E(config)# show ip pim interface port 26

Address     IfName/IfId Ver/Mode  Nbr   Qry         DR-Address   DR-Prio
                        Count Interval
-------     ----------- --------  ----- --------    ----------   -------
26.26.26.1    Rport  26   2/Sparse  0     30         26.26.26.1      1
```

Configuring BGP on Routed Ports

The routed port can be used to establish a TCP connection to form peer relationship with another BGP router. See Chapter 26, "Border Gateway Protocol," for details on the BGP protocol and its configuration.

The following BGP configurations are not supported on routed ports:

● Update source - configuring a local IP interface

Configuring IGMP on Routed Ports

IGMP querier and snooping can be configured on routed ports. For details, see Chapter 24, "Internet Group Management Protocol."

To configure IGMP snooping on a routed port, enter the following command in the Global Configuration mode:

```
RS G8124-E(config)# ip igmp snoop port <routed port ID>
```

To configure fastleave on routed ports, enter the following command in the Global Configuration mode:

```
RS G8124-E(config)# ip igmp snoop port <routed port ID> fast-leave
```

The following IGMP querier commands are supported on routed ports:

```
RS G8124-E(config)# ip igmp querier port <routed port ID> ?

  election-type     Set IGMP querier type
  enable            Turn IGMP Querier on
  max-response      Set Queriers max response time
  query-interval    Set general query interval for IGMP Querier only
  robustness        Set IGMP robustness
  source-ip         Set source IP to be used for IGMP
  startup-count     Set startupcount for IGMP
  startup-interval  Set startup query interval for IGMP
  version           Sets the operating version of the IGMP snooping
                     switch
```

Limitations

Following features/configurations are not supported on routed ports:

- Subinterfaces
- BPDU Guard
- Broadcast Threshold
- Multicast Threshold
- Link Aggregation Control Protocol (LACP)
- Static Aggregation
- Fibre Channel over Ethernet (FCoE)
- Converged Enhanced Ethernet (CEE)
- IPv6
- IP Security (IPsec)
- Internet Key Exchange version 2 (IKEv2)
- Virtual Router Redundancy Protocol (VRRP)
- Policy-based Routing (PBR)
- Hotlinks
- Failover
- 802.1X
- Dynamic Host Configuration Protocol (DHCP)
- BOOTP
- Simple Network Management Protocol (SNMP)
- IGMP Relay
- Static Mrouter Port
- Management Port

Chapter 21. Internet Protocol Version 6

Internet Protocol version 6 (IPv6) is a network layer protocol intended to expand the network address space. IPv6 is a robust and expandable protocol that meets the need for increased physical address space. The switch supports the following RFCs for IPv6-related features:

● RFC 1981	● RFC 3306	● RFC 4007	● RFC 4443
● RFC 2404	● RFC 3307	● RFC 4213	● RFC 4552
● RFC 2410	● RFC 3411	● RFC 4291	● RFC 4718
● RFC 2451	● RFC 3412	● RFC 4292	● RFC 4835
● RFC 2460	● RFC 3413	● RFC 4293	● RFC 4861
● RFC 2474	● RFC 3414	● RFC 4301	● RFC 4862
● RFC 2526	● RFC 3484	● RFC 4302	● RFC 5095
● RFC 2711	● RFC 3602	● RFC 4303	● RFC 5114
● RFC 2740	● RFC 3810	● RFC 4306	● RFC 5340
● RFC 3289	● RFC 3879	● RFC 4307	

This chapter describes the basic configuration of IPv6 addresses and how to manage the switch via IPv6 host management.

IPv6 Limitations

The following IPv6 features are not supported in this release:

● Dynamic Host Control Protocol for IPv6 (DHCPv6)
● Border Gateway Protocol for IPv6 (BGP)
● Routing Information Protocol for IPv6 (RIPng)

Most other Lenovo Networking OS 8.3 features permit IP addresses to be configured using either IPv4 or IPv6 address formats. However, the following switch features support IPv4 only:

● Bootstrap Protocol (BOOTP) and DHCP
● RADIUS, TACACS+ and LDAP
● VMware Virtual Center (vCenter) for VMready
● Routing Information Protocol (RIP)
● Border Gateway Protocol (BGP)
● Protocol Independent Multicast (PIM)
● Virtual Router Redundancy Protocol (VRRP)
● sFlow

IPv6 Address Format

The IPv6 address is 128 bits (16 bytes) long and is represented as a sequence of eight 16-bit hex values, separated by colons.

Each IPv6 address has two parts:

- Subnet prefix representing the network to which the interface is connected

- Local identifier, either derived from the MAC address or user-configured

The preferred hexadecimal format is as follows:

```
xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx
```

Example IPv6 address:

```
FEDC:BA98:7654:BA98:FEDC:1234:ABCD:5412
```

Some addresses can contain long sequences of zeros. A single contiguous sequence of zeros can be compressed to :: (two colons). For example, consider the following IPv6 address:

```
FE80:0:0:0:2AA:FF:FA:4CA2
```

The address can be compressed as follows:

```
FE80::2AA:FF:FA:4CA2
```

Unlike IPv4, a subnet mask is not used for IPv6 addresses. IPv6 uses the subnet prefix as the network identifier. The prefix is the part of the address that indicates the bits that have fixed values or are the bits of the subnet prefix. An IPv6 prefix is written in address/prefix-length notation. For example, in the following address, 64 is the network prefix:

```
21DA:D300:0000:2F3C::/64
```

IPv6 addresses can be either user-configured or automatically configured. Automatically configured addresses always have a 64-bit subnet prefix and a 64-bit interface identifier. In most implementations, the interface identifier is derived from the switch's MAC address, using a method called EUI-64.

Most Lenovo N/OS 8.3 features permit IP addresses to be configured using either IPv4 or IPv6 address formats. Throughout this manual, *IP address* is used in places where either an IPv4 or IPv6 address is allowed. In places where only one type of address is allowed, the type (IPv4 or IPv6) is specified.

IPv6 Address Types

IPv6 supports three types of addresses: unicast (one-to-one), multicast (one-to-many), and anycast (one-to-nearest). Multicast addresses replace the use of broadcast addresses.

Unicast Address

Unicast is a communication between a single host and a single receiver. Packets sent to a unicast address are delivered to the interface identified by that address. IPv6 defines the following types of unicast address:

● Global Unicast address: An address that can be reached and identified globally. Global Unicast addresses use the high-order bit range up to FF00, therefore all non-multicast and non-link-local addresses are considered to be global unicast. A manually configured IPv6 address must be fully specified. Autoconfigured IPv6 addresses are comprised of a prefix combined with the 64-bit EUI. RFC 4291 defines the IPv6 addressing architecture.

The interface ID must be unique within the same subnet.

● Link-local unicast address: An address used to communicate with a neighbor on the same link. Link-local addresses use the format `FE80::EUI`

Link-local addresses are designed to be used for addressing on a single link for purposes such as automatic address configuration, neighbor discovery, or when no routers are present.

Routers must not forward any packets with link-local source or destination addresses to other links.

Multicast

Multicast is communication between a single host and multiple receivers. Packets are sent to all interfaces identified by that address. An interface may belong to any number of multicast groups.

A multicast address (FF00 - FFFF) is an identifier for a group interface. The multicast address most often encountered is a solicited-node multicast address using prefix `FF02::1:FF00:0000/104` with the low-order 24 bits of the unicast or anycast address.

The following well-known multicast addresses are pre-defined. The group IDs defined in this section are defined for explicit scope values, as follows:

`FF00:::::::0` through `FF0F:::::::0`

Anycast

Packets sent to an anycast address or list of addresses are delivered to the nearest interface identified by that address. Anycast is a communication between a single sender and a list of addresses.

Anycast addresses are allocated from the unicast address space, using any of the defined unicast address formats. Thus, anycast addresses are syntactically indistinguishable from unicast addresses. When a unicast address is assigned to

more than one interface, thus turning it into an anycast address, the nodes to which the address is assigned must be explicitly configured to know that it is an anycast address.

IPv6 Address Autoconfiguration

IPv6 supports the following types of address autoconfiguration:

- **Stateful address configuration**

  Address configuration is based on the use of a stateful address configuration protocol, such as DHCPv6, to obtain addresses and other configuration options.

- **Stateless address configuration**

  Address configuration is based on the receipt of Router Advertisement messages that contain one or more Prefix Information options.

N/OS 8.3 supports stateless address configuration.

Stateless address configuration allows hosts on a link to configure themselves with link-local addresses and with addresses derived from prefixes advertised by local routers. Even if no router is present, hosts on the same link can configure themselves with link-local addresses and communicate without manual configuration.

IPv6 Interfaces

Each IPv6 interface supports multiple IPv6 addresses. You can manually configure up to two IPv6 addresses for each interface, or you can allow the switch to use stateless autoconfiguration.

You can manually configure two IPv6 addresses for each interface, as follows:

● Initial IPv6 address is a global unicast or anycast address.

```
RS G8124-E(config)# interface ip <interface number>
RS G8124-E(config-ip-if)# ipv6 address <IPv6 address>
```

Note that you cannot configure both addresses as anycast. If you configure an anycast address on the interface you must also configure a global unicast address on that interface.

● Second IPv6 address can be a unicast or anycast address.

```
RS G8124-E(config-ip-if)# ipv6 secaddr6 <IPv6 address>
RS G8124-E(config-ip-if)# exit
```

You cannot configure an IPv4 address on an IPv6 management interface. Each interface can be configured with only one address type: either IPv4 or IPv6, but not both. When changing between IPv4 and IPv6 address formats, the prior address settings for the interface are discarded.

Each IPv6 interface can belong to only one VLAN. Each VLAN can support only one IPv6 interface. Each VLAN can support multiple IPv4 interfaces.

Use the following commands to configure the IPv6 gateway:

```
RS G8124-E(config)# ip gateway6 1 address <IPv6 address>
RS G8124-E(config)# ip gateway6 1 enable
```

IPv6 gateway 1 is reserved for IPv6 data interfaces.IPv6 gateway 4 is the default IPv6 management gateway.

Neighbor Discovery

The switch uses Neighbor Discovery protocol (ND) to gather information about other router and host nodes, including the IPv6 addresses. Host nodes use ND to configure their interfaces and perform health detection. ND allows each node to determine the link-layer addresses of neighboring nodes and to keep track of each neighbor's information. A neighboring node is a host or a router linked directly to the switch. The switch supports Neighbor Discovery as described in RFC 4861.

Neighbor Discovery Overview

Neighbor Discover messages allow network nodes to exchange information, as follows:

- *Neighbor Solicitations* allow a node to discover information about other nodes.

- *Neighbor Advertisements* are sent in response to Neighbor Solicitations. The Neighbor Advertisement contains information required by nodes to determine the link-layer address of the sender, and the sender's role on the network.

- IPv6 hosts use *Router Solicitations* to discover IPv6 routers. When a router receives a Router Solicitation, it responds immediately to the host.

- Routers uses *Router Advertisements* to announce its presence on the network, and to provide its address prefix to neighbor devices. IPv6 hosts listen for Router Advertisements, and uses the information to build a list of default routers. Each host uses this information to perform autoconfiguration of IPv6 addresses.

- *Redirect messages* are sent by IPv6 routers to inform hosts of a better first-hop address for a specific destination. Redirect messages are only sent by routers for unicast traffic, are only unicast to originating hosts, and are only processed by hosts.

ND configuration for general advertisements, flags, and interval settings, as well as for defining prefix profiles for router advertisements, is performed on a per-interface basis using the following commands:

```
RS G8124-E(config)# interface ip <interface number>
RS G8124-E(config-ip-if)# [no] ipv6 nd ?
RS G8124-E(config-ip-if)# exit
```

To add or remove entries in the static neighbor cache, use the following command:

```
RS G8124-E(config)# [no] ip neighbors ?
```

To view the neighbor cache table, use the following command:

```
RS G8124-E(config)# show ipv6 neighbors ?
```

To view the neighbor cache counters, use the following command:

```
RS G8124-E(config)# show ipv6 neighbors counters
```

To clear the neighbor cache counters, use the following command:

```
RS G8124-E(config)# clear ipv6 neighbors counters
```

Host vs. Router

Each IPv6 interface can be configured as a router node or a host node, as follows:

- A router node's IP address is configured manually. Router nodes can send Router Advertisements.

- A host node's IP address can be autoconfigured. Host nodes listen for Router Advertisements that convey information about devices on the network.

Note: When IP forwarding is turned on, all IPv6 interfaces configured on the switch can forward packets.

You can configure each IPv6 interface as either a host node or a router node. You can manually assign an IPv6 address to an interface in host mode, or the interface can be assigned an IPv6 address by an upstream router, using information from router advertisements to perform stateless auto-configuration.

To set an interface to host mode, use the following command:

```
RS G8124-E(config)# interface ip <interface number>
RS G8124-E(config-ip-if)# ip6host
RS G8124-E(config-ip-if)# exit
```

The G8124-E supports up to 1156 IPv6 routes.

Supported Applications

The following applications have been enhanced to provide IPv6 support.

- **Ping**

  The **ping** command supports IPv6 addresses. Use the following format to ping an IPv6 address:

  ping *<host name>|<IPv6 address>* [**-n** *<tries (0-4294967295)>*]
  [**-w** *<msec delay (0-4294967295)>*] [**-l** *<length (0/32-65500/2080)>*]
  [**-s** *<IP source>*] [**-v** *<TOS (0-255)>*] [**-f**] [**-t**]

  To ping a link-local address (begins with FE80), provide an interface index, as follows:

  ping *<IPv6 address>%<Interface index>* [**-n** *<tries (0-4294967295)>*]
  [**-w** *<msec delay (0-4294967295)>*] [**-l** *<length (0/32-65500/2080)>*]
  [**-s** *<IP source>*] [**-v** *<TOS (0-255)>*] [**-f**] [**-t**]

- **Traceroute**

  The **traceroute** command supports IPv6 addresses (but not link-local addresses).
  Use the following format to perform a traceroute to an IPv6 address:

  traceroute *<host name>| <IPv6 address>* [*<max-hops (1-32)>*
  [*<msec delay (1-4294967295)>*]]

- **Telnet server**

  The **telnet** command supports IPv6 addresses (but not link-local addresses).
  Use the following format to Telnet into an IPv6 interface on the switch:

  telnet *<host name>| <IPv6 address>* [*<port>*]

- **Telnet client**

  The **telnet** command supports IPv6 addresses (but not link-local addresses).
  Use the following format to Telnet to an IPv6 address:

  telnet *<host name>| <IPv6 address>* [*<port>*]

- **HTTP/HTTPS**

  The HTTP/HTTPS servers support both IPv4 and IPv6 connections.

- **SSH**

  Secure Shell (SSH) connections over IPv6 are supported (but not link-local addresses). The following syntax is required from the client:

  ssh -u *<IPv6 address>*

  Example:

  ssh -u 2001:2:3:4:0:0:0:142

- **TFTP**

  The TFTP commands support both IPv4 and IPv6 addresses. Link-local addresses are not supported.

- **FTP**

  The FTP commands support both IPv4 and IPv6 addresses. Link-local addresses are not supported.

- **DNS client**

  DNS commands support both IPv4 and IPv6 addresses. Link-local addresses are not supported. Use the following command to specify the type of DNS query to be sent first:

  ```
  RS G8124-E(config)# ip dns ipv6 request-version {ipv4|ipv6}
  ```

  If you set the request version to `ipv4`, the DNS application sends an *A* query first, to resolve the hostname with an IPv4 address. If no *A* record is found for that hostname (no IPv4 address for that hostname) an *AAAA* query is sent to resolve the hostname with a IPv6 address.

  If you set the request version to `ipv6`, the DNS application sends an *AAAA* query first, to resolve the hostname with an IPv6 address. If no *AAAA* record is found for that hostname (no IPv6 address for that hostname) an *A* query is sent to resolve the hostname with an IPv4 address.

Configuration Guidelines

When you configure an interface for IPv6, consider the following guidelines:

- Support for subnet router anycast addresses is not available.
- A single interface can accept either IPv4 or IPv6 addresses, but not both IPv4 and IPv6 addresses.
- A single interface can accept multiple IPv6 addresses.
- A single interface can accept only one IPv4 address.
- If you change the IPv6 address of a configured interface to an IPv4 address, all IPv6 settings are deleted.
- A single VLAN can support only one IPv6 interface.
- Health checks are not supported for IPv6 gateways.
- IPv6 interfaces support Path MTU Discovery. The CPU's MTU is fixed at 1500 bytes.
- Support for jumbo frames (1,500 to 9,216 byte MTUs) is limited. Any jumbo frames intended for the CPU must be fragmented by the remote node. The switch can re-assemble fragmented packets up to 9k. It can also fragment and transmit jumbo packets received from higher layers.

IPv6 Configuration Examples

This section provides steps to configure IPv6 on the switch.

IPv6 Example 1

The following example uses IPv6 host mode to autoconfigure an IPv6 address for the interface. By default, the interface is assigned to VLAN 1.

1. Enable IPv6 host mode on an interface.

```
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip6host
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

2. Configure the IPv6 default gateway.

```
RS G8124-E(config)# ip gateway6 1 address
    2001:BA98:7654:BA98:FEDC:1234:ABCD:5412
RS G8124-E(config)# ip gateway6 1 enable
```

3. Verify the interface address.

```
RS G8124-E(config)# show interface ip 2
```

IPv6 Example 2

Use the following example to manually configure IPv6 on an interface.

1. Assign an IPv6 address and prefix length to the interface.

```
RS G8124-E(config)# interface ip 3
RS G8124-E(config-ip-if)# ipv6 address
    2001:BA98:7654:BA98:FEDC:1234:ABCD:5214
RS G8124-E(config-ip-if)# ipv6 prefixlen 64
RS G8124-E(config-ip-if)# ipv6 seccaddr6 2003::1 32
RS G8124-E(config-ip-if)# vlan 2
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

The secondary IPv6 address is compressed, and the prefix length is 32.

2. Configure the IPv6 default gateway.

```
RS G8124-E(config)# ip gateway6 1 address
    2001:BA98:7654:BA98:FEDC:1234:ABCD:5412
RS G8124-E(config)# ip gateway6 1 enable
```

3. Configure router advertisements for the interface (optional)

```
RS G8124-E(config)# interface ip 3
RS G8124-E(config-ip-if)# no ipv6 nd suppress-ra
```

4. Verify the configuration.

```
RS G8124-E(config-ip-if)# show layer3
```

Chapter 22. IPsec with IPv6

Internet Protocol Security (IPsec) is a protocol suite for securing Internet Protocol (IP) communications by authenticating and encrypting each IP packet of a communication session. IPsec also includes protocols for establishing mutual authentication between agents at the beginning of the session and negotiation of cryptographic keys to be used during the session.

Since IPsec was implemented in conjunction with IPv6, all implementations of IPv6 must contain IPsec. To support the National Institute of Standards and Technology (NIST) recommendations for IPv6 implementations, Lenovo Networking OS IPv6 feature compliance has been extended to include the following IETF RFCs, with an emphasis on IP Security (IPsec), Internet Key Exchange version 2, and authentication/confidentiality for OSPFv3:

- RFC 4301 for IPv6 security
- RFC 4302 for the IPv6 Authentication Header
- RFCs 2404, 2410, 2451, 3602, and 4303 for IPv6 Encapsulating Security Payload (ESP), including NULL encryption, CBC-mode 3DES and AES ciphers, and HMAC-SHA-1-96.
- RFCs 4306, 4307, 4718, and 4835 for IKEv2 and cryptography
- RFC 4552 for OSPFv3 IPv6 authentication
- RFC 5114 for Diffie-Hellman groups

Note: This implementation of IPsec supports DH groups 1, 2, 5, 14, and 24.

The following topics are discussed in this chapter:

- "IPsec Protocols" on page 316
- "Using IPsec with the RackSwitch G8124-E" on page 317

IPsec Protocols

The Lenovo N/OS implementation of IPsec supports the following protocols:

- Authentication Header (AH)
  AHs provide connectionless integrity out and data origin authentication for IP packets. They also provide protection against replay attacks. In IPv6, the AH protects the AH itself, the Destination Options extension header after the AH, and the IP payload. It also protects the fixed IPv6 header and all extension headers before the AH, except for the mutable fields DSCP, ECN, Flow Label, and Hop Limit. AH is defined in RFC 4302.

- Encapsulating Security Payload (ESP)
  ESPs provide confidentiality, data origin authentication, integrity, an anti-replay service (a form of partial sequence integrity), and some traffic flow confidentiality. ESPs may be applied alone or in combination with an AH. ESP is defined in RFC 4303.

- Internet Key Exchange Version 2 (IKEv2)
  IKEv2 is used for mutual authentication between two network elements. An IKE establishes a security association (SA) that includes shared secret information to efficiently establish SAs for ESPs and AHs, and a set of cryptographic algorithms to be used by the SAs to protect the associated traffic. IKEv2 is defined in RFC 4306.

Using IKEv2 as the foundation, IPsec supports ESP for encryption and/or authentication, and/or AH for authentication of the remote partner.

Both ESP and AH rely on security associations. A security association (SA) is the bundle of algorithms and parameters (such as keys) that encrypt and authenticate a particular flow in one direction.

Using IPsec with the RackSwitch G8124-E

IPsec supports the fragmentation and reassembly of IP packets that occurs when data goes to and comes from an external device. The RackSwitch G8124-E acts as an end node that processes any fragmentation and reassembly of packets but does not forward the IPsec traffic. :You must authenticate the IKEv2 key following the directions in before you can use IPsec.

The security protocol for the session key is either ESP or AH. Outgoing packets are labeled with the SA SPI (Security Parameter Index), which the remote device will use in its verification and decryption process.

Every outgoing IPv6 packet is checked against the IPsec policies in force. For each outbound packet, after the packet is encrypted, the software compares the packet size with the MTU size that it either obtains from the default minimum maximum transmission unit (MTU) size (1500) or from path MTU discovery. If the packet size is larger than the MTU size, the receiver drops the packet and sends a message containing the MTU size to the sender. The sender then fragments the packet into smaller pieces and retransmits them using the correct MTU size.

The maximum traffic load for each IPsec packet is limited to the following:

- IKEv2 SAs: 5

- IPsec SAs: 10 (5 SAs in each direction)

- SPDs: 20 (10 policies in each direction)

IPsec is implemented as a software cryptography engine designed for handling control traffic, such as network management. IPsec is not designed for handling data traffic, such as a VPN.

Setting up Authentication

Before you can use IPsec, you need to have key policy authentication in place. There are two types of key policy authentication:

- Preshared key (default)

  The parties agree on a shared, secret key that is used for authentication in an IPsec policy. During security negotiation, information is encrypted before transmission by using a session key created by using a Diffie-Hellman calculation and the shared, secret key. Information is decrypted on the receiving end using the same key. One IPsec peer authenticates the other peer's packet by decryption and verification of the hash inside the packet (the hash inside the packet is a hash of the preshared key). If authentication fails, the packet is discarded.

- Digital certificate (using RSA algorithms)

  The peer being validated must hold a digital certificate signed by a trusted Certificate Authority and the private key for that digital certificate. The side performing the authentication only needs a copy of the trusted certificate authorities digital certificate. During IKEv2 authentication, the side being validated sends a copy of the digital certificate and a hash value signed using the private key. The certificate can be either generated or imported.

Note: During the IKEv2 negotiation phase, the digital certificate takes precedence over the preshared key.

Creating an IKEv2 Proposal

With IKEv2, a single policy can have multiple encryption and authentication types, as well as multiple integrity algorithms.

To create an IKEv2 proposal:

1. Enter IKEv2 proposal mode.

```
RS G8124-E(config)# ikev2 proposal
```

2. Set the DES encryption algorithm.

```
RS G8124-E(config-ikev2-prop)# encryption 3des|aes-cbc  (default: 3des)
```

3. Set the authentication integrity algorithm type.

```
RS G8124-E(config-ikev2-prop)# integrity sha1 (default: sha1)
```

4. Set the Diffie-Hellman group.

```
RS G8124-E(config-ikev2-prop)# group 1|2|5|14|24 (default: 24)
```

Importing an IKEv2 Digital Certificate

To import an IKEv2 digital certificate for authentication:

1. Import the CA certificate file.

```
RS G8124-E(config)# copy tftp ca-cert address <hostname or IPv4 address>

Source file name: <path and filename of CA certificate file>
Confirm download operation [y/n]: y
```

2. Import the host key file.

```
RS G8124-E(config)# copy tftp host-key address <hostname or IPv4 address>

Source file name: <path and filename of host private key file>
Confirm download operation [y/n]: y
```

3. Import the host certificate file.

```
RS G8124-E(config)# copy tftp host-cert address <hostname or IPv4 address>

Source file name: <path and filename of host certificate file>
Confirm download operation [y/n]: y
```

Note: When prompted for the port to use for download the file, if you used a management port to connect the switch to the server, enter **mgt**, otherwise enter **data**.

Generating an IKEv2 Digital Certificate

To create an IKEv2 digital certificate for authentication:

1. Create an HTTPS certificate defining the information you want to be used in the various fields.

```
RS G8124-E(config)# access https generate-certificate
Country Name (2 letter code) [US]:
State or Province Name (full name) [CA]:
Locality Name (eg, city) [Santa Clara]:
Organization Name (eg, company) [Lenovo]:
Organizational Unit Name (eg, section) [Engineering]:
Common Name (eg, YOUR name) [10.240.226.241]:
Email (eg, email address) []:
Confirm generat'eywing certificate? [y/n]: y
Generating certificate. Please wait (approx 30 seconds)
restarting SSL agent
```

2. Save the HTTPS certificate.

   The certificate is valid only until the switch is rebooted. To save the certificate so that it is retained beyond reboot or power cycles, use the following command:

```
RS G8124-E(config)# access https save-certificate
```

3. Enable IKEv2 RSA-signature authentication:

```
RS G8124-E(config)# access https enable
```

Enabling IKEv2 Preshared Key Authentication

To set up IKEv2 preshared key authentication:

1. Enter the local preshared key.

```
RS G8124-E(config)# ikev2 preshare-key local <preshared key, a string of 1-256 chars>
```

2. If asymmetric authentication is supported, enter the remote key:

```
RS G8124-E(config)# ikev2 preshare-key remote <preshared key> <IPv6 host>
```

   where the following parameters are used:

   - *preshared key*   A string of 1-256 characters

   - *IPv6 host*   An IPv6-format host, such as "3000::1"

3. Set up the IKEv2 identification type by entering *one* of the following commands:

```
RS G8124-E(config)# ikev2 identity local address (use an IPv6 address)
RS G8124-E(config)# ikev2 identity local email <email address>
RS G8124-E(config)# ikev2 identity local fqdn <domain name>
```

   To disable IKEv2 RSA-signature authentication method and enable preshared key authentication, enter:

```
RS G8124-E(config)# access https disable
```

Setting Up a Key Policy

When configuring IPsec, you must define a key policy. This key policy can be either manual or dynamic. Either way, configuring a policy involves the following steps:

- Create a transform set—This defines which encryption and authentication algorithms are used.

- Create a traffic selector—This describes the packets to which the policy applies.

- Establish an IPsec policy.

- Apply the policy.

1. To define which encryption and authentication algorithms are used, create a transform set:

```
RS G8124-E(config)# ipsec transform-set <transform ID> <encryption method>
    <integrity algorithm> <AH authentication algorithm>
```

where the following parameters are used:

- *transform ID*   A number from 1-10

- *encryption method*   One of the following: **esp-3des** | **esp-aes-cbc** | **esp-null**

- *integrity algorithm*   One of the following: **esp-sha1** | **none**

- *AH authentication algorithm*   One of the following: **ah-sha1** | **ah-md5** | **none**

2. Decide whether to use tunnel or transport mode. The default mode is transport.

```
RS G8124-E(config)# ipsec transform-set tunnel|transport
```

3. To describe the packets to which this policy applies, create a traffic selector using the following command:

```
RS G8124-E(config)# ipsec traffic-selector <traffic selector number> permit|deny
    any|icmp <type/any>|tcp ><source IP address/any> <destination IP address/any> [<prefix
    length>]
```

where the following parameters are used:

- *traffic selector number*   an integer from 1-10

- **permit**|**deny**   whether or not to permit IPsec encryption of traffic that meets the criteria specified in this command

- **any**   apply the selector to any type of traffic

- **icmp** *<type>*|**any**   only apply the selector only to ICMP traffic of the specified *type* (an integer from 1-255) or to any ICMP traffic

- **tcp**   only apply the selector to TCP traffic

- *source IP address*|**any**   the source IP address in IPv6 format or "any" source

- *destination IP address* | **any**   the destination IP address in IPv6 format or "any"
  destination

- *prefix length*   (Optional) the length of the destination IPv6
  prefix; an integer from 1-128

Permitted traffic that matches the policy in force is encrypted, while denied traffic that matches the policy in force is dropped. Traffic that does not match the policy bypasses IPsec and passes through *clear* (unencrypted).

4. Choose whether to use a manual or a dynamic policy.

Using a Manual Key Policy

A manual policy involves configuring policy and manual SA entries for local and remote peers.

To configure a manual key policy, you need:

- The IP address of the peer in IPv6 format (for example, "3000::1").
- Inbound/Outbound session keys for the security protocols.

You can then assign the policy to an interface. The peer represents the other end of the security association. The security protocol for the session key can be either ESP or AH.

To create and configure a manual policy:

1. Enter a manual policy to configure.

```
RS G8124-E(config)#ipsec manual-policy <policy number>
```

2. Configure the policy.

```
RS G8124-E(config-ipsec-manual)#peer <peer's IPv6 address>
RS G8124-E(config-ipsec-manual)#traffic-selector <IPsec traffic selector>
RS G8124-E(config-ipsec-manual)#transform-set <IPsec transform set>
RS G8124-E(config-ipsec-manual)#in-ah auth-key <inbound AH IPsec key>
RS G8124-E(config-ipsec-manual)#in-ah auth-spi <inbound AH IPsec SPI>
RS G8124-E(config-ipsec-manual)#in-esp cipher-key <inbound ESP cipher key>
RS G8124-E(config-ipsec-manual)#in-esp auth-spi <inbound ESP SPI>
RS G8124-E(config-ipsec-manual)#in-esp auth-key <inbound ESP authenticator key>
RS G8124-E(config-ipsec-manual)#out-ah auth-key <outbound AH IPsec key>
RS G8124-E(config-ipsec-manual)#out-ah auth-spi <outbound AH IPsec SPI>
RS G8124-E(config-ipsec-manual)#out-esp cipher-key <outbound ESP cipher key>
RS G8124-E(config-ipsec-manual)#out-esp auth-spi <outbound ESP SPI>
RS G8124-E(config-ipsec-manual)#out-esp auth-key <outbound ESP authenticator key>
```

where the following parameters are used:

- *peer's IPv6 address*   The IPv6 address of the peer (for example,
  3000::1)

- *IPsec traffic-selector*   A number from 1-10

- *IPsec of transform-set*   A number from 1-10

- *inbound AH IPsec key*   The inbound AH key code, in hexadecimal

- *inbound AH IPsec SPI*   A number from 256-4294967295

- *inbound ESP cipher key*   The inbound ESP key code, in hexadecimal

- *inbound ESP SPI*   A number from 256-4294967295

- *inbound ESP authenticator key*   The inbound ESP authenticator key code, in hexadecimal

- *outbound AH IPsec key*   The outbound AH key code, in hexadecimal

- *outbound AH IPsec SPI*   A number from 256-4294967295

- *outbound ESP cipher key*   The outbound ESP key code, in hexadecimal

- *outbound ESP SPI*   A number from 256-4294967295

- *outbound ESP authenticator key*   The outbound ESP authenticator key code, in hexadecimal

Note:

- When configuring a manual policy ESP, the ESP authenticator key is optional.

- If using third-party switches, the IPsec manual policy session key must be of fixed length as follows:
  - For AH key: SHA1 is 20 bytes; MD5 is 16 bytes
  - For ESP cipher key: 3DES is 24 bytes; AES-cbc is 24 bytes; DES is 8 bytes
  - For ESP auth key: SHA1 is 20 bytes; MD5 is 16 bytes

3. After you configure the IPSec policy, you need to apply it to the interface to enforce the security policies on that interface and save it to keep it in place after a reboot. To accomplish this, enter:

```
RS G8124-E(config-ip)#interface ip <IP interface number, 1-128>
RS G8124-E(config-ip-if)#address <IPv6 address>
RS G8124-E(config-ip-if)#ipsec manual-policy <policy index, 1-10>
RS G8124-E(config-ip-if)#enable (enable the IP interface)
RS G8124-E#write (save the current configuration)
```

Using a Dynamic Key Policy

When you use a dynamic key policy, the first packet triggers IKE and sets the IPsec SA and IKEv2 SA. The initial packet negotiation also determines the lifetime of the algorithm, or how long it stays in effect. When the key expires, a new key is automatically created. This helps prevent break-ins.

To configure a dynamic key policy:

1. Choose a dynamic policy to configure.

```
RS G8124-E(config)#ipsec dynamic-policy <policy number>
```

2. Configure the policy.

```
RS G8124-E(config-ipsec-dynamic)#peer <peer's IPv6 address>
RS G8124-E(config-ipsec-dynamic)#traffic-selector <index of traffic selector>
RS G8124-E(config-ipsec-dynamic)#transform-set <index of transform set>
RS G8124-E(config-ipsec-dynamic)#sa-lifetime <SA lifetime, in seconds>
RS G8124-E(config-ipsec-dynamic)#pfs enable|disable
```

where the following parameters are used:

- *peer's IPv6 address*   The IPv6 address of the peer (for example, 3000::1)

- *index of traffic-selector*   A number from1-10

- *index of transform-set*   A number from1-10

- *SA lifetime, in seconds*   The length of time the SA is to remain in effect; an integer from120-86400

- **pfs enable|disable**   Whether to enable or disable the perfect forward security feature. The default is **disable**.

Note: In a dynamic policy, the AH and ESP keys are created by IKEv2.

3. After you configure the IPSec policy, you need to apply it to the interface to enforce the security policies on that interface and save it to keep it in place after a reboot. To accomplish this, enter:

```
RS G8124-E(config-ip)#interface ip <IP interface number, 1-128>
RS G8124-E(config-ip-if)#address <IPv6 address>
RS G8124-E(config-ip-if)#ipsec dynamic-policy <policy index, 1-10>
RS G8124-E(config-ip-if)#enable (enable the IP interface)
RS G8124-E#write (save the current configuration)
```

Chapter 23. Routing Information Protocol

In a routed environment, routers communicate with one another to keep track of available routes. Routers can learn about available routes dynamically using the Routing Information Protocol (RIP). Lenovo Networking OS software supports RIP version 1 (RIPv1) and RIP version 2 (RIPv2) for exchanging TCP/IPv4 route information with other routers.

Note: Lenovo N/OS 8.3 does not support IPv6 for RIP.

Distance Vector Protocol

RIP is known as a distance vector protocol. The vector is the network number and next hop, and the distance is the metric associated with the network number. RIP identifies network reachability based on metric, and metric is defined as hop count. One hop is considered to be the distance from one switch to the next, which typically is 1.

When a switch receives a routing update that contains a new or changed destination network entry, the switch adds 1 to the metric value indicated in the update and enters the network in the routing table. The IPv4 address of the sender is used as the next hop.

Stability

RIP includes a number of other stability features that are common to many routing protocols. For example, RIP implements the split horizon and hold-down mechanisms to prevent incorrect routing information from being propagated.

RIP prevents routing loops from continuing indefinitely by implementing a limit on the number of hops allowed in a path from the source to a destination. The maximum number of hops in a path is 15. The network destination network is considered unreachable if increasing the metric value by 1 causes the metric to be 16 (that is infinity). This limits the maximum diameter of a RIP network to less than 16 hops.

RIP is often used in stub networks and in small autonomous systems that do not have many redundant paths.

Routing Updates

RIP sends routing-update messages at regular intervals and when the network topology changes. Each router "advertises" routing information by sending a routing information update every 30 seconds. If a router doesn't receive an update from another router for 180 seconds, those routes provided by that router are declared invalid. The routes are removed from the routing table, but they remain in the RIP routes table. After another 120 seconds without receiving an update for those routes, the routes are removed from respective regular updates.

When a router receives a routing update that includes changes to an entry, it updates its routing table to reflect the new route. The metric value for the path is increased by 1, and the sender is indicated as the next hop. RIP routers maintain only the best route (the route with the lowest metric value) to a destination.

For more information, see the Configuration section, Routing Information Protocol Configuration in the *Lenovo Networking OS Command Reference*.

RIPv1

RIP version 1 use broadcast User Datagram Protocol (UDP) data packets for the regular routing updates. The main disadvantage is that the routing updates do not carry subnet mask information. Hence, the router cannot determine whether the route is a subnet route or a host route. It is of limited usage after the introduction of RIPv2. For more information about RIPv1 and RIPv2, refer to RFC 1058 and RFC 2453.

RIPv2

RIPv2 is the most popular and preferred configuration for most networks. RIPv2 expands the amount of useful information carried in RIP messages and provides a measure of security. For a detailed explanation of RIPv2, refer to RFC 1723 and RFC 2453.

RIPv2 improves efficiency by using multicast UDP (address 224.0.0.9) data packets for regular routing updates. Subnet mask information is provided in the routing updates. A security option is added for authenticating routing updates, by using a shared password. N/OS supports using clear password for RIPv2.

RIPv2 in RIPv1 Compatibility Mode

N/OS allows you to configure RIPv2 in RIPv1compatibility mode, for using both RIPv2 and RIPv1 routers within a network. In this mode, the regular routing updates use broadcast UDP data packet to allow RIPv1 routers to receive those packets. With RIPv1 routers as recipients, the routing updates have to carry natural or host mask. Hence, it is not a recommended configuration for most network topologies.

Note: When using both RIPv1 and RIPv2 within a network, use a single subnet mask throughout the network.

RIP Features

N/OS provides the following features to support RIPv1 and RIPv2:

Poison

Simple split horizon in RIP scheme omits routes learned from one neighbor in updates sent to that neighbor. That is the most common configuration used in RIP, that is setting this Poison to DISABLE. Split horizon with poisoned reverse includes such routes in updates, but sets their metrics to 16. The disadvantage of using this feature is the increase of size in the routing updates.

Triggered Updates

Triggered updates are an attempt to speed up convergence. When Triggered Updates is enabled, whenever a router changes the metric for a route, it sends update messages almost immediately, without waiting for the regular update interval. It is recommended to enable Triggered Updates.

Multicast

RIPv2 messages use IPv4 multicast address (224.0.0.9) for periodic broadcasts. Multicast RIPv2 announcements are not processed by RIPv1 routers. IGMP is not needed since these are inter-router messages which are not forwarded.

To configure RIPv2 in RIPv1 compatibility mode, set multicast to `disable`, and set version to `both`.

Default

The RIP router can listen and supply a default route, usually represented as IPv4 0.0.0.0 in the routing table. When a router does not have an explicit route to a destination network in its routing table, it uses the default route to forward those packets.

Metric

The metric field contains a configurable value between 1 and 15 (inclusive) which specifies the current metric for the interface. The metric value typically indicates the total number of hops to the destination. The metric value of 16 represents an unreachable destination.

Authentication

RIPv2 authentication uses plaintext password for authentication. If configured using Authentication password, then it is necessary to enter an authentication key value.

The following method is used to authenticate an RIP message:

- If the router is not configured to authenticate RIPv2 messages, then RIPv1 and unauthenticated RIPv2 messages are accepted; authenticated RIPv2 messages are discarded.

- If the router is configured to authenticate RIPv2 messages, then RIPv1 messages and RIPv2 messages which pass authentication testing are accepted; unauthenticated and failed authentication RIPv2 messages are discarded.

For maximum security, RIPv1 messages are ignored when authentication is enabled; otherwise, the routing information from authenticated messages is propagated by RIPv1 routers in an unauthenticated manner.

RIP Configuration Example

The following is an example of RIP configuration.

Note: An interface with RIP disabled uses all the default values of the RIP, no matter how the RIP parameters are configured for that interface. RIP sends out RIP regular updates to include an UP interface, but not a DOWN interface.

1. Add VLANs for routing interfaces.

```
>> (config)# vlan 2
>> (config-vlan)# exit
>> (config)# interface port 2
>> (config-if)# switchport mode trunk
>> (config-if)# switchport trunk allowed vlan add 2

Port 2 is an UNTAGGED port and its current PVID is 1.
Confirm changing PVID from 1 to 2 [y/n]: y
>> (config-if)# exit

>> (config)# vlan 3
>> (config-vlan)# exit
>> (config)# interface port 3
>> (config-if)# switchport mode trunk
>> (config-if)# switchport trunk allowed vlan add 3
>> (config-if)# exit
Port 3 is an UNTAGGED port and its current PVID is 1.
Confirm changing PVID from 1 to 3 [y/n]: y
```

2. Add IP interfaces with IPv4 addresses to VLANs.

```
>> # interface ip 2
>> (config-ip-if)# enable
>> (config-ip-if)# ip address 102.1.1.1
>> (config-ip-if)# vlan 2
>> (config-ip-if)# exit
>> # interface ip 3
>> (config-ip-if)# enable
>> (config-ip-if)# ip address 103.1.1.1
>> (config-ip-if)# vlan 3
```

3. Turn on RIP globally and enable RIP for each interface.

```
>> # router rip
>> (config-router-rip)# enable
>> (config-router-rip)# exit
>> # interface ip 2
>> (config-ip-if)# ip rip enable
>> (config-ip-if)# exit
>> # interface ip 3
>> (config-ip-if)# ip rip enable
>> (config-ip-if)# exit
```

Use the following command to check the current valid routes in the routing table of the switch:

```
>> # show ip route
```

For those RIP routes learned within the garbage collection period, that are routes phasing out of the routing table with metric 16, use the following command:

```
>> # show ip rip routes
```

Locally configured static routes do not appear in the RIP Routes table.

Chapter 24. Internet Group Management Protocol

Internet Group Management Protocol (IGMP) is used by IPv4 Multicast routers (Mrouters) to learn about the existence of host group members on their directly attached subnet. The IPv4 Mrouters get this information by broadcasting IGMP Membership Queries and listening for IPv4 hosts reporting their host group memberships. This process is used to set up a client/server relationship between an IPv4 multicast source that provides the data streams and the clients that want to receive the data. The switch supports three versions of IGMP:

- IGMPv1: Defines the method for hosts to join a multicast group. However, this version does not define the method for hosts to leave a multicast group. See RFC 1112 for details.

- IGMPv2: Adds the ability for a host to signal its desire to leave a multicast group. See RFC 2236 for details.

- IGMPv3: Adds support for source filtering by which a host can report interest in receiving packets only from specific source addresses, or from all but specific source addresses, sent to a particular multicast address. See RFC 3376 for details.

The G8124-E can perform IGMP Snooping, and connect to static Mrouters. The G8124-E can act as a Querier, and participate in the IGMP Querier election process.

The following topics are discussed in this chapter:

IGMP Terms

The following are commonly used IGMP terms:

- Multicast traffic: Flow of data from one source to multiple destinations.

- Group: A multicast stream to which a host can join. Multicast groups have IP addresses in the range: 224.0.1.0 to 239.255.255.255.

- IGMP Querier: A router or switch in the subnet that generates *Membership Queries.*

- IGMP Snooper: A Layer 3 device that forwards multicast traffic only to hosts that are interested in receiving multicast data. This device can be a router or a Layer 3 switch.

- Multicast Router: A router configured to make routing decisions for multicast traffic. The router identifies the type of packet received (unicast or multicast) and forwards the packet to the intended destination.

- IGMP Proxy: A device that filters Join messages and Leave messages sent upstream to the Mrouter to reduce the load on the Mrouter.

- Membership Report: A report sent by the host that indicates an interest in receiving multicast traffic from a multicast group.

- Leave: A message sent by the host when it wants to leave a multicast group.

- FastLeave: A process by which the switch stops forwarding multicast traffic to a port as soon as it receives a Leave message.

- Membership Query: Message sent by the Querier to verify if hosts are listening to a group.

- General Query: A *Membership Query* sent to all hosts. The Group address field for general queries is 0.0.0.0 and the destination address is 224.0.0.1.

- Group-specific Query: A *Membership Query* sent to all hosts in a multicast group.

How IGMP Works

When IGMP is not configured, switches forward multicast traffic through all ports, increasing network load. When IGMPv2 is configured on a switch, multicast traffic flows as follows:

- A server sends multicast traffic to a multicast group.

- The Mrouter sends *Membership Queries* to the switch, which forwards them to all ports in a given VLAN.

- Hosts respond with *Membership Reports* if they want to join a group. The switch forwards these reports to the Mrouter.

- The switch forwards multicast traffic only to hosts that have joined a group.

- The Mrouter periodically sends *Membership Queries* to ensure that a host wants to continue receiving multicast traffic. If a host does not respond, the IGMP Snooper stops sending traffic to the host.

- The host can initiate the Leave process by sending an IGMP Leave packet to the IGMP Snooper.

- When a host sends an IGMPv2 Leave packet, the IGMP Snooper queries to find out if any other host connected to the port is interested in receiving the multicast traffic. If it does not receive a Join message in response, the IGMP Snooper removes the group entry and passes on the information to the Mrouter.

The G8124-E supports the following:

- IGMP version 1, 2, and 3
- 20 Mrouters (in default switch profile) / 128 Mrouters (in HFT switch profile)

Note: Unknown multicast traffic is sent to all ports if the flood option is enabled and no Membership Report was learned for that specific IGMP group. If the flood option is disabled, unknown multicast traffic is discarded if no hosts or Mrouters are learned on a switch.

To enable or disable IGMP flood, use the following command:
```
RS G8124-E(config)# [no] ip ipmcfld
```

IGMP Capacity and Default Values

The following table lists the maximum and minimum values of the G8124-E variables.

Table 27. *G8124-E Capacity Table*

Variable	Maximum
IGMP Entries - Snoop	1000
VLANs - Snoop	1023
Static Mrouters	20
Dynamic Mrouters	20
Number of IGMP Filters	16

The following table lists the default settings for IGMP features and variables.

Table 28. *IGMP Default Configuration Settings*

Field	Default Value
Global IGMP State	Disabled
IGMP Querier	Disabled
IGMP Snooping	Disabled
IGMP Filtering	Disabled
IP Multicast (IPMC) Flood	Enabled
IGMP FastLeave	Disabled for all VLANs
IGMP Mrouter Timeout	255 Seconds
IGMP Report Timeout Variable	10 Seconds
IGMP Query-Interval Variable	125 Seconds
IGMP Robustness Variable	2
IGMPv3	Disabled
IGMPv3 number of sources	8 (The switch processes only the first eight sources listed in the IGMPv3 group record.) Valid range: 1 - 64
IGMPv3 - allow v1v2 Snooping	Enabled

IGMP Snooping

IGMP Snooping allows a switch to listen to the IGMP conversation between hosts and Mrouters. By default, a switch floods multicast traffic to all ports in a broadcast domain. With IGMP Snooping enabled, the switch learns the ports interested in receiving multicast data and forwards it only to those ports. IGMP Snooping conserves network resources.

The switch can sense IGMP *Membership Reports* from attached hosts and acts as a proxy to set up a dedicated path between the requesting host and a local IPv4 Mrouter. After the path is established, the switch blocks the IPv4 multicast stream from flowing through any port that does not connect to a host member, thus conserving bandwidth.

IGMP Querier

For IGMP Snooping to function, you must have an Mrouter on the network that generates IGMP Query packets. Enabling the IGMP Querier feature on the switch allows it to participate in the Querier election process. If the switch is elected as the Querier, it will send IGMP Query packets for the LAN segment.

Querier Election

If multiple Mrouters exist on the network, only one can be elected as a Querier. The Mrouters elect the one with the lowest source IPv4 address or MAC address as the Querier. The Querier performs all periodic membership queries. All other Mrouters (non-Queriers) do not send IGMP Query packets.

Note: When IGMP Querier is enabled on a VLAN, the switch performs the role of an IGMP Querier only if it meets the IGMP Querier election criteria.

Each time the Querier switch sends an IGMP Query packet, it initializes a *general query timer*. If a Querier receives a General Query packet from an Mrouter with a lower IP address or MAC address, it transitions to a non-Querier state and initializes an *other querier present timer*. When this timer expires, the Mrouter transitions back to the Querier state and sends a General Query packet.

Follow this procedure to configure IGMP Querier.

1. Enable IGMP and configure the source IPv4 address for IGMP Querier on a VLAN.

```
RS G8124-E(config)# ip igmp enable
RS G8124-E(config)# ip igmp querier vlan 2 source-ip 10.10.10.1
```

2. Enable IGMP Querier on the VLAN.

```
RS G8124-E(config)# ip igmp querier vlan 2 enable
```

3. Configure the querier election type and define the address.

```
RS G8124-E(config)# ip igmp querier vlan 2 election-type ipv4
```

4. Verify the configuration.

```
RS G8124-E# show ip igmp querier vlan 2

Current IGMP snooping Querier information:
 IGMP Querier information for vlan 2:
 Other IGMP querier - none
 Switch-querier enabled, current state: Querier
 Switch-querier type: Ipv4, address 10.10.10.1,
 Switch-querier general query interval: 125 secs,
 Switch-querier max-response interval: 100 'tenths of secs',
 Switch-querier startup interval: 31 secs, count: 2
 Switch-querier robustness: 2
 IGMP configured version is v3
 IGMP Operating version is v3
```

IGMP Groups

One IGMP entry is allocated for each unique join request, based on the VLAN and IGMP group address. If multiple ports join the same IGMP group using the same VLAN, only a single IGMP entry is used.

IGMPv3 Snooping

IGMPv3 includes new Membership Report messages that extend IGMP functionality. The switch provides snooping capability for all types of IGMPv3 *Membership Reports*.

IGMPv3 supports Source-Specific Multicast (SSM). SSM identifies session traffic by both source and group addresses.

The IGMPv3 implementation keeps records on the multicast hosts present in the network. If a host is already registered, when it receives a new IS_INC, TO_INC, IS_EXC, or TO_EXC report from same host, the switch makes the correct transition to new (port-host-group) registration based on the IGMPv3 RFC. The registrations of other hosts for the same group on the same port are not changed.

The G8124-E supports the following IGMPv3 filter modes:

● INCLUDE mode: The host requests membership to a multicast group and pro-vides a list of IPv4 addresses from which it wants to receive traffic.
● EXCLUDE mode: The host requests membership to a multicast group and provides a list of IPv4 addresses from which it does not want to receive traffic. This indicates that the host wants to receive traffic only from sources that are not part of the Exclude list. To disable snooping on EXCLUDE mode reports, use the following command:

```
RS G8124-E(config)# no ip igmp snoop igmpv3 exclude
```

By default, the G8124-E snoops the first eight sources listed in the IGMPv3 Group Record. Use the following command to change the number of snooping sources:

```
RS G8124-E(config)# ip igmp snoop igmpv3 sources <1-64>
```

IGMPv3 Snooping is compatible with IGMPv1 and IGMPv2 Snooping. To disable snooping on version 1 and version 2 reports, use the following command:

```
RS G8124-E(config)# no ip igmp snoop igmpv3 v1v2
```

IGMP Snooping Configuration Guidelines

Consider the following guidelines when you configure IGMP Snooping:

- IGMP operation is independent of the routing method. You can use RIP, OSPF, or static routes for Layer 3 routing.

- When multicast traffic flood is disabled, the multicast traffic sent by the multicast server is discarded if no hosts or Mrouters are learned on the switch.

- The Mrouter periodically sends IGMP Queries.

- The switch learns the Mrouter on the port connected to the router when it sees Query messages. The switch then floods the IGMP queries on all other ports including a LAG, if any.

- Multicast hosts send IGMP Reports as a reply to the IGMP Queries sent by the Mrouter.

- The switch can also learn an Mrouter when it receives a PIM `hello` packet from another device. However, an Mrouter learned from a PIM packet has a lower priority than an Mrouter learned from an IGMP Query. A switch overwrites an Mrouter learned from a PIM packet when it receives an IGMP Query on the same port.

- A host sends an IGMP Leave message to its multicast group. The expiration timer for this group is updated to 10 seconds. The Layer 3 switch sends IGMP Group-Specific Query to the host that had sent the Leave message. If the host does not respond with an IGMP Report during these 10 seconds, all the groups expire and the switch deletes the host from the IGMP groups table. The switch then proxies the IGMP Leave messages to the Mrouter.

IGMP Snooping Configuration Example

This section provides steps to configure IGMP Snooping on the G8124-E.

1. Configure port and VLAN membership on the switch.

2. Add VLANs to IGMP Snooping.

```
RS G8124-E(config)# ip igmp snoop vlan 1
```

3. Enable IGMPv3 Snooping (optional).

```
RS G8124-E(config)# ip igmp snoop igmpv3 enable
```

4. Enable the IGMP feature.

```
RS G8124-E(config)# ip igmp enable
```

5. View dynamic IGMP information.

```
RS G8124-E# show ip igmp groups

Command mode: All
Total entries: 5 Total IGMP groups: 2
Note: The <Total IGMP groups> number is computed as
the number of unique (Group, Vlan) entries!
Note: Local groups (224.0.0.x) are not snooped/relayed and will not
appear.
     Source          Group      VLAN   Port   Version Mode  Expires Fwd
-------------- --------------- ------- ------ -------- ----- ------- ---
    10.1.1.1       232.1.1.1      2      4       V3    INC    4:16   Yes
    10.1.1.5       232.1.1.1      2      4       V3    INC    4:16   Yes
        *          232.1.1.1      2      4       V3    INC     -     No
   10.10.10.43     235.0.0.1      9      1       V3    EXC    2:26   No
        *          235.0.0.1      9      1       V3    EXC     -     Yes

RS G8124-E# show ip igmp mrouter

SrcIP              VLAN    Port      Version   Expires   MRT       QRV
QQIC

------------------- ------- -------    --------- --------  -------
---- ----
10.1.1.1              2      21        V3         4:09     128         2
125
10.1.1.5              2      23        V2         4:09     125         -
-
10.10.10.43           9      24        V2         static   unknown     -
-
```

These commands display information about IGMP Groups and Mrouters learned by the switch.

Advanced Configuration Example: IGMP Snooping

Figure 34 shows an example topology. Switches B and C are configured with IGMP Snooping.

Figure 34. Topology
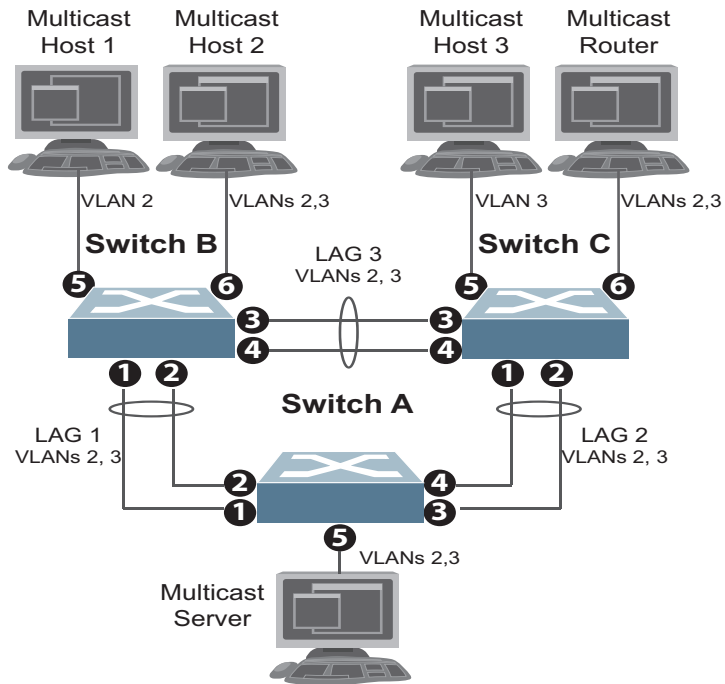
Devices in this topology are configured as follows:

- STG2 includes VLAN2; STG3 includes VLAN3.
- The multicast server sends IP multicast traffic for the following groups:
    - VLAN 2, 225.10.0.11 – 225.10.0.12, Source: 22.10.0.11
    - VLAN 2, 225.10.0.13 – 225.10.0.15, Source: 22.10.0.13
    - VLAN 3, 230.0.2.1 – 230.0.2.2, Source: 22.10.0.1
    - VLAN 3, 230.0.2.3 – 230.0.2.5, Source: 22.10.0.3
- The Mrouter sends IGMP Query packets in VLAN 2 and VLAN 3. The Mrouter's IP address is 10.10.10.10.
- The multicast hosts send the following IGMP Reports:
    - IGMPv2 Report, VLAN 2, Group: 225.10.0.11, Source: *
    - IGMPv2 Report, VLAN 3, Group: 230.0.2.1, Source: *
    - IGMPv3 IS_INCLUDE Report, VLAN 2, Group: 225.10.0.13, Source: 22.10.0.13
    - IGMPv3 IS_INCLUDE Report, VLAN 3, Group: 230.0.2.3, Source: 22.10.0.3

- The hosts receive multicast traffic as follows:
  - Host 1 receives multicast traffic for groups (*, 225.10.0.11), (22.10.0.13, 225.10.0.13)
  - Host 2 receives multicast traffic for groups (*, 225.10.0.11), (*, 230.0.2.1), (22.10.0.13, 225.10.0.13), (22.10.0.3, 230.0.2.3)
  - Host 3 receives multicast traffic for groups (*, 230.0.2.1), (22.10.0.3, 230.0.2.3)
- The Mrouter receives all the multicast traffic.

Prerequisites

Before you configure IGMP Snooping, ensure you have performed the following actions:

- Configured VLANs.
- Enabled IGMP.
- Added VLANs to IGMP Snooping.
- Configured a switch or Mrouter as the Querier.
- Identified the IGMP version(s) you want to enable.
- Disabled IGMP flooding.

Configuration

This section provides the configuration details of the switches shown in Figure 34.

Switch A Configuration

1. Configure VLANs and tagging.

```
RS G8124-E(config)# interface port 1-5
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 2,3
RS G8124-E(config-if)# exit
```

2. Configure an IP interface with IPv4 address, and assign a VLAN.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip address 10.10.10.1 enable
RS G8124-E(config-ip-if)# vlan 2
RS G8124-E(config-ip-if)# exit
```

3. Assign a bridge priority lower than the default bridge priority to enable the switch to become the STP root in STG 2 and 3.

```
RS G8124-E(config)# spanning-tree stp 2 bridge priority 4096
RS G8124-E(config)# spanning-tree stp 3 bridge priority 4096
```

4. Configure LACP dynamic LAGs (portchannels).

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# lacp key 100
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# exit

RS G8124-E(config)# interface port 2
RS G8124-E(config-if)# lacp key 100
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# exit

RS G8124-E(config)# interface port 3
RS G8124-E(config-if)# lacp key 200
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# exit

RS G8124-E(config)# interface port 4
RS G8124-E(config-if)# lacp key 200
RS G8124-E(config-if)# lacp mode active
```

Switch B Configuration

1. Configure VLANs and tagging.

```
RS G8124-E(config)# vlan 2,3
RS G8124-E(config-vlan)# interface port 1-4,6
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 2,3
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 5
RS G8124-E(config-if)# switchport access vlan 2
```

2. Configure an IP interface with IPv4 address, and assign a VLAN.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip address 10.10.10.2 enable
RS G8124-E(config-ip-if)# vlan 2
RS G8124-E(config-ip-if)# exit
```

3. Configure STP. Reset the ports to make the edge configuration operational.

```
RS G8124-E(config)# interface port 5,6
RS G8124-E(config-if)# spanning-tree portfast
RS G8124-E(config-if)# shutdown
RS G8124-E(config-if)# no shutdown
RS G8124-E(config-if)# exit
```

4. Configure an LACP dynamic LAG (portchannel).

```
RS G8124-E(config)# interface port 1,2
RS G8124-E(config-if)# lacp key 300
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# exit
```

5. Configure a static LAG (portchannel).

```
RS G8124-E(config)# portchannel 1 port 3,4 enable
```

6. Configure IGMP Snooping.

```
RS G8124-E(config)# ip igmp enable
RS G8124-E(config)# ip igmp snoop vlan 2,3
RS G8124-E(config)# ip igmp snoop source-ip 10.10.10.2
RS G8124-E(config)# ip igmp snoop igmpv3 enable
RS G8124-E(config)# ip igmp snoop igmpv3 sources 64
RS G8124-E(config)# ip igmp snoop enable
RS G8124-E(config)# no ip ipmcfld
```

Switch C Configuration

1. Configure VLANs and tagging.

```
RS G8124-E(config)# vlan 2,3
RS G8124-E(config-vlan)# interface port 1-4,6
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 2,3
RS G8124-E(config-if)# exit
RS G8124-E(config)# interface port 5
RS G8124-E(config-if)# switchport access vlan 3
```

2. Configure an IP interface with IPv4 address, and assign a VLAN.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip address 10.10.10.3 enable
RS G8124-E(config-ip-if)# vlan 2
RS G8124-E(config-ip-if)# exit
```

3. Configure STP. Reset the ports to make the edge configuration operational.

```
RS G8124-E(config)# interface port 5,6
RS G8124-E(config-if)# spanning-tree portfast
RS G8124-E(config-if)# shutdown
RS G8124-E(config-if)# no shutdown
RS G8124-E(config-if)# exit
```

4. Configure an LACP dynamic LAG (portchannel).

```
RS G8124-E(config)# interface port 1,2
RS G8124-E(config-if)# lacp key 400
RS G8124-E(config-if)# lacp mode active
RS G8124-E(config-if)# exit
```

5. Configure a static LAG (portchannel).

```
RS G8124-E(config)# portchannel 1 port 3,4 enable
```

6. Configure IGMP Snooping.

```
RS G8124-E(config)# ip igmp enable
RS G8124-E(config)# ip igmp snoop vlan 2,3
RS G8124-E(config)# ip igmp snoop source-ip 10.10.10.3
RS G8124-E(config)# ip igmp snoop igmpv3 enable
RS G8124-E(config)# ip igmp snoop igmpv3 sources 64
RS G8124-E(config)# ip igmp snoop enable
RS G8124-E(config)# no ip ipmcfld
```

Troubleshooting

This section provides the steps to resolve common IGMP Snooping configuration issues. The topology described in Figure 34 is used as an example.

Multicast traffic from non-member groups reaches the host or Mrouter

● Check if traffic is unregistered. For unregistered traffic, an IGMP entry is not displayed in the IGMP groups table.

```
RS G8124-E# show ip igmp groups
```

● Ensure IPMC flooding is disabled.

```
RS G8124-E(config)# no ip ipmcfld
```

● Check the egress port's VLAN membership. The ports to which the hosts and Mrouter are connected must be used only for VLAN 2 and VLAN 3.

```
RS G8124-E# show vlan
```

Note: To avoid such a scenario, disable IPMC flooding for all VLANs enabled on the switches (if this is an acceptable configuration).

● Check IGMP Reports on switches B and C for information about the IGMP groups.

```
RS G8124-E# show ip igmp groups
```

If the non-member IGMP groups are displayed in the table, close the application that may be sending the IGMP Reports for these groups.

Identify the traffic source by using a sniffer on the hosts and reading the source IP/MAC address. If the source IP/MAC address is unknown, check the port statistics to find the ingress port.

```
RS G8124-E# show interface port <port id> interface-counters
```

● Ensure no static multicast MACs, static multicast groups, or static Mrouters are configured.

● Ensure PIM is not configured.

Not all multicast traffic reaches the appropriate receivers.

● Ensure hosts are sending IGMP Reports for all the groups. Check the VLAN on which the groups are learned.

```
RS G8124-E# show ip igmp groups
```

If some of the groups are not displayed, ensure the multicast application is running on the host device and the generated IGMP Reports are correct.

- Ensure multicast traffic reaches the switch to which the host is connected.

  Close the application sending the IGMP Reports. Clear the IGMP groups by disabling, then re-enabling the port.

  Note: To clear all IGMP groups, use the following command:

  ```
  RS G8124-E(config)# clear ip igmp groups
  ```

  However, this will clear all the IGMP groups and will influence other hosts.

  Check if the multicast traffic reaches the switch.

  ```
  RS G8124-E# show ip igmp ipmcgrp
  ```

  If the multicast traffic group is not displayed in the table, check the link state, VLAN membership, and STP convergence.

- Ensure multicast server is sending all the multicast traffic.

- Ensure no static multicast MACs, static multicast groups, or static multicast routes are configured.

IGMP queries sent by the Mrouter do not reach the host.

- Ensure the Mrouter is learned on switches B and C.

  ```
  RS G8124-E# show ip igmp mrouter
  ```

  If it is not learned on switch B but is learned on switch C, check the link state of the LAG, VLAN membership, and STP convergence.

  If it is not learned on any switch, ensure the multicast application is running and is sending correct IGMP Query packets.

  If it is learned on both switches, check the link state, VLAN membership, and STP port states for the ports connected to the hosts.

IGMP Reports/Leaves sent by the hosts do not reach the Mrouter

- Ensure IGMP Queries sent by the Mrouter reach the hosts.

- Ensure the Mrouter is learned on both switches. Note that the Mrouter may not be learned on switch B immediately after a LAG failover/failback.

  ```
  RS G8124-E# show ip igmp mrouter
  ```

- Ensure the host's multicast application is started and is sending correct IGMP Reports/Leaves.

  ```
  RS G8124-E# show ip igmp groups
  RS G8124-E# show ip igmp counters
  ```

A host receives multicast traffic from the incorrect VLAN

- Check port VLAN membership.

- Check IGMP Reports sent by the host.

- Check multicast data sent by the server.

The Mrouter is learned on the incorrect LAG

- Check link state. LAG 1 might be down or in STP discarding state.

- Check STP convergence.

- Check port VLAN membership.

Hosts receive multicast traffic at a lower rate than normal

Note: This behavior is expected if IPMC flood is disabled. As soon as the IGMP/IPMC entries are installed on ASIC, the IPMC traffic recovers and is forwarded at line rate. This applies to unregistered IPMC traffic.

- Ensure a storm control is not configured on the LAGs.

```
RS G8124-E(config)# interface port <port id>
RS G8124-E(config-if)# no storm-control multicast
```

- Check link speeds and network congestion.

Additional IGMP Features

The following topics are discussed in this section:

FastLeave

In normal IGMP operation, when the switch receives an IGMPv2 Leave message, it sends a Group-Specific Query to determine if any other devices in the same group (and on the same port) are still interested in the specified multicast group traffic. The switch removes the affiliated port from that particular group, if the switch does not receive an IGMP Membership Report within the query-response-interval.

With FastLeave enabled on the VLAN, a port can be removed immediately from the port list of the group entry when the IGMP Leave message is received.

Note: Only IGMPv2 supports FastLeave. Enable FastLeave on ports that have only one host connected. If more than one host is connected to a port, you may lose some hosts unexpectedly.

Use the following command to enable FastLeave.

```
RS G8124-E(config)# ip igmp snoop vlan <VLAN number> fast-leave
```

IGMP Filtering

With IGMP filtering, you can allow or deny certain IGMP groups to be learned on a port.

If access to a multicast group is denied, IGMP *Membership Reports* from the port are dropped, and the port is not allowed to receive IPv4 multicast traffic from that group. If access to the multicast group is allowed, Membership Reports from the port are forwarded for normal processing.

To configure IGMP filtering, you must globally enable IGMP filtering, define an IGMP filter, assign the filter to a port, and enable IGMP filtering on the port. To define an IGMP filter, you must configure a range of IPv4 multicast groups, choose whether the filter will allow or deny multicast traffic for groups within the range, and enable the filter.

Configuring the Range

Each IGMP filter allows you to set a start and end point that defines the range of IPv4 addresses upon which the filter takes action. Each IPv4 address in the range must be between 224.0.0.0 and 239.255.255.255.

Configuring the Action

Each IGMP filter can allow or deny IPv4 multicasts to the range of IPv4 addresses configured. If you configure the filter to deny IPv4 multicasts, then IGMP *Membership Reports* from multicast groups within the range are dropped. You can configure a secondary filter to allow IPv4 multicasts to a small range of addresses within a larger range that a primary filter is configured to deny. The two filters work together to allow IPv4 multicasts to a small subset of addresses within the larger range of addresses.

Note: Lower-numbered filters take precedence over higher-number filters. For example, the action defined for IGMP filter 1 supersedes the action defined for IGMP filter 2.

Configure IGMP Filtering

1. Enable IGMP filtering on the switch.

```
RS G8124-E(config)# ip igmp filtering
```

2. Define an IGMP filter with IPv4 information.

```
RS G8124-E(config)# ip igmp profile 1 range 224.0.0.0 226.0.0.0
RS G8124-E(config)# ip igmp profile 1 action deny
RS G8124-E(config)# ip igmp profile 1 enable
```

3. Assign the IGMP filter to a port.

```
RS G8124-E(config)# interface port 3
RS G8124-E(config-if)# ip igmp profile 1
RS G8124-E(config-if)# ip igmp filtering
```

Static Multicast Router

A static Mrouter can be configured for a particular port on a particular VLAN. A static Mrouter does not have to be learned through IGMP Snooping. Any data port can accept a static Mrouter.

When you configure a static Mrouter on a VLAN, it replaces any dynamic Mrouters learned through IGMP Snooping.

Configure a Static Multicast Router

1. For each Mrouter, configure a port, VLAN, and IGMP version of the multicast router.

```
RS G8124-E(config)# ip igmp mrouter 5 1 2
```

2. Verify the configuration.

```
RS G8124-E(config)# show ip igmp mrouter
```

Chapter 25. Multicast Listener Discovery

Multicast Listener Discovery (MLD) is an IPv6 protocol that a host uses to request multicast data for a multicast group. An IPv6 router uses MLD to discover the presence of multicast listeners (nodes that want to receive multicast packets) on its directly attached links, and to discover specifically the multicast addresses that are of interest to those neighboring nodes.

MLD version 1 is derived from Internet Group Management Protocol version 2 (IGMPv2) and MLDv2 is derived from IGMPv3. MLD uses ICMPv6 (IP Protocol 58) message types. See RFC 2710 and RFC 3810 for details.

MLDv2 protocol, when compared to MLDv1, adds support for source filtering—the ability for a node to report interest in listening to packets only from specific source addresses, or from all but specific source addresses, sent to a particular multicast address. MLDv2 is interoperable with MLDv1. See RFC 3569 for details on Source-Specific Multicast (SSM).

The following topics are discussed in this chapter:

MLD Terms

Following are the commonly used MLD terms:

- Multicast traffic: Flow of data from one source to multiple destinations.

- Group: A multicast stream to which a host can join.

- Multicast Router (Mrouter): A router configured to make routing decisions for multicast traffic. The router identifies the type of packet received (unicast or multicast) and forwards the packet to the intended destination.

- Querier: An Mrouter that sends periodic query messages. Only one Mrouter on the subnet can be elected as the Querier.

- Multicast Listener Query: Messages sent by the Querier. There are three types of queries:
  - General Query: Sent periodically to learn multicast address listeners from an attached link. G8124-E uses these queries to build and refresh the Multicast Address Listener state. General Queries are sent to the link-scope all-nodes multicast address (FF02::1), with a multicast address field of 0, and a maximum response delay of *query response interval*.
  - Multicast Address Specific Query: Sent to learn if a specific multicast address has any listeners on an attached link. The multicast address field is set to the IPv6 multicast address.
  - Multicast Address and Source Specific Query: Sent to learn if, for a specified multicast address, there are nodes still listening to a specific set of sources. Supported only in MLDv2.

Note: Multicast Address Specific Queries and Multicast Address and Source Specific Queries are sent only in response to State Change Reports, and never in response to Current State Reports.

- Multicast Listener Report: Sent by a host when it joins a multicast group, or in response to a Multicast Listener Query sent by the Querier. Hosts use these reports to indicate their current multicast listening state, or changes in the multicast listening state of their interfaces. These reports are of two types:
  - Current State Report: Contains the current Multicast Address Listening State of the host.
  - State Change Report: If the listening state of a host changes, the host immediately reports these changes through a State Change Report message. These reports contain either Filter Mode Change records and/or Source List Change records. State Change Reports are retransmitted several times to ensure all Mrouters receive it.

- Multicast Listener Done: Sent by a host when it wants to leave a multicast group. This message is sent to the link-scope all-routers IPv6 destination address of FF02::2. When an Mrouter receives a Multicast Listener Done message from the last member of the multicast address on a link, it stops forwarding traffic to this multicast address.

How MLD Works

The software uses the information obtained through MLD to maintain a list of multicast group memberships for each interface and forwards the multicast traffic only to interested listeners.

Without MLD, the switch forwards IPv6 multicast traffic through all ports, increasing network load. Following is an overview of operations when MLD is configured on the G8124-E:

- The switch acts as an Mrouter when MLDv1/v2 is configured and enabled on each of its directly attached links. If the switch has multiple interfaces connected to the same link, it operates the protocol on any one of the interfaces.

- If there are multiple Mrouters on the subnet, the Mrouter with the numerically lowest IPv6 address is elected as the Querier.

- The Querier sends general queries at short intervals to learn multicast address listener information from an attached link.

- Hosts respond to these queries by reporting their per-interface Multicast Address Listening state, through Current State Report messages sent to a specific multicast address that all MLD routers on the link listen to.

- If the listening state of a host changes, the host immediately reports these changes through a State Change Report message.

- The Querier sends a Multicast Address Specific Query to verify if hosts are listening to a specified multicast address or not. Similarly, if MLDv2 is configured, the Querier sends a Multicast Address and Source Specific Query to verify, for a specified multicast address, if hosts are listening to a specific set of sources, or not. MLDv2 listener report messages consists of Multicast Address Records:

  - INCLUDE: to receive packets from source specified in the MLDv2 message
  - EXCLUDE: to receive packets from all sources except the ones specified in the MLDv2 message

- A host can send a State Change Report to indicate its desire to stop listening to a particular multicast address (or source in MLDv2). The Querier then sends a multicast address specific query to verify if there are other listeners of the multicast address. If there aren't any, the Mrouter deletes the multicast address from its Multicast Address Listener state and stops sending multicast traffic. Similarly in MLDv2, the Mrouter sends a Multicast Address and Source Specific Query to verify if, for a specified multicast address, there are hosts still listening to a specific set of sources.

G8124-E supports MLD versions 1 and 2.

Note: MLDv2 operates in version 1 compatibility mode when, in a specific network, not all hosts are configured with MLDv2.

How Flooding Impacts MLD

By default, the flood option is enabled to allow hardware flooding in VLAN for all unknown IP multicast (IPMC) traffic. When the flood option is disabled, unknown IPMC is sent only to the Mrouter ports.

Enter the following command to set the flood option:

```
RS G8124-E(config)# [no] ip ipmcfld
```

MLD Querier

An Mrouter acts as a Querier and periodically (at short query intervals) sends query messages in the subnet. If there are multiple Mrouters in the subnet, only one can be the Querier. All Mrouters on the subnet listen to the messages sent by the multicast address listeners, and maintain the same multicast listening information state.

All MLDv2 queries are sent with the FE80::/64 link-local source address prefix.

Querier Election

Only one Mrouter can be the Querier per subnet. All other Mrouters will be non-Queriers. MLD versions 1 and 2 elect the Mrouter with the numerically lowest IPv6 address as the Querier.

If the switch is configured as an Mrouter on a subnet, it also acts as a Querier by default and sends multiple general queries. If the switch receives a general query from another Querier with a numerically lower IPv6 address, it sets the *other querier present timer* to the *other querier present timeout*, and changes its state to non-Querier. When the *other querier present timer* expires, it regains the Querier state and starts sending general queries.

Note: When MLD Querier is enabled on a VLAN, the switch performs the role of an MLD Querier only if it meets the MLD Querier election criteria.

Dynamic Mrouters

The switch learns Mrouters on the ingress VLANs of the MLD-enabled interface. All report or done messages are forwarded to these Mrouters. By default, the option of dynamically learning Mrouters is disabled. To enable it, use the following command:

```
RS G8124-E(config)# interface ip <interface number>
RS G8124-E(config-ip-if)# ipv6 mld dmrtr enable
```

MLD Capacity and Default Values

Table 29 lists the maximum and minimum values of the G8124-E variables.

Table 29. *G8124-E Capacity Table*

Variable	Maximum Value
IPv6 Multicast Entries	256
IPv6 Interfaces for MLD	8

Table 30 lists the default settings for MLD features and variables.

Table 30. *MLD Timers and Default Values*

Field	Default Value
Robustness Variable (RV)	2
Query Interval (QI)	125 seconds
Query Response Interval (QRI)	10 seconds
Multicast Address Listeners Interval (MALI)	260 seconds [derived: RV*QI+QRI]
Other Querier Present Interval [OQPT]	255 seconds [derived: RV*QI + ½ QRI]
Start up Query Interval [SQI]	31.25 seconds [derived: ¼ * QI]
Startup Query Count [SQC]	2 [derived: RV]
Last Listener Query Interval [LLQI]	1 second
Last Listener Query Count [LLQC]	2 [derived: RV]
Last Listener Query Time [LLQT]	2 seconds [derived: LLQI * LLQT]
Older Version Querier Present Timeout: [OVQPT]	260 seconds [derived: RV*QI+ QRI]
Older Version Host Present Interval [OVHPT]	260 seconds [derived: RV* QI+QRI]

Configuring MLD

Following are the steps to enable MLD and configure the interface parameters:

1. Enable IPv6 boot profile.

```
RS G8124-E(config)# boot profile ipv6
RS G8124-E(config)# exit
RS G8124-E# reload
```

2. Turn on MLD globally.

```
RS G8124-E(config)# ipv6 mld
RS G8124-E(config-router-mld)# enable
RS G8124-E(config-router-mld)# exit
```

3. Create an IPv6 interface.

```
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# ipv6 address 2002:1:0:0:0:0:0:3
RS G8124-E(config-ip-if)# ipv6 prefixlen 64
```

4. Enable MLD on the IPv6 interface.

```
RS G8124-E(config-ip-if)# ipv6 mld enable
```

5. Configure the MLD parameters on the interface: version, robustness, query response interval, MLD query interval, and last listener query interval.

```
RS G8124-E(config-ip-if)# ipv6 mld version <1-2>(MLD version)
RS G8124-E(config-ip-if)# ipv6 mld robust <1-10>(Robustness)
RS G8124-E(config-ip-if)# ipv6 mld qri <1-256>(In seconds)
RS G8124-E(config-ip-if)# ipv6 mld qintrval <1-608>(In seconds)
RS G8124-E(config-ip-if)# ipv6 mld llistnr <1-32>(In seconds)
```

Chapter 26. Border Gateway Protocol

Border Gateway Protocol (BGP) is an Internet protocol that enables routers on an IPv4 network to share and advertise routing information with each other about the segments of the IPv4 address space they can access within their network and with routers on external networks. BGP allows you to decide what is the "best" route for a packet to take from your network to a destination on another network rather than simply setting a default route from your border router(s) to your upstream provider(s). BGP is defined in RFC 1771.

RackSwitch G8124-Ees can advertise their IP interfaces and IPv4 addresses using BGP and take BGP feeds from as many as 96 BGP router peers. This allows more resilience and flexibility in balancing traffic from the Internet.

Note: Lenovo Networking OS 8.3 does not support IPv6 for BGP.

The following topics are discussed in this section:

Internal Routing Versus External Routing

To ensure effective processing of network traffic, every router on your network needs to know how to send a packet (directly or indirectly) to any other location/destination in your network. This is referred to as *internal routing* and can be done with static routes or using active, internal dynamic routing protocols, such as RIP, RIPv2, and OSPF.

Static routes must have a higher degree of precedence than dynamic routing protocols. If the destination route is not in the route cache, the packets are forwarded to the default gateway which may be incorrect if a dynamic routing protocol is enabled.

It is also useful to tell routers outside your network (upstream providers or *peers*) about the routes you can access in your network. External networks (those outside your own) that are under the same administrative control are referred to as *autonomous systems* (AS). Sharing of routing information between autonomous systems is known as *external routing*.

External BGP (eBGP) is used to exchange routes between different autonomous systems whereas internal BGP (iBGP) is used to exchange routes within the same autonomous system. An iBGP is a type of internal routing protocol you can use to do active routing inside your network. It also carries AS path information, which is important when you are an ISP or doing BGP transit.

The iBGP peers have to maintain reciprocal sessions to every other iBGP router in the same AS (in a full-mesh manner) to propagate route information throughout the AS. If the iBGP session shown between the two routers in AS 20 was not present (as indicated in Figure 35), the top router would not learn the route to AS 50, and the bottom router would not learn the route to AS 11, even though the two AS 20 routers are connected via the RackSwitch G8124-E.

Figure 35. iBGP and eBGP
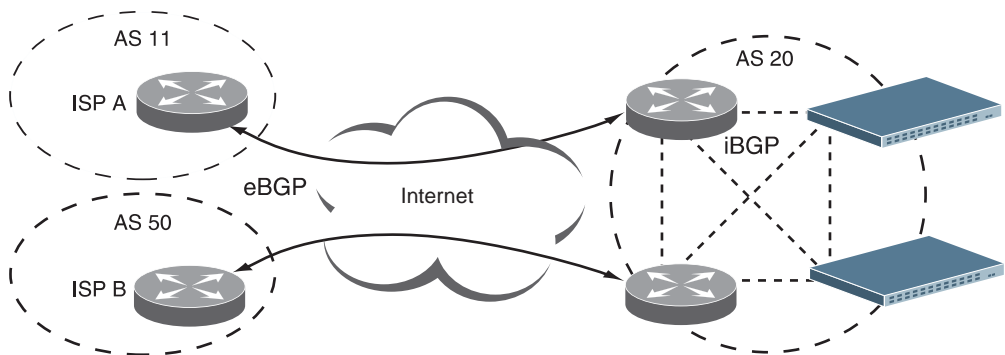
When there are many iBGP peers, having a full-mesh configuration results in large number of sessions between the iBGP peers. In such situations, configuring a route reflector eliminates the full-mesh configuration requirement, prevents route propagation loops, and provides better scalability to the peers. For details, see "Route Reflector" on page 357.

Typically, an AS has one or more *border routers*—peer routers that exchange routes with other ASs—and an internal routing scheme that enables routers in that AS to reach every other router and destination within that AS. When you *advertise* routes to border routers on other autonomous systems, you are effectively committing to carry data to the IPv4 space represented in the route being advertised. For example, if you advertise 192.204.4.0/24, you are declaring that if another router sends you data destined for any address in 192.204.4.0/24, you know how to carry that data to its destination.

Route Reflector

The Lenovo N/OS implementation conforms to the BGP Route Reflection specification defined in RFC 4456.

As per RFC 1771 specification, a route received from an iBGP peer cannot be advertised to another iBGP peer. This makes it mandatory to have full-mesh iBGP sessions between all BGP routers within an AS. A route reflector—a BGP router—breaks this iBGP loop avoidance rule. It does not affect the eBGP behavior. A route reflector is a BGP speaker that advertises a route learnt from an iBGP peer to another iBGP peer. The advertised route is called the reflected route.

A route reflector has two groups of internal peers: clients and non-clients. A route reflector reflects between these groups and among the clients. The non-client peers must be fully meshed. The route reflector and its clients form a cluster.

When a route reflector receives a route from an iBGP peer, it selects the best path based on its path selection rule. It then does the following based on the type of peer it received the best path from:

● A route received from a non-client iBGP peer is reflected to all clients.

● A route received from an iBGP client peer is reflected to all iBGP clients and iBGP non-clients.

In Figure 36, the G8124-E is configured as a route reflector. All clients and non-clients are in the same AS.

Figure 36. iBGP Route Reflector
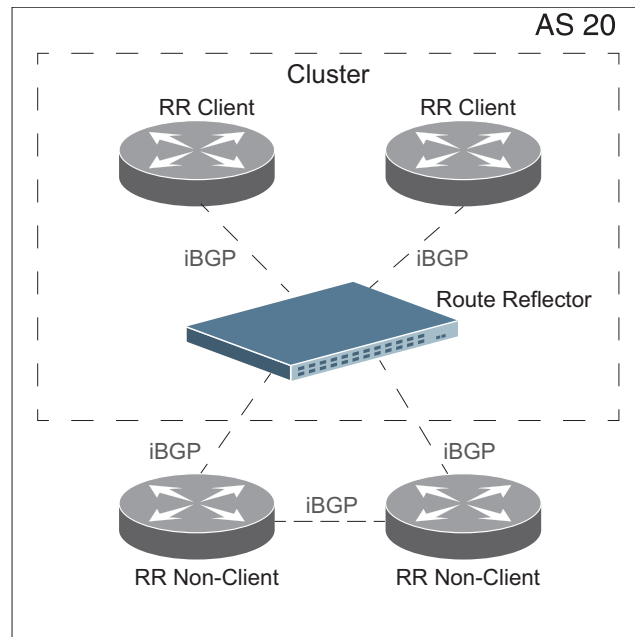
The following attributes are used by the route reflector functionality:

- ORIGINATOR ID: BGP identifier (BGP router ID) of the route originator in the local AS. If the route does not have the ORIGINATOR ID attribute (it has not been reflected before), the router ID of the iBGP peer from which the route has been received is copied into the Originator ID attribute.This attribute is never modified by subsequent route reflectors. A router that identifies its own ID as the ORIGINATOR ID, it ignores the route.

- CLUSTER LIST: Sequence of the CLUSTER ID (the router ID) values representing the reflection path that the route has passed. The value configured with the `cluster-id` command (or the router ID of the route reflector if the `cluster-id` is not configured) is prepended to the Cluster list attribute. If a route reflector detects its own CLUSTER ID in the CLUSTER LIST, it ignores the route. Up to 10 CLUSTER IDs can be added to a CLUSTER LIST.

Route reflection functionality can be configured as follows:

1. Configure an AS.

```
RS G8124-E(config)# router bgp
RS G8124-E(config-router-bgp)# as 22
RS G8124-E(config-router-bgp)# enable
```

2. Configure a route reflector client.

```
RS G8124-E(config-router-bgp)# neighbor 2 remote-address 10.1.50.1
RS G8124-E(config-router-bgp)# neighbor 2 remote-as 22
RS G8124-E(config-router-bgp)# neighbor 2 route-reflector-client
RS G8124-E(config-router-bgp)# no neighbor 2 shutdown
```

Note: When a client is configured on the G8124-E, the switch automatically gets configured as a route reflector.

3. Verify configuration.

```
RS G8124-E(config)# show ip bgp neighbor 2 information

BGP Peer 2 Information:
  2: 10.1.50.1          , version 0, TTL 255, TTL Security hops 0
    Remote AS: 0, Local AS: 22, Link type: IBGP
    Remote router ID: 0.0.0.0,    Local router ID: 9.9.9.9
    next-hop-self disabled
    RR client enabled
    BGP status: connect, Old status: connect
    Total received packets: 0, Total sent packets: 0
    Received updates: 0, Sent updates: 0
    Keepalive: 0, Holdtime: 0, MinAdvTime: 60
    LastErrorCode: unknown(0), LastErrorSubcode: unspecified(0)
    Established state transitions: 0
```

Once configured as a route reflector, the switch, by default, passes routes between clients. If required, you can disable this by using the following command:

```
RS G8124-E(config-router-bgp)# no client-to-client reflection
```

You can view the route reflector BGP attributes attached to a BGP route using the following command:

```
RS G8124-E(config-router-bgp)# show ip bgp information 5.0.0.0
255.255.255.0
BGP routing table entry for 5.0.0.0/255.255.255.0
Paths: (1 available, best #1)
Multipath: eBGP
 Local
        30.1.1.1 (metric 0) from 22.22.1.1(17.17.17.17)
        Origin: IGP, localpref 0, valid, internal, best
        Originator: 1.16.0.195
         Cluster list: 17.17.17.17
```

You can view BGP advertised routes to a specific neighbor or to all neighbors using the command:

```
[Prompt](config)# show ip bgp neighbor advertised-routes
```

Restrictions

Consider the following restrictions when configuring route reflection functionality:

- When a CLUSTER ID is changed, all iBGP sessions are restarted.
- When a route reflector client is enabled/disabled, the session is restarted.

Forming BGP Peer Routers

Two BGP routers become peers or neighbors once you establish a TCP connection between them. You can configure BGP peers statically or dynamically. While it may be desirable to configure static peers for security reasons, dynamic peers prove to be useful in cases where the remote address of the peer is unknown. For example in B-RAS applications, where subscriber interfaces are dynamically created and the address is assigned dynamically from a local pool or by using RADIUS.

For each new route, if a peer is interested in that route (for example, if a peer would like to receive your static routes and the new route is static), an update message is sent to that peer containing the new route. For each route removed from the route table, if the route has already been sent to a peer, an update message containing the route to withdraw is sent to that peer.

For each Internet host, you must be able to send a packet to that host, and that host has to have a path back to you. This means that whoever provides Internet connectivity to that host must have a path to you. Ultimately, this means that they must "hear a route" which covers the section of the IPv4 space you are using; otherwise, you will not have connectivity to the host in question.

Static Peers

You can configure BGP static peers by using the following commands:

```
RS G8124-E(config)# router bgp
RS G8124-E(config-router-bgp)# neighbor <1-96> remote-address <IP address>
RS G8124-E(config-router-bgp)# neighbor <1-96> remote-as <1-65535>
RS G8124-E(config-router-bgp)# no neighbor <1-96> shutdown
```

Static peers always take precedence over dynamic peers. Consider the following:

- If the remote address of an incoming BGP connection matches both a static peer address and an IP address from a dynamic group, the peer is configured statically and not dynamically.

- If a new static peer is enabled while a dynamic peer for the same remote address exists, BGP automatically removes the dynamic peer.

- If a new static peer is enabled when the maximum number of BGP peers were already configured, then BGP deletes the dynamic peer that was last created and adds the newly created static peer. A syslog will be generated for the peer that was deleted.

Dynamic Peers

To configure dynamic peers, you must define a range of IP addresses for a group. BGP waits to receive an open message initiated from BGP speakers within that range. Dynamic peers are automatically created when a peer group member accepts the incoming BGP connection. Dynamic peers are passive. When they are not in the established state, they accept inbound connections but do not initiate outbound connections.

You can configure up to 6 AS numbers per group. When the BGP speaker receives an open message from a dynamic peer, the AS number from the packet must match one of the remote AS numbers configured on the corresponding group.

When you delete a remote AS number, all dynamic peers established from that remote AS will be deleted.

You can define attributes for the dynamic peers only at the group level. You cannot configure attributes for any one dynamic peer. All static peer attributes, except the BGP passive mode, can also be configured for groups.

To set the maximum number of dynamic peers for a group that can simultaneously be in an established state, enter the following command:

```
RS G8124-E(config-router-bgp)# neighbor group <1-8> listen limit <1-96>
```

If you reset this limit to a lower number, and if the dynamic peers already established for the group are higher than this new limit, then BGP deletes the last created dynamic peer(s) until the new limit is reached.

Note: The maximum number of static and dynamic peers established simultaneously cannot exceed the maximum peers, i.e. 96, that the switch can support. If the maximum peers are established, no more dynamic peers will be enabled even if the maximum dynamic peers limit you had configured for the groups was not reached.

Given below are the basic commands for configuring dynamic peers:

```
RS G8124-E(config)# router bgp
RS G8124-E(config-router-bgp)# neighbor group <1-8> listen range
<IP address> <subnet mask>                                        (Define
IP address range)
RS G8124-E(config-router-bgp)# neighbor group <1-8> remote-as <1-65535>
alternate-as <1-65535>                                           (Enter up to 5
alternate AS numbers)
RS G8124-E(config-router-bgp)# no neighbor group <1-96> shutdown
```

Removing Dynamic Peers

You cannot remove dynamic peers manually. However, you can stop a dynamic peer using the following command:

```
RS G8124-E(config)# router bgp stop <neighbor number>
```

The stop command interrupts the BGP connection until the peer tries to re-establish the connection.

Also, when a dynamic peer state changes from established to idle, BGP removes the dynamic peer.

Loopback Interfaces

In many networks, multiple connections may exist between network devices. In such environments, it may be useful to employ a loopback interface for a common BGP router address, rather than peering the switch to each individual interface.

Note: To ensure that the loopback interface is reachable from peer devices, it must be advertised using an interior routing protocol (such as OSPF), or a static route must be configured on the peer.

To configure an existing loopback interface for BGP neighbor, use the following commands:

```
RS G8124-E(config)# router bgp
RS G8124-E(config-router-bgp)# neighbor <#> update-source loopback <1-5>
RS G8124-E(config-router-bgp)# exit
```

What is a Route Map?

A route map is used to control and modify routing information. Route maps define conditions for redistributing routes from one routing protocol to another or controlling routing information when injecting it in and out of BGP. For example, a route map is used to set a preference value for a specific route from a peer router and another preference value for all other routes learned via the same peer router. For example, the following command is used to enter the Route Map mode for defining a route map:

```
RS G8124-E(config)# route-map <map number>(Select a route map)
```

A route map allows you to match attributes, such as metric, network address, and AS number. It also allows users to overwrite the local preference metric and to append the AS number in the AS route. See "BGP Failover Configuration" on page 372.

Lenovo N/OS allows you to configure 64 route maps. Each route map can have up to eight access lists. Each access list consists of a network filter. A network filter defines an IPv4 address and subnet mask of the network that you want to include in the filter. Figure 37 illustrates the relationship between route maps, access lists, and network filters.

Figure 37. Distributing Network Filters in Access Lists and Route Maps
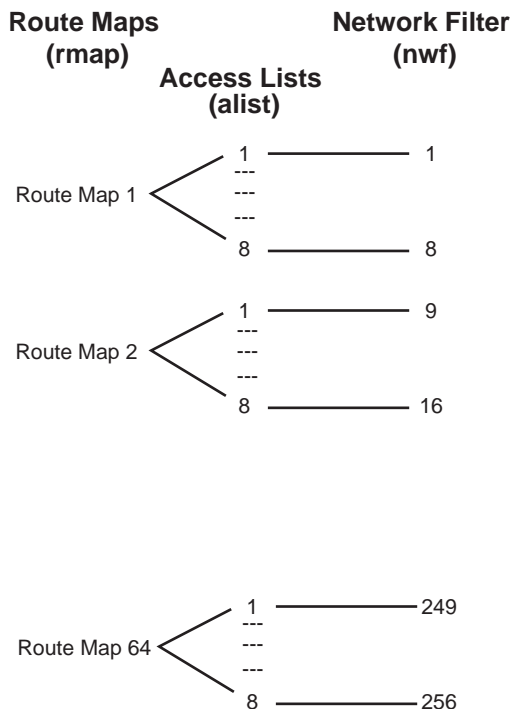
Next Hop Peer IP Address

Next hop peer IP address can be configured only for route maps used in BGP. When a route map is applied on ingress, the next hop of learnt routes is replaced with peer IP address. When applied on egress, the next hop of the redistributed routes is replaced with the local IP address.

```
RS G8124-E(config)# route-map <map number>
RS G8124-E(config-router-map)# set ip next-hop <peer IP address>
```

Incoming and Outgoing Route Maps

You can have two types of route maps: incoming and outgoing. A BGP peer router can be configured to support up to eight route maps in the incoming route map list and outgoing route map list.

If a route map is not configured in the incoming route map list, the router imports all BGP updates. If a route map is configured in the incoming route map list, the router ignores all unmatched incoming updates. If you set the action to **deny**, you must add another route map to permit all unmatched updates.

Route maps in an outgoing route map list behave similar to route maps in an incoming route map list. If a route map is not configured in the outgoing route map list, all routes are advertised or permitted. If a route map in the outgoing route map list is set to **permit**, matched routes are advertised and unmatched routes are ignored.

Precedence

You can set a priority to a route map by specifying a precedence value with the following command (Route Map mode):

```
RS G8124-E(config)# route-map <map number>(Select a route map)
RS G8124-E(config-route-map)# precedence <1-255>(Specify a precedence)
RS G8124-E(config-route-map)# exit
```

The smaller the value the higher the precedence. If two route maps have the same precedence value, the smaller number has higher precedence.

Configuration Overview

To configure route maps, you need to do the following:

1. Define a network filter.

```
RS G8124-E(config)# ip match-address 1 <IPv4 address> <IPv4 subnet mask>
RS G8124-E(config)# ip match-address 1 enable
```

Enter a filter number from 1 to 256. Specify the IPv4 address and subnet mask of the network that you want to match. Enable the network filter. You can distribute up to 256 network filters among 64 route maps each containing eightaccess lists.

Steps 2 and 3 are optional, depending on the criteria that you want to match. In Step 2, the network filter number is used to match the subnets defined in the network filter. In Step 3, the autonomous system number is used to match the subnets. Or, you can use both (Step 2 and Step 3) criteria: access list (network filter) and access path (AS filter) to configure the route maps.

2. (Optional) Define the criteria for the access list and enable it.

Specify the access list and associate the network filter number configured in Step 1.

```
RS G8124-E(config)# route-map 1
RS G8124-E(config-route-map)# access-list 1 match-address 1
RS G8124-E(config-route-map)# access-list 1 metric <metric value>
RS G8124-E(config-route-map)# access-list 1 action deny
RS G8124-E(config-route-map)# access-list 1 enable
```

3. (Optional) Configure the AS filter attributes.

```
RS G8124-E(config-route-map)# as-path-list 1 as 1
RS G8124-E(config-route-map)# as-path-list 1 action deny
RS G8124-E(config-route-map)# as-path-list 1 enable
```

4. Set up the BGP attributes.

If you want to overwrite the attributes that the peer router is sending, define the following BGP attributes:

- Specify up to 32 AS numbers that you want to prepend to a matched route. Use one space between each of the entries in the list.

- Specify the local preference for the matched route.

- Specify the Multi Exit Discriminator (MED) metric for the matched route.

```
RS G8124-E(config-route-map)# as-path-preference <AS number list (up to 32 entries)>
RS G8124-E(config-route-map)# local-preference <local preference value>
RS G8124-E(config-route-map)# metric <metric value>
```

5. Enable the route map.

```
RS G8124-E(config-route-map)# enable
RS G8124-E(config-route-map)# exit
```

6. Turn BGP on.

```
RS G8124-E(config)# router bgp
RS G8124-E(config-router-bgp)# enable
```

7. Assign the route map to a peer router.

    Select the peer router and then add the route map to the incoming route map list,

```
RS G8124-E(config-router-bgp)# neighbor 1 route-map in <1-64>
```

    or to the outgoing route map list.

```
RS G8124-E(config-router-bgp)# neighbor 1 route-map out <1-64>
```

8. Exit Router BGP mode.

```
RS G8124-E(config-router-bgp)# exit
```

Aggregating Routes

Aggregation is the process of combining several different routes in such a way that a single route can be advertised, which minimizes the size of the routing table. You can configure aggregate routes in BGP either by redistributing an aggregate route into BGP or by creating an aggregate entry in the BGP routing table.

To define an aggregate route in the BGP routing table, use the following commands:

```
RS G8124-E(config)# router bgp
RS G8124-E(config-router-bgp)# aggregate-address <1-16> <IPv4 address> <mask>
RS G8124-E(config-router-bgp)# aggregate-address <1-16> enable
```

An example of creating a BGP aggregate route is shown in "Default Redistribution and Route Aggregation Example" on page 374.

Redistributing Routes

In addition to running multiple routing protocols simultaneously, N/OS software can redistribute information from one routing protocol to another. For example, you can instruct the switch to use BGP to re-advertise static routes. This applies to all of the IP-based routing protocols.

You can also conditionally control the redistribution of routes between routing domains by defining a method known as route maps between the two domains. For more information on route maps, see "What is a Route Map?" on page 363. Redistributing routes is another way of providing policy control over whether to export OSPF routes, fixed routes, and static routes. For an example configuration, see "Default Redistribution and Route Aggregation Example" on page 374.

Default routes can be configured using the following methods:

- Import
- Originate—The router sends a default route to peers if it does not have any default routes in its routing table.
- Redistribute—Default routes are either configured through the default gateway or learned via other protocols and redistributed to peer routers. If the default routes are from the default gateway, enable the static routes because default routes from the default gateway are static routes. Similarly, if the routes are learned from another routing protocol, make sure you enable that protocol for redistribution.
- None

BGP Communities

BGP communities are attribute tags that allow controlled distribution of routing information based on an agreement between BGP peers. Communities are commonly used by transit service providers to enable peering customers to choose specific routing destinations for their outgoing routes. The transit service provider would typically publish a list of well-known or proprietary communities along with their descriptions, and take it upon itself to advertise incoming routes accordingly. For instance, an ISP may advertise that incoming routes tagged with community XY:01 will be advertised only to European peers while incoming routes tagged with community XY:02 will be advertised only to Asian peers.

The RackSwitch G8124-E can be configured to manage the community tags applied to the outgoing route updates. It does not, however, modify any routing decisions based on the community tags.

Up to 32 community tags can be applied to prefixes that pass a route-map. Valid values are between 0:0 and 65535:65535. Newly added communities will be appended to any existing configured communities. To append communities to prefixes that pass the route-map, use the following commands:

```
RS G8124-E(config)# route-map <map number>
RS G8124-E(config-route-map)# set community {aa:nn [aa:nn]}
```

To remove all community tags from prefixes that pass the route-map, use the following command:

```
RS G8124-E(config-route-map)# set community none
```

To remove configured communities on a routemap, use the following command:

```
RS G8124-E(config-route-map)# no set community
```

Prefixes with communities are propagated unchanged if there is no routemap that alters the community attribute and if the neighbor has community tags forwarding enabled.

To enable or disable community tags forwarding for specific neighbors or neighbor groups, use the following commands:

```
RS G8124-E(config)# router bgp
RS G8124-E(config-router-bgp)# neighbor 5 send-community
RS G8124-E(config-router-bgp)# no neighbor 6 send-community
RS G8124-E(config-router-bgp)# neighbor group 1 send-community
RS G8124-E(config-router-bgp)# no neighbor group 2 send-community
```

BGP Attributes

The following BGP attributes are discussed in this section: Local preference, metric (Multi-Exit Discriminator), and Next hop.

Local Preference Attribute

When there are multiple paths to the same destination, the local preference attribute indicates the preferred path. The path with the higher preference is preferred (the default value of the local preference attribute is 100). Unlike the weight attribute, which is only relevant to the local router, the local preference attribute is part of the routing update and is exchanged among routers in the same AS.

The local preference attribute can be set in one of two ways:

- The following commands use the BGP default local preference method, affecting the outbound direction only.

```
RS G8124-E(config)# router bgp
RS G8124-E(config_router_bgp)# local-preference
RS G8124-E(config_router_bgp)# exit
```

- The following commands use the route map local preference method, which affects both inbound and outbound directions.

```
RS G8124-E(config)# route-map 1
RS G8124-E(config_route_map)# local-preference
RS G8124-E(config_router_map)# exit
```

Metric (Multi-Exit Discriminator) Attribute

This attribute is a hint to external neighbors about the preferred path into an AS when there are multiple entry points. A lower metric value is preferred over a higher metric value. The default value of the metric attribute is 0.

Unlike local preference, the metric attribute is exchanged between ASs; however, a metric attribute that comes into an AS does not leave the AS.

When an update enters the AS with a certain metric value, that value is used for decision making within the AS. When BGP sends that update to another AS, the metric is reset to 0.

Unless otherwise specified, the router compares metric attributes for paths from external neighbors that are in the same AS.

Next Hop Attribute

BGP routing updates sent to a neighbor contain the next hop IP address used to reach a destination. In eBGP, the edge router, by default, sends its own IP address as the next hop address. However, this can sometimes cause routing path failures in Non-Broadcast Multiaccess Networks (NBMA) and when the edge router sends iBGP updates.

To avoid routing failures, you can manually configure the next hop IP address. In case of NBMA networks, you can configure the external BGP speaker to advertise its own IP address as the next hop. In case of iBGP updates, you can configure the edge iBGP router to send its IP address as the next hop.

Next hop can be configured on a BGP peer or a peer group. Use the following commands:

- Next Hop for a BGP Peer

```
RS G8124-E(config)# router bgp
RS G8124-E(config-router-bgp)# neighbor <number> next-hop-self
```

- Next Hop for a BGP Peer Group:

```
RS G8124-E(config)# router bgp
RS G8124-E(config-router-bgp)# neighbor group <number> next-hop-self
```

Selecting Route Paths in BGP

BGP selects only one path as the best path. It does not rely on metric attributes to determine the best path. When the same network is learned via more than one BGP peer, BGP uses its policy for selecting the best route to that network. The BGP implementation on the G8124-E uses the following criteria to select a path when the same route is received from multiple peers.

1. Local fixed and static routes are preferred over learned routes.

2. With iBGP peers, routes with higher local preference values are selected.

3. In the case of multiple routes of equal preference, the route with lower AS path weight is selected.

   AS path weight = 128 x AS path length (number of autonomous systems traversed).

4. In the case of equal weight and routes learned from peers that reside in the same AS, the lower metric is selected.

   Note: A route with a metric is preferred over a route without a metric.

5. The lower cost to the next hop of routes is selected.

6. In the case of equal cost, the eBGP route is preferred over iBGP.

7. If all routes have same route type (eBGP or iBGP), the route with the lower router ID is selected.

When the path is selected, BGP puts the selected path in its routing table and propagates the path to its neighbors.

Multipath Relax

BGP multipath relax functionality allows load balancing across routes with different autonomous system paths, but equal in length (same as-path length). With this option disabled, both autonomous system paths and as-path length must be identical for load sharing. This functionality can be enabled using the command:

```
RS G8124-E(config-router-bgp)# bestpath as-path multipath-relax
```

BGP Failover Configuration

Use the following example to create redundant default gateways for a G8124-E at a Web Host/ISP site, eliminating the possibility, if one gateway goes down, that requests will be forwarded to an upstream router unknown to the switch.

As shown in Figure 38, the switch is connected to ISP 1 and ISP 2. The customer negotiates with both ISPs to allow the switch to use their peer routers as default gateways. The ISP peer routers will then need to announce themselves as default gateways to the G8124-E.

Figure 38. BGP Failover Configuration Example
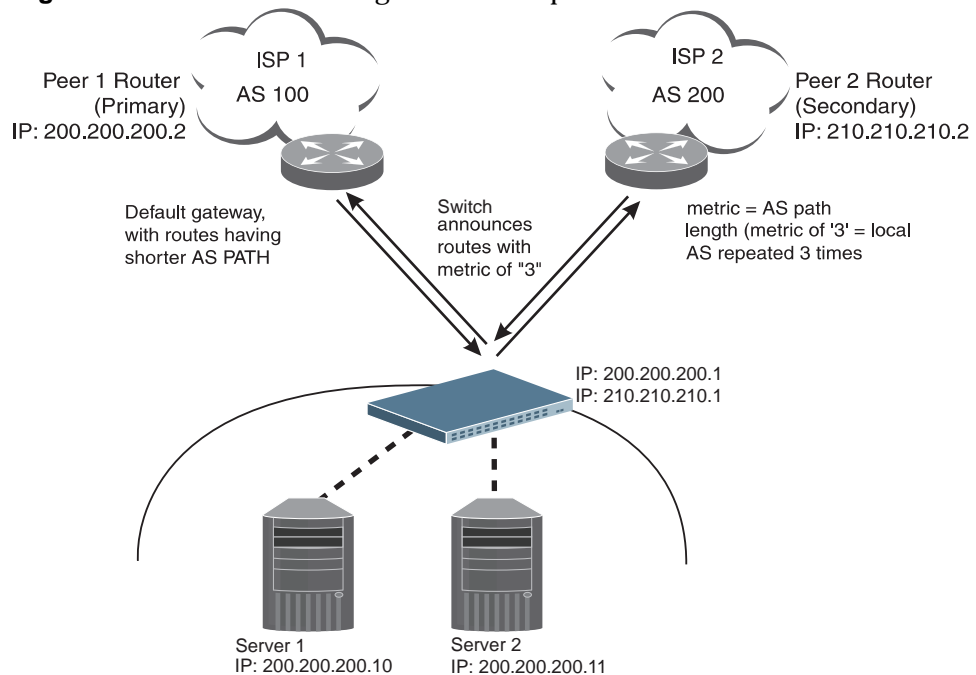
On the G8124-E, one peer router (the secondary one) is configured with a longer AS path than the other, so that the peer with the shorter AS path will be seen by the switch as the primary default gateway. ISP 2, the secondary peer, is configured with a metric of "3," thereby appearing to the switch to be three router *hops* away.

1. Define the VLANs.

   For simplicity, both default gateways are configured in the same VLAN in this example. The gateways could be in the same VLAN or different VLANs.

   ```
   RS G8124-E(config)# vlan 1
   ```

2. Define the IP interfaces with IPv4 addresses.

   The switch will need an IP interface for each default gateway to which it will be connected. Each interface must be placed in the appropriate VLAN. These interfaces will be used as the primary and secondary default gateways for the switch.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip address 200.200.200.1
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip address 210.210.210.1
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

3. Enable IP forwarding.

   IP forwarding is turned on by default and is used for VLAN-to-VLAN (non-BGP) routing. Make sure IP forwarding is on if the default gateways are on different subnets or if the switch is connected to different subnets and those subnets need to communicate through the switch (which they almost always do).

```
RS G8124-E(config)# ip routing
```

   Note: To help eliminate the possibility for a Denial of Service (DoS) attack, the forwarding of directed broadcasts is disabled by default.

4. Configure BGP peer router 1 and 2 with IPv4 addresses.

```
RS G8124-E(config)# router bgp
RS G8124-ERS G8124-E(config-router-bgp)# neighbor 1 remote-address
200.200.200.2
RS G8124-E(config-router-bgp)# neighbor 1 remote-as 100
RS G8124-E(config-router-bgp)# no neighbor 1 shutdown
RS G8124-ERS G8124-E(config-router-bgp)# neighbor 2 remote-address
210.210.210.2
RS G8124-E(config-router-bgp)# neighbor 2 remote-as 200
RS G8124-E(config-router-bgp)# no neighbor 2 shutdown
```

Default Redistribution and Route Aggregation Example

This example shows you how to configure the switch to redistribute information from one routing protocol to another and create an aggregate route entry in the BGP routing table to minimize the size of the routing table.

As illustrated in Figure 39, you have two peer routers: an internal and an external peer router. Configure the G8124-E to redistribute the default routes from AS 200 to AS 135. At the same time, configure for route aggregation to allow you to condense the number of routes traversing from AS 135 to AS 200.

Figure 39. Route Aggregation and Default Route Redistribution
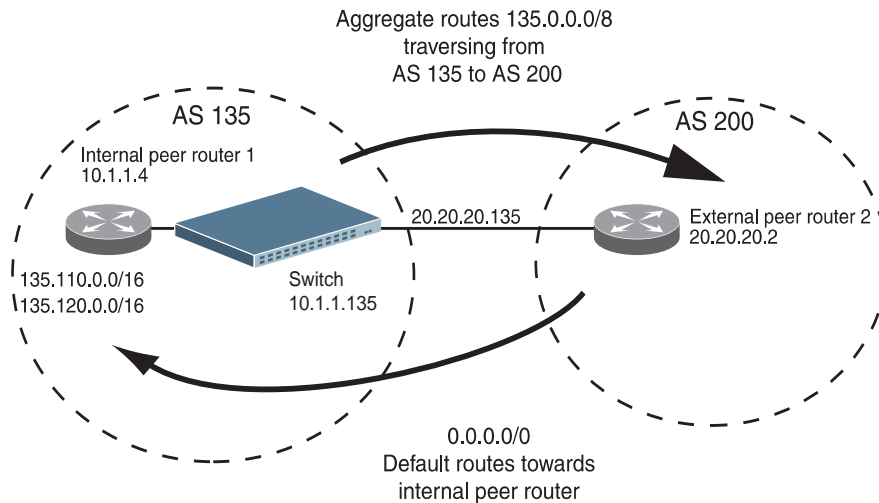
1. Configure the IP interface.

2. Configure the AS number (AS 135) and router ID (10.1.1.135).

```
RS G8124-E(config)# router bgp
RS G8124-E(config-router-bgp)# as 135
RS G8124-E(config-router-bgp)# exit
RS G8124-E(config)# ip router-id 10.1.1.135
```

3. Configure internal peer router 1 and external peer router 2 with IPv4 addresses.

```
RS G8124-E(config)# router bgp
RS G8124-ERS G8124-E(config-router-bgp)# neighbor 1 remote-address
10.1.1.4
RS G8124-E(config-router-bgp)# neighbor 1 remote-as 135
RS G8124-E(config-router-bgp)# no neighbor 1 shutdown
RS G8124-ERS G8124-E(config-router-bgp)# neighbor 2 remote-address
20.20.20.2
RS G8124-E(config-router-bgp)# neighbor 2 remote-as 200
RS G8124-E(config-router-bgp)# no neighbor 2 shutdown
```

4. Configure redistribution for Peer 1.

```
RS G8124-E(config-router-bgp)# neighbor 1 redistribute default-action
    redistribute
RS G8124-E(config-router-bgp)# neighbor 1 redistribute fixed
```

5. Configure aggregation policy control.

Configure the IPv4 routes that you want aggregated.

```
RS G8124-E(config-router-bgp)# aggregate-address 1 135.0.0.0 255.0.0.0
RS G8124-E(config-router-bgp)# aggregate-address 1 enable
```

Chapter 27. Open Shortest Path First

Lenovo Networking OS supports the Open Shortest Path First (OSPF) routing protocol. The Lenovo N/OS implementation conforms to the OSPF version 2 specifications detailed in Internet RFC 1583, and OSPF version 3 specifications in RFC 5340. The following sections discuss OSPF support for the RackSwitch G8124-E:

- "OSPFv2 Overview" on page 378. This section provides information on OSPFv2 concepts, such as types of OSPF areas, types of routing devices, neighbors, adjacencies, link state database, authentication, and internal versus external routing.

- "OSPFv2 Implementation in Lenovo N/OS" on page 383. This section describes how OSPFv2 is implemented in N/OS, such as configuration parameters, electing the designated router, summarizing routes, defining route maps and so forth.

- "OSPFv2 Configuration Examples" on page 393. This section provides step-by-step instructions on configuring different OSPFv2 examples:

  o Creating a simple OSPF domain

  o Creating virtual links

  o Summarizing routes

- "OSPFv3 Implementation in Lenovo N/OS" on page 402. This section describes differences and additional features found in OSPFv3.

OSPFv2 Overview

OSPF is designed for routing traffic within a single IP domain called an Autonomous System (AS). The AS can be divided into smaller logical units known as *areas*.

All routing devices maintain link information in their own Link State Database (LSDB). OSPF allows networks to be grouped together into an area. The topology of an area is hidden from the rest of the AS, thereby reducing routing traffic. Routing within an area is determined only by the area's own topology, thus protecting it from bad routing data. An area can be generalized as an IP subnetwork.

The following sections describe key OSPF concepts.

Types of OSPF Areas

An AS can be broken into logical units known as *areas*. In any AS with multiple areas, one area must be designated as area 0, known as the *backbone*. The backbone acts as the central OSPF area. All other areas in the AS must be connected to the backbone. Areas inject summary routing information into the backbone, which then distributes it to other areas as needed.

As shown in Figure 40, OSPF defines the following types of areas:

- Stub Area—an area that is connected to only one other area. External route information is not distributed into stub areas.

- Not-So-Stubby-Area (NSSA)—similar to a stub area with additional capabilities. Routes originating from within the NSSA can be propagated to adjacent transit and backbone areas. External routes from outside the AS can be advertised within the NSSA but can be configured to not be distributed into other areas.

- Transit Area—an area that carries data traffic which neither originates nor terminates in the area itself.

Figure 40. OSPF Area Types
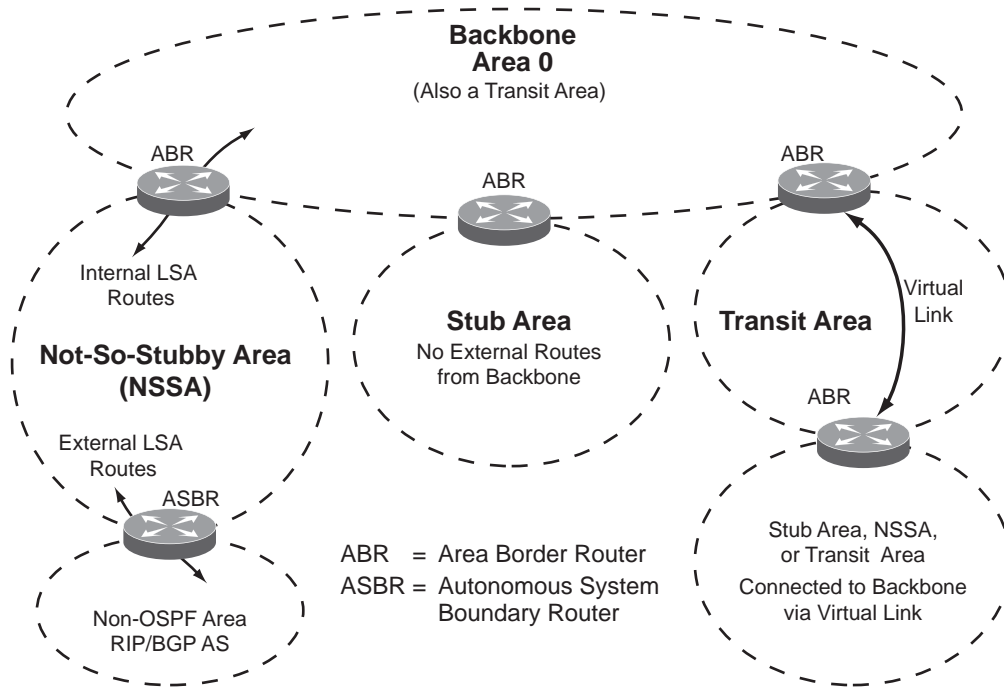
Types of OSPF Routing Devices

As shown in Figure 41, OSPF uses the following types of routing devices:

- Internal Router (IR)—a router that has all of its interfaces within the same area. IRs maintain LSDBs identical to those of other routing devices within the local area.

- Area Border Router (ABR)—a router that has interfaces in multiple areas. ABRs maintain one LSDB for each connected area and disseminate routing information between areas.

- Autonomous System Boundary Router (ASBR)—a router that acts as a gateway between the OSPF domain and non-OSPF domains, such as RIP, BGP, and static routes.

Figure 41. OSPF Domain and an Autonomous System
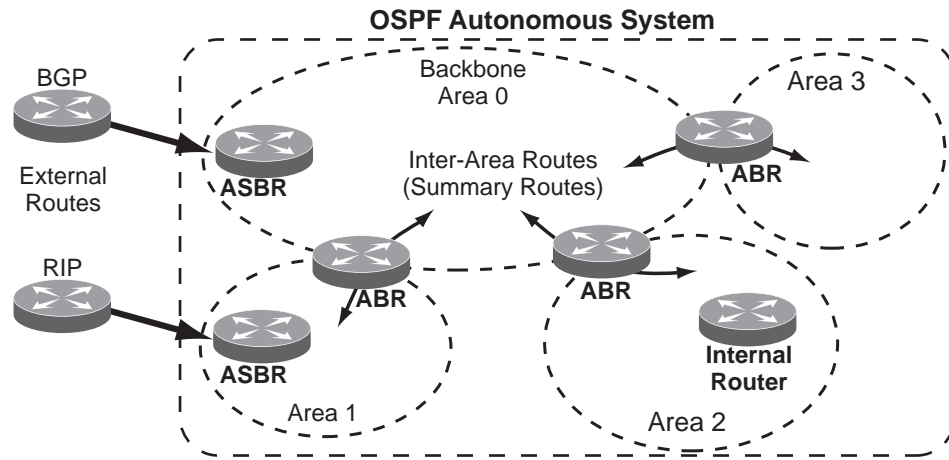
Neighbors and Adjacencies

In areas with two or more routing devices, *neighbors* and *adjacencies* are formed.

Neighbors are routing devices that maintain information about each others' state. To establish neighbor relationships, routing devices periodically send hello packets on each of their interfaces. All routing devices that share a common network segment, appear in the same area, and have the same health parameters (`hello` and `dead` intervals) and authentication parameters respond to each other's hello packets and become neighbors. Neighbors continue to send periodic hello packets to advertise their health to neighbors. In turn, they listen to hello packets to determine the health of their neighbors and to establish contact with new neighbors.

The hello process is used for electing one of the neighbors as the network segment's Designated Router (DR) and one as the network segment's Backup Designated Router (BDR). The DR is adjacent to all other neighbors on that specific network segment and acts as the central contact for database exchanges. Each neighbor sends its database information to the DR, which relays the information to the other neighbors.

The BDR is adjacent to all other neighbors (including the DR). Each neighbor sends its database information to the BDR just as with the DR, but the BDR merely stores this data and does not distribute it. If the DR fails, the BDR will take over the task of distributing database information to the other neighbors.

The Link-State Database

OSPF is a link-state routing protocol. A *link* represents an interface (or routable path) from the routing device. By establishing an adjacency with the DR, each routing device in an OSPF area maintains an identical Link-State Database (LSDB) describing the network topology for its area.

Each routing device transmits a Link-State Advertisement (LSA) on each of its *active* interfaces. LSAs are entered into the LSDB of each routing device. OSPF uses *flooding* to distribute LSAs between routing devices. Interfaces may also be *passive*. Passive interfaces send LSAs to active interfaces, but do not receive LSAs, hello packets, or any other OSPF protocol information from active interfaces. Passive interfaces behave as stub networks, allowing OSPF routing devices to be aware of devices that do otherwise participate in OSPF (either because they do not support it, or because the administrator chooses to restrict OSPF traffic exchange or transit).

When LSAs result in changes to the routing device's LSDB, the routing device forwards the changes to the adjacent neighbors (the DR and BDR) for distribution to the other neighbors.

OSPF routing updates occur only when changes occur, instead of periodically. For each new route, if a neighbor is interested in that route (for example, if configured to receive static routes and the new route is indeed static), an update message containing the new route is sent to the adjacency. For each route removed from the route table, if the route has already been sent to a neighbor, an update message containing the route to withdraw is sent.

The Shortest Path First Tree

The routing devices use a link-state algorithm (Dijkstra's algorithm) to calculate the shortest path to all known destinations, based on the cumulative *cost* required to reach the destination.

The cost of an individual interface in OSPF is an indication of the overhead required to send packets across it.

Internal Versus External Routing

To ensure effective processing of network traffic, every routing device on your network needs to know how to send a packet (directly or indirectly) to any other location/destination in your network. This is referred to as *internal routing* and can be done with static routes or using active internal routing protocols, such as OSPF, RIP, or RIPv2.

It is also useful to tell routers outside your network (upstream providers or *peers*) about the routes you have access to in your network. Sharing of routing information between autonomous systems is known as *external routing*.

Typically, an AS will have one or more border routers (peer routers that exchange routes with other OSPF networks) as well as an internal routing system enabling every router in that AS to reach every other router and destination within that AS.

When a routing device *advertises* routes to boundary routers on other autonomous systems, it is effectively committing to carry data to the IP space represented in the route being advertised. For example, if the routing device advertises 192.204.4.0/24, it is declaring that if another router sends data destined for any address in the 192.204.4.0/24 range, it will carry that data to its destination.

OSPFv2 Implementation in Lenovo N/OS

N/OS supports a single instance of OSPF and up to 4K routes on the network. The following sections describe OSPF implementation in N/OS:

Configurable Parameters

In N/OS, OSPF parameters can be configured through the Industry Standard Command Line Interfaces (ISCLI), Browser-Based Interface (BBI), or through SNMP. For more information, see Chapter 1, "Switch Administration."

The ISCLI supports the following parameters: interface output cost, interface priority, dead and hello intervals, retransmission interval, and interface transmit delay.

In addition to the preceding parameters, you can specify the following:

- Shortest Path First (SPF) interval—Time interval between successive calculations of the shortest path tree using the Dijkstra's algorithm.

- Stub area metric—A stub area can be configured to send a numeric metric value such that all routes received via that stub area carry the configured metric to potentially influence routing decisions.

- Default routes—Default routes with weight metrics can be manually injected into transit areas. This helps establish a preferred route when multiple routing devices exist between two areas. It also helps route traffic to external networks.

- Passive—When enabled, the interface sends LSAs to upstream devices, but does not otherwise participate in OSPF protocol exchanges.

- Point-to-Point—For LANs that have only two OSPF routing agents (the G8124-E and one other device), this option allows the switch to significantly reduce the amount of routing information it must carry and manage.

Defining Areas

If you are configuring multiple areas in your OSPF domain, one of the areas must be designated as area 0, known as the *backbone*. The backbone is the central OSPF area and is usually physically connected to all other areas. The areas inject routing information into the backbone which, in turn, disseminates the information into other areas.

Since the backbone connects the areas in your network, it must be a contiguous area. If the backbone is partitioned (possibly as a result of joining separate OSPF networks), parts of the AS will be unreachable, and you will need to configure *virtual links* to reconnect the partitioned areas (see "Virtual Links" on page 388).

Up to 20 OSPF areas can be connected to the G8124-E with N/OS software. To configure an area, the OSPF number must be defined and then attached to a network interface on the switch. The full process is explained in the following sections.

An OSPF area is defined by assigning **two** pieces of information: an *area index* and an *area ID*. The commands to define and enable an OSPF area are as follows:

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area <area index> area-id <n.n.n.n>
RS G8124-E(config-router-ospf)# area <area index> enable
RS G8124-E(config-router-ospf)# exit
```

Note: The area option is an arbitrary index used only on the switch and does not represent the actual OSPF area number. The actual OSPF area number is defined in the area portion of the command as explained in the following sections.

Assigning the Area Index

The area *<area index>* option is actually just an arbitrary index (0-19) used only by the G8124-E. This index number does not necessarily represent the OSPF area number, though for configuration simplicity, it ought to where possible.

For example, both of the following sets of commands define OSPF area 0 (the backbone) and area 1 because that information is held in the area ID portion of the command. However, the first set of commands is easier to maintain because the arbitrary area indexes agree with the area IDs:

- Area index and area ID agree

  area 0 area-id 0.0.0.0   *(Use index 0 to set area 0 in ID octet format)*

  area 1 area-id 0.0.0.1   *(Use index 1 to set area 1 in ID octet format)*

- Area index set to an arbitrary value

  area 1 area-id 0.0.0.0   *(Use index 1 to set area 0 in ID octet format)*

  area 2 area-id 0.0.0.1   *(Use index 2 to set area 1 in ID octet format)*

Using the Area ID to Assign the OSPF Area Number

The OSPF area number is defined in the `areaid` *<IP address>* option. The octet format is used to be compatible with two different systems of notation used by other OSPF network vendors. There are two valid ways to designate an area ID:

- Single Number

  Most common OSPF vendors express the area ID number as a single number. For example, the Cisco IOS-based router command "`network 1.1.1.0 0.0.0.255 area 1`" defines the area number simply as "`area 1`."

- Multi-octet (*IP address*): Placing the area number in the last octet (0.0.0.*n*)

  Some OSPF vendors express the area ID number in multi-octet format. For example, "`area 0.0.0.2`" represents OSPF area 2 and can be specified directly on the G8124-E as "`area-id 0.0.0.2`".

On the G8124-E, using the last octet in the area ID, "`area 1`" is equivalent to "`area-id 0.0.0.1`".

Note: Although both types of area ID formats are supported, be sure that the area IDs are in the same format throughout an area.

Attaching an Area to a Network

Once an OSPF area has been defined, it must be associated with a network. To attach the area to a network, you must assign the OSPF area index to an IP interface that participates in the area. The format for the command is as follows:

```
RS G8124-E(config)# interface ip <interface number>
RS G8124-E(config-ip-if)# ip ospf area <area index>
RS G8124-E(config-ip-if)# exit
```

For example, the following commands could be used to configure IP interface 14 to use 10.10.10.1 on the 10.10.10.0/24 network, to define OSPF area 1, and to attach the area to the network:

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area 1 area-id 0.0.0.1
RS G8124-E(config-router-ospf)# area 1 enable
RS G8124-E(config-router-ospf)# enable
RS G8124-E(config-router-ospf)# exit
RS G8124-E(config)# interface ip 14
RS G8124-E(config-ip-if)# ip address 10.10.10.1
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# ip ospf area 1
RS G8124-E(config-ip-if)# ip ospf enable
```

Note: OSPFv2 supports IPv4 only. IPv6 is supported in OSPFv3 (see "OSPFv3 Implementation in Lenovo N/OS" on page 402).

Interface Cost

The OSPF link-state algorithm (Dijkstra's algorithm) places each routing device at the root of a tree and determines the cumulative *cost* required to reach each destination. You can manually enter the cost for the output route with the following command (Interface IP mode):

```
RS G8124-E(config-ip-if)# ip ospf cost <cost value (1-65535)>
```

Electing the Designated Router and Backup

In any broadcast type subnet, a Designated Router (DR) is elected as the central contact for database exchanges among neighbors. On subnets with more the one device, a Backup Designated Router (BDR) is elected in case the DR fails.

DR and BDR elections are made through the `hello` process. The election can be influenced by assigning a priority value to the OSPF interfaces on the G8124-E. The command is as follows:

```
RS G8124-E(config-ip-if)# ip ospf priority <priority value (0-255)>
```

A priority value of 255 is the highest, and 1 is the lowest. A priority value of 0 specifies that the interface cannot be used as a DR or BDR. In case of a tie, the routing device with the highest router ID wins. Interfaces configured as *passive* do not participate in the DR or BDR election process:

```
RS G8124-E(config-ip-if)# ip ospf passive-interface
RS G8124-E(config-ip-if)# exit
```

Summarizing Routes

Route summarization condenses routing information. Without summarization, each routing device in an OSPF network would retain a route to every subnet in the network. With summarization, routing devices can reduce some sets of routes to a single advertisement, reducing both the load on the routing device and the perceived complexity of the network. The importance of route summarization increases with network size.

Summary routes can be defined for up to 16 IP address ranges using the following command:

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area-range <range number> address <IP address>
        <mask>
```

where *<range number>* is a number 1 to 16, *<IP address>* is the base IP address for the range, and *<mask>* is the IP address mask for the range. For a detailed configuration example, see "Example 3: Summarizing Routes" on page 400.

Default Routes

When an OSPF routing device encounters traffic for a destination address it does not recognize, it forwards that traffic along the *default route*. Typically, the default route leads upstream toward the backbone until it reaches the intended area or an external router.

Each G8124-E acting as an ABR automatically inserts a default route into stub and NSSA attached areas. In simple OSPF stub areas or NSSAs with only one ABR leading upstream (see Area 1 in Figure 42), any traffic for IP address destinations outside the area is forwarded to the switch's IP interface, and then into the connected transit area (usually the backbone). Since this is automatic, no further configuration is required for such areas.

Figure 42. Injecting Default Routes
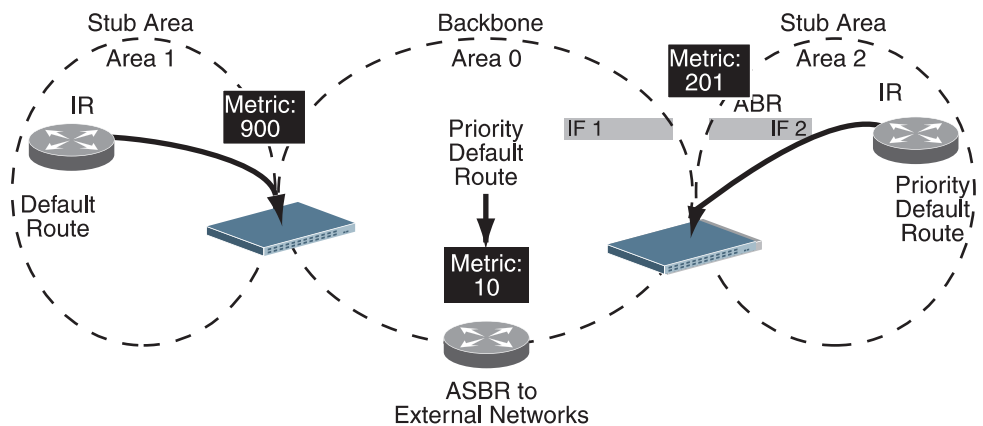
If the switch is in a transit area and has a configured default gateway, it can inject a default route into rest of the OSPF domain. Use the following command to configure the switch to inject OSPF default routes (Router OSPF mode):

```
RS G8124-E(config-router-ospf)# default-information <metric value> <metric type (1 or 2)>
```

In this command, *<metric value>* sets the priority for choosing this switch for default route. The value none sets no default and 1 sets the highest priority for default route. Metric type determines the method for influencing routing decisions for external routes.

When the switch is configured to inject a default route, an AS-external LSA with link state ID 0.0.0.0 is propagated throughout the OSPF routing domain. This LSA is sent with the configured metric value and metric type.

The OSPF default route configuration can be removed with the command:

```
RS G8124-E(config-router-ospf)# no default-information
```

Virtual Links

Usually, all areas in an OSPF AS are physically connected to the backbone. In some cases where this is not possible, you can use a *virtual link*. Virtual links are created to connect one area to the backbone through another non-backbone area (see Figure 40 on page 379).

The area which contains a virtual link must be a transit area and have full routing information. Virtual links cannot be configured inside a stub area or NSSA. The area type must be defined as transit using the following command:

```
RS G8124-E(config-router-ospf)# area <area index> type transit
```

The virtual link must be configured on the routing devices at each endpoint of the virtual link, though they may traverse multiple routing devices. To configure a G8124-E as one endpoint of a virtual link, use the following command:

```
RS G8124-E(config-router-ospf)# area-virtual-link <link number>
neighbor-router <router ID>
```

where *<link number>* is a value between 1 and 3, *<area index>* is the OSPF area index of the transit area, and *<router ID>* is the router ID of the virtual neighbor, the routing device at the target endpoint. Another router ID is needed when configuring a virtual link in the other direction. To provide the G8124-E with a router ID, see the following section Router ID.

For a detailed configuration example on Virtual Links, see "Example 2: Virtual Links" on page 396.

Router ID

Routing devices in OSPF areas are identified by a router ID. The router ID is expressed in IP address format. The IP address of the router ID is not required to be included in any IP interface range or in any OSPF area, and may even use the G8124-E loopback interface.

The router ID can be configured in one of the following two ways:

- Dynamically—OSPF protocol configures the lowest IP interface IP address as the router ID (loopback interface has priority over the IP interface). This is the default.
- Statically—Use the following command to manually configure the router ID:

```
RS G8124-E(config-router-ospf)# ip router-id <IPv4 address>
```

If there is a loopback interface, its IP address is always preferred as the router ID, instead of an IP interface address. The **ip router-id** command is the preferred method to set the router ID and it is always used in preference to the other methods.

- To modify the router ID from static to dynamic, set the router ID to 0.0.0.0 and re-initialize the OSPF process.
- To view the router ID, use the following command:

```
RS G8124-E(config-router-ospf)# show ip ospf
```

Authentication

OSPF protocol exchanges can be authenticated so that only trusted routing devices can participate. This ensures less processing on routing devices that are not listening to OSPF packets.

OSPF allows packet authentication and uses IP multicast when sending and receiving packets. Routers participate in routing domains based on pre-defined passwords. N/OS supports simple password (type 1 plain text passwords) and MD5 cryptographic authentication. This type of authentication allows a password to be configured per interface.

We strongly recommend that you implement MD5 cryptographic authentication as a best practice.

Figure 43 shows authentication configured for area 0 with the password test. Simple authentication is also configured for the virtual link between area 2 and area 0. Area 1 is not configured for OSPF authentication.

Figure 43. OSPF Authentication
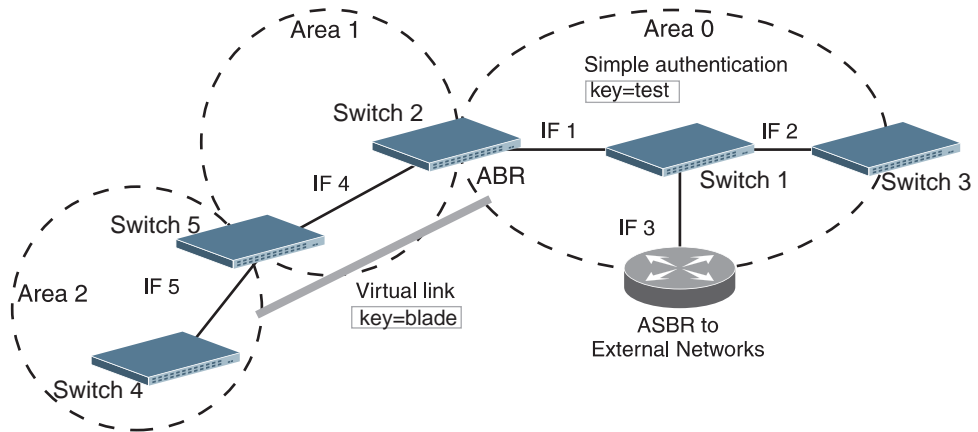
Configuring Plain Text OSPF Passwords

To configure simple plain text OSPF passwords on the switches shown in Figure 43 use the following commands:

1. Enable OSPF authentication for Area 0 on switches 1, 2, and 3.

```
RS G8124-E(config-router-ospf)# area 0 authentication-type password
RS G8124-E(config-router-ospf)# exit
```

2. Configure a simple text password up to eight characters for each OSPF IP interface in Area 0 on switches 1, 2, and 3.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip ospf key test
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip ospf key test
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 3
RS G8124-E(config-ip-if)# ip ospf key test
RS G8124-E(config-ip-if)# exit
```

3. Enable OSPF authentication for Area 2 on switch 4.

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area 2 authentication-type password
```

4. Configure a simple text password up to eight characters for the virtual link between Area 2 and Area 0 on switches 2 and 4.

```
RS G8124-E(config-router-ospf)# area-virtual-link 1 key blade
```

Configuring MD5 Authentication

Use the following commands to configure MD5 authentication on the switches shown in Figure 43:

1. Enable OSPF MD5 authentication for Area 0 on switches 1, 2, and 3.

```
RS G8124-E(config-router-ospf)# area 0 authentication-type md5
```

2. Configure MD5 key ID for Area 0 on switches 1, 2, and 3.

```
RS G8124-E(config-router-ospf)# message-digest-key 1 md5-key test
RS G8124-E(config-router-ospf)# exit
```

3. Assign MD5 key ID to OSPF interfaces on switches 1, 2, and 3.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip ospf message-digest-key 1
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip ospf message-digest-key 1
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 3
RS G8124-E(config-ip-if)# ip ospf message-digest-key 1
RS G8124-E(config-ip-if)# exit
```

4. Enable OSPF MD5 authentication for Area 2 on switch 4.

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area 1 authentication-type md5
```

5. Configure MD5 key for the virtual link between Area 2 and Area 0 on switches 2 and 4.

```
RS G8124-E(config-router-ospf)# message-digest-key 2 md5-key test
```

6. Assign MD5 key ID to OSPF virtual link on switches 2 and 4.

```
RS G8124-E(config-router-ospf)# area-virtual-link 1 message-digest-key 2
RS G8124-E(config-router-ospf)# exit
```

Host Routes for Load Balancing

N/OS implementation of OSPF includes host routes. Host routes are used for advertising network device IP addresses to external networks, accomplishing the following goals:

- ABR Load Sharing

  As a form of load balancing, host routes can be used for dividing OSPF traffic among multiple ABRs. To accomplish this, each switch provides identical services but advertises a host route for a different IP address to the external network. If each IP address serves a different and equal portion of the external world, incoming traffic from the upstream router must be split evenly among ABRs.

- ABR Failover

  Complementing ABR load sharing, identical host routes can be configured on each ABR. These host routes can be given different costs so that a different ABR is selected as the preferred route for each server and the others are available as backups for failover purposes.

- Equal Cost Multipath (ECMP)

  With equal cost multipath, a router potentially has several available next hops towards any given destination. ECMP allows separate routes to be calculated for each IP Type of Service. All paths of equal cost to a given destination are calculated, and the next hops for all equal-cost paths are inserted into the routing table.

Loopback Interfaces in OSPF

A loopback interface is an IP interface which has an IP address, but is not associated with any particular physical port. The loopback interface is thus always available to the general network, regardless of which specific ports are in operation. Because loopback interfaces are always available on the switch, loopback interfaces may present an advantage when used as the router ID.

If dynamic router ID selection is used (see "Router ID" on page 388) loopback interfaces can be used to force router ID selection. If a loopback interface is configured, its IP address is automatically selected as the router ID, even if other IP interfaces have lower IP addresses. If more than one loopback interface is configured, the lowest loopback interface IP address is selected.

Loopback interfaces can be advertised into the OSPF domain by specifying an OSPF host route with the loopback interface IP address.

To enable OSPF on an existing loopback interface:

```
RS G8124-E(config)# interface loopback <1-5>
RS G8124-E(config-ip-loopback)# ip ospf area <area ID>
RS G8124-E(config-ip-loopback)# ip ospf enable
RS G8124-E(config-ip-loopback)# exit
```

OSPF Features Not Supported in This Release

The following OSPF features are not supported in this release:

● Summarizing external routes

● Filtering OSPF routes

● Using OSPF to forward multicast routes

● Configuring OSPF on non-broadcast multi-access networks (such as frame relay, X.25, or ATM)

OSPFv2 Configuration Examples

A summary of the basic steps for configuring OSPF on the G8124-E is listed here. Detailed instructions for each of the steps is covered in the following sections:

1. Configure IP interfaces.

   One IP interface is required for each desired network (range of IP addresses) being assigned to an OSPF area on the switch.

2. (Optional) Configure the router ID.

3. Enable OSPF on the switch.

4. Define the OSPF areas.

5. Configure OSPF interface parameters.

   IP interfaces are used for attaching networks to the various areas.

6. (Optional) Configure route summarization between OSPF areas.

7. (Optional) Configure virtual links.

8. (Optional) Configure host routes.

Example 1: Simple OSPF Domain

In this example, two OSPF areas are defined—one area is the backbone and the other is a stub area. A stub area does not allow advertisements of external routes, thus reducing the size of the database. Instead, a default summary route of IP address 0.0.0.0 is automatically inserted into the stub area. Any traffic for IP address destinations outside the stub area will be forwarded to the stub area's IP interface, and then into the backbone.

Figure 44. A Simple OSPF Domain
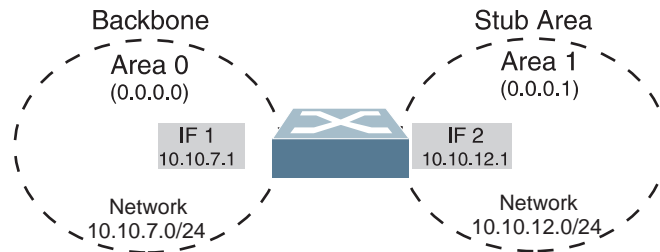
Follow this procedure to configure OSPF support as shown in Figure 44:

1. Configure IP interfaces on each network that will be attached to OSPF areas.

   In this example, two IP interfaces are needed:

   ● Interface 1 for the backbone network on 10.10.7.0/24

   ● Interface 2 for the stub area network on 10.10.12.0/24

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip address 10.10.7.1
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip address 10.10.12.1
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

   Note: OSPFv2 supports IPv4 only. IPv6 is supported in OSPFv3 (see "OSPFv3 Implementation in Lenovo N/OS" on page 402).

2. Enable OSPF.

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# enable
```

3. Define the backbone.

   The backbone is always configured as a transit area using areaid 0.0.0.0.

```
RS G8124-E(config-router-ospf)# area 0 area-id 0.0.0.0
RS G8124-E(config-router-ospf)# area 0 type transit
RS G8124-E(config-router-ospf)# area 0 enable
```

4. Define the stub area.

```
RS G8124-E(config-router-ospf)# area 1 area-id 0.0.0.1
RS G8124-E(config-router-ospf)# area 1 type stub
RS G8124-E(config-router-ospf)# area 1 enable
RS G8124-E(config-router-ospf)# exit
```

5. Attach the network interface to the backbone.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip ospf area 0
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
```

6. Attach the network interface to the stub area.

```
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip ospf area 1
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
```

Example 2: Virtual Links

In the example shown in Figure 45, area 2 is not physically connected to the backbone as is usually required. Instead, area 2 will be connected to the backbone via a virtual link through area 1. The virtual link must be configured at each endpoint.

Figure 45. Configuring a Virtual Link

Note: OSPFv2 supports IPv4 only. IPv6 is supported in OSPFv3 (see "OSPFv3 Implementation in Lenovo N/OS" on page 402).

Configuring OSPF for a Virtual Link on Switch #1

1. Configure IP interfaces on each network that will be attached to the switch.

   In this example, two IP interfaces are needed:

   ● Interface 1 for the backbone network on 10.10.7.0/24

   ● Interface 2 for the transit area network on 10.10.12.0/24

   ```
   RS G8124-E(config)# interface ip 1
   RS G8124-E(config-ip-if)# ip address 10.10.7.1
   RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
   RS G8124-E(config-ip-if)# enable
   RS G8124-E(config-ip-if)# exit
   RS G8124-E(config)# interface ip 2
   RS G8124-E(config-ip-if)# ip address 10.10.12.1
   RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
   RS G8124-E(config-ip-if)# enable
   RS G8124-E(config-ip-if)# exit
   ```

2. Configure the router ID.

   A router ID is required when configuring virtual links. Later, when configuring the other end of the virtual link on Switch 2, the router ID specified here will be used as the target virtual neighbor (nbr) address.

   ```
   RS G8124-E(config)# ip router-id 10.10.10.1
   ```

3. Enable OSPF.

   ```
   RS G8124-E(config)# router ospf
   RS G8124-E(config-router-ospf)# enable
   ```

4. Define the backbone.

```
RS G8124-E(config-router-ospf)# area 0 area-id 0.0.0.0
RS G8124-E(config-router-ospf)# area 0 type transit
RS G8124-E(config-router-ospf)# area 0 enable
```

5. Define the transit area.

   The area that contains the virtual link must be configured as a transit area.

```
RS G8124-E(config-router-ospf)# area 1 area-id 0.0.0.1
RS G8124-E(config-router-ospf)# area 1 type transit
RS G8124-E(config-router-ospf)# area 1 enable
RS G8124-E(config-router-ospf)# exit
```

6. Attach the network interface to the backbone.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip ospf area 0
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
```

7. Attach the network interface to the transit area.

```
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip ospf area 1
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
```

8. Configure the virtual link.

   The nbr router ID configured in this step must be the same as the router ID that will be configured for Switch #2 in .

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area-virtual-link 1 area 1
RS G8124-E(config-router-ospf)# area-virtual-link 1 neighbor-router
      10.10.14.1
RS G8124-E(config-router-ospf)# area-virtual-link 1 enable
```

Configuring OSPF for a Virtual Link on Switch #2

1. Configure IP interfaces on each network that will be attached to OSPF areas.

   In this example, two IP interfaces are needed:

   ● Interface 1 for the transit area network on 10.10.12.0/24

Chapter 27: Open Shortest Path First **397**

- Interface 2 for the stub area network on 10.10.24.0/24

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip address 10.10.12.2
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# vlan 10
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip address 10.10.24.1
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# vlan 20
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

2. Configure the router ID.

A router ID is required when configuring virtual links. This router ID must be the same one specified as the target virtual neighbor (nbr) on switch 1 in .

```
RS G8124-E(config)# ip router-id 10.10.14.1
```

3. Enable OSPF.

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# enable
```

4. Define the backbone.

This version of N/OS requires that a backbone index be configured on the non-backbone end of the virtual link as follows:

```
RS G8124-E(config-router-ospf)# area 0 area-id 0.0.0.0
RS G8124-E(config-router-ospf)# area 0 enable
```

5. Define the transit area.

```
RS G8124-E(config-router-ospf)# area 1 area-id 0.0.0.1
RS G8124-E(config-router-ospf)# area 1 type transit
RS G8124-E(config-router-ospf)# area 1 enable
```

6. Define the stub area.

```
RS G8124-E(config-router-ospf)# area 2 area-id 0.0.0.2
RS G8124-E(config-router-ospf)# area 2 type stub
RS G8124-E(config-router-ospf)# area 2 enable
RS G8124-E(config-router-ospf)# exit
```

7. Attach the network interface to the transmit area.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip ospf area 1
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
```
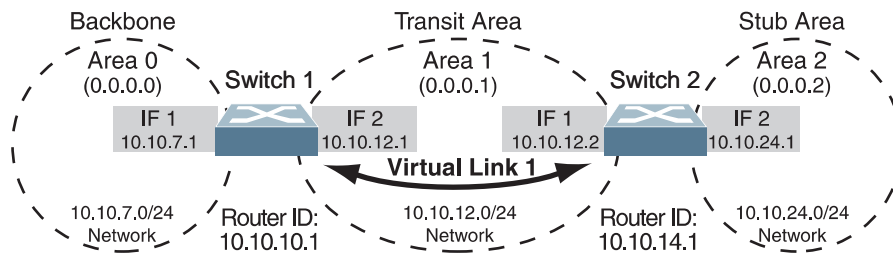
8. Attach the network interface to the stub area.

```
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip ospf area 2
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
```

9. Configure the virtual link.

The nbr router ID configured in this step must be the same as the router ID that was configured for switch #1 in Step 2 on page 396.

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area-virtual-link 1 area 1
RS G8124-E(config-router-ospf)# area-virtual-link 1 neighbor-router
      10.10.10.1
RS G8124-E(config-router-ospf)# area-virtual-link 1 enable
```

Other Virtual Link Options

- You can use redundant paths by configuring multiple virtual links.

- Only the endpoints of the virtual link are configured. The virtual link path may traverse multiple routers in an area as long as there is a routable path between the endpoints.

Example 3: Summarizing Routes

By default, ABRs advertise all the network addresses from one area into another area. Route summarization can be used for consolidating advertised addresses and reducing the perceived complexity of the network.

If network IP addresses in an area are assigned to a contiguous subnet range, you can configure the ABR to advertise a single summary route that includes all individual IP addresses within the area.

The following example shows one summary route from area 1 (stub area) injected into area 0 (the backbone). The summary route consists of all IP addresses from 36.128.192.0 through 36.128.254.255 except for the routes in the range 36.128.200.0 through 36.128.200.255.

Figure 46. Summarizing Routes
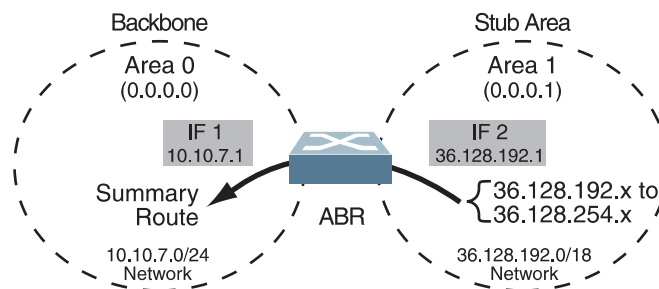
Notes:

- OSPFv2 supports IPv4 only. IPv6 is supported in OSPFv3 (see "OSPFv3 Implementation in Lenovo N/OS" on page 402).

- You can specify a range of addresses to prevent advertising by using the hide option. In this example, routes in the range 36.128.200.0 through 36.128.200.255 are kept private.

Use the following procedure to configure OSPF support as shown in Figure 46:

1. Configure IP interfaces for each network which will be attached to OSPF areas.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip address 10.10.7.1
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# vlan 10
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip address 36.128.192.1
RS G8124-E(config-ip-if)# ip netmask 255.255.255.0
RS G8124-E(config-ip-if)# vlan 20
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

2. Enable OSPF.

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# enable
```

3. Define the backbone.

```
RS G8124-E(config-router-ospf)# area 0 area-id 0.0.0.0
RS G8124-E(config-router-ospf)# area 0 type transit
RS G8124-E(config-router-ospf)# area 0 enable
```

4. Define the stub area.

```
RS G8124-E(config-router-ospf)# area 1 area-id 0.0.0.1
RS G8124-E(config-router-ospf)# area 1 type stub
RS G8124-E(config-router-ospf)# area 1 enable
RS G8124-E(config-router-ospf)# exit
```

5. Attach the network interface to the backbone.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip ospf area 0
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
```

6. Attach the network interface to the stub area.

```
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip ospf area 1
RS G8124-E(config-ip-if)# ip ospf enable
RS G8124-E(config-ip-if)# exit
```

7. Configure route summarization by specifying the starting address and mask of the range of addresses to be summarized.

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area-range 1 address 36.128.192.0
     255.255.192.0
RS G8124-E(config-router-ospf)# area-range 1 area 1
RS G8124-E(config-router-ospf)# area-range 1 enable
RS G8124-E(config-router-ospf)# exit
```

8. Use the hide command to prevent a range of addresses from advertising to the backbone.

```
RS G8124-E(config)# router ospf
RS G8124-E(config-router-ospf)# area-range 2 address 36.128.200.0
     255.255.255.0
RS G8124-E(config-router-ospf)# area-range 2 area 1
RS G8124-E(config-router-ospf)# area-range 2 hide
RS G8124-E(config-router-ospf)# exit
```

Verifying OSPF Configuration

Use the following commands to verify the OSPF configuration on your switch:

- **show ip ospf**
- **show ip ospf neighbor**
- **show ip ospf database database-summary**
- **show ip ospf routes**

Refer to the *Lenovo Networking OS Command Reference* for information on the preceding commands.

OSPFv3 Implementation in Lenovo N/OS

OSPF version 3 is based on OSPF version 2, but has been modified to support IPv6 addressing. In most other ways, OSPFv3 is similar to OSPFv2: They both have the same packet types and interfaces, and both use the same mechanisms for neighbor discovery, adjacency formation, LSA flooding, aging, and so on. The administrator must be familiar with the OSPFv2 concepts covered in the preceding sections of this chapter before implementing the OSPFv3 differences as described in the following sections.

Although OSPFv2 and OSPFv3 are very similar, they represent independent features on the G8124-E. They are configured separately, and both can run in parallel on the switch with no relation to one another, serving different IPv6 and IPv4 traffic, respectively.

The Lenovo N/OS implementation conforms to the OSPF version 3 authentication/confidentiality specifications in RFC 4552.

OSPFv3 Differences from OSPFv2

Note: When OSPFv3 is enabled, the OSPF backbone area (0.0.0.0) is created by default and is always active.

OSPFv3 Requires IPv6 Interfaces

OSPFv3 is designed to support IPv6 addresses. This requires IPv6 interfaces to be configured on the switch and assigned to OSPF areas, in much the same way IPv4 interfaces are assigned to areas in OSPFv2. This is the primary configuration difference between OSPFv3 and OSPFv2.

See Chapter 21, "Internet Protocol Version 6," for configuring IPv6 interfaces.

OSPFv3 Uses Independent Command Paths

Though OSPFv3 and OSPFv2 are very similar, they are configured independently. OSPFv3 command paths are located as follows:

```
RS G8124-E(config)# ipv6 router ospf     (OSPFv3 router config mode)
RS G8124-E(config-router-ospf3)# ?

RS G8124-E(config)# interface ip <Interface number>(Configure OSPFv3)
RS G8124-E(config-ip-if)# ipv6 ospf ?   (OSPFv3 interface config)

RS G8124-E# show ipv6 ospf ?             (Show OSPFv3 information)
```

OSPFv3 Identifies Neighbors by Router ID

Where OSPFv2 uses a mix of IPv4 interface addresses and Router IDs to identify neighbors, depending on their type, OSPFv3 configuration consistently uses a Router ID to identify all neighbors.

Although Router IDs are written in dotted decimal notation, and may even be based on IPv4 addresses from an original OSPFv2 network configuration, it is important to realize that Router IDs are not IP addresses in OSPFv3, and can be assigned independently of IP address space. However, maintaining Router IDs consistent with any legacy OSPFv2 IPv4 addressing allows for easier implementation of both protocols.

Other Internal Improvements

OSPFv3 has numerous improvements that increase the protocol efficiency in addition to supporting IPv6 addressing. These improvements change some of the behaviors in the OSPFv3 network and may affect topology consideration, but have little direct impact on configuration. For example:

- Addressing fields have been removed from Router and Network LSAs.

- Flexible treatment of unknown LSA types to make integration of OSPFv3 easier.

- Interface network type can be specified using the command:
  `RS G8124-E(config-ip-if)# ipv6 ospf network {broadcast|non-broadcast|point-to-multipoint|point-to-point}`

- For an interface network type that is not broadcast or NBMA, link LSA suppression can be enabled so link LSA is not originated for the interface. Use the command: `RS G8124-E(config-ip-if)# ipv6 ospf linklsasuppress`

OSPFv3 Limitations

N/OS 8.3 does not currently support the following OSPFv3 features:

- Multiple interfaces of OSPFv3 on the same link.

OSPFv3 Configuration Example

The following example depicts the OSPFv3 equivalent configuration of for OSPFv2.

In this example, one summary route from area 1 (stub area) is injected into area 0 (the backbone). The summary route consists of all IP addresses from the 36::0/32 portion of the 36::0/56 network, except for the routes in the 36::0/8 range.

Figure 47. Summarizing Routes
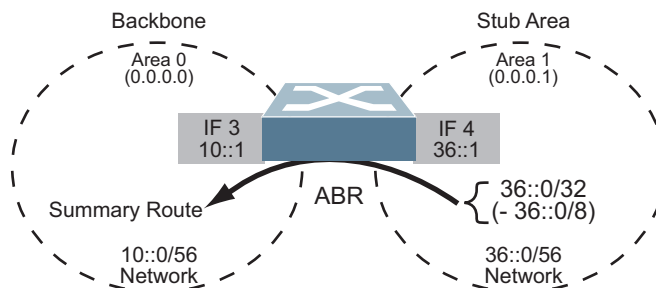
Note: You can specify a range of addresses to prevent advertising by using the hide option. In this example, routes in the 36::0/8 range are kept private.

Use the following procedure to configure OSPFv3 support as shown in Figure 46:

1. Configure IPv6 interfaces for each link which will be attached to OSPFv3 areas.

```
RS G8124-ERS G8124-E(config)# interface ip 3
(config-ip-if)# ipv6 address 10:0:0:0:0:0:0:1
RS G8124-E(config-ip-if)# ipv6 prefixlen 56
RS G8124-E(config-ip-if)# vlan 10
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 4
RS G8124-E(config-ip-if)# ip address 36:0:0:0:0:0:1
RS G8124-E(config-ip-if)# ipv6 prefixlen 56
RS G8124-E(config-ip-if)# vlan 20
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

This is equivalent to configuring the IP address and netmask for IPv4 interfaces.

2. Enable OSPFv3.

```
RS G8124-E(config)# ipv6 router ospf
RS G8124-E(config-router-ospf3)# enable
```

This is equivalent to the OSPFv2 enable option in the router ospf command path.

3. Define the backbone.

```
RS G8124-E(config-router-ospf3)# area 0 area-id 0.0.0.0
RS G8124-E(config-router-ospf3)# area 0 type transit
RS G8124-E(config-router-ospf3)# area 0 enable
```

This is identical to OSPFv2 configuration.

4. Define the stub area.

```
RS G8124-E(config-router-ospf3)# area 1 area-id 0.0.0.1
RS G8124-E(config-router-ospf3)# area 1 type stub
RS G8124-E(config-router-ospf3)# area 1 enable
RS G8124-E(config-router-ospf3)# exit
```

This is identical to OSPFv2 configuration.

5. Attach the network interface to the backbone.

```
RS G8124-E(config)# interface ip 3
RS G8124-E(config-ip-if)# ipv6 ospf area 0
RS G8124-E(config-ip-if)# ipv6 ospf enable
RS G8124-E(config-ip-if)# exit
```

The ipv6 command path is used instead of the OSPFv2 ip command path

6. Attach the network interface to the stub area.

```
RS G8124-E(config)# interface ip 4
RS G8124-E(config-ip-if)# ipv6 ospf area 1
RS G8124-E(config-ip-if)# ipv6 ospf enable
RS G8124-E(config-ip-if)# exit
```

The ipv6 command path is used instead of the OSPFv2 ip command path

7. Configure route summarization by specifying the starting address and prefix length of the range of addresses to be summarized.

```
RS G8124-E(config)# ipv6 router ospf
RS G8124-E(config-router-ospf3)# area-range 1 address 36:0:0:0:0:0:0:0 32
RS G8124-E(config-router-ospf3)# area-range 1 area 0
RS G8124-E(config-router-ospf3)# area-range 1 enable
```

This differs from OSPFv2 only in that the OSPFv3 command path is used, and the address and prefix are specified in IPv6 format.

8. Use the hide command to prevent a range of addresses from advertising to the backbone.

```
RS G8124-E(config-router-ospf)# area-range 2 address 36:0:0:0:0:0:0:0 8
RS G8124-E(config-router-ospf)# area-range 2 area 0
RS G8124-E(config-router-ospf)# area-range 2 hide
RS G8124-E(config-router-ospf)# exit
```

This differs from OSPFv2 only in that the OSPFv3 command path is used, and the address and prefix are specified in IPv6 format.

Neighbor Configuration Example

When using NBMA or point to multipoint interfaces, you must manually configure neighbors. Following example includes the steps for neighbor configuration.

1. Configure IPv6 interface parameters:

```
RS G8124-E(config# interface ip 10
RS G8124-E(config-ip-if)# ipv6 address 10:0:0:0:0:0:0:12 64
RS G8124-E(config-ip-if)# vlan 10
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# ipv6 ospf area 0
RS G8124-E(config-ip-if)# ipv6 ospf retransmit-interval 5
RS G8124-E(config-ip-if)# ipv6 ospf transmit-delay 1
RS G8124-E(config-ip-if)# ipv6 ospf priority 1
RS G8124-E(config-ip-if)# ipv6 ospf hello-interval 10
RS G8124-E(config-ip-if)# ipv6 ospf dead-interval 40
RS G8124-E(config-ip-if)# ipv6 ospf network point-to-multipoint
RS G8124-E(config-ip-if)# ipv6 ospf poll-interval 120
RS G8124-E(config-ip-if)# ipv6 ospf enable
RS G8124-E(config-ip-if)# exit
```

2. Enable OSPFv3:

```
RS G8124-E(config# ipv6 router ospf
RS G8124-E(config-router-ospf3)# router-id 12.12.12.12
RS G8124-E(config-router-ospf3)# enable
```

3. Define the backbone.

```
RS G8124-E(config-router-ospf3)# area 0 area-id 0.0.0.0
RS G8124-E(config-router-ospf3)# area 0 stability-interval 40
RS G8124-E(config-router-ospf3)# area 0 default-metric 1
RS G8124-E(config-router-ospf3)# area 0 default-metric type 1
RS G8124-E(config-router-ospf3)# area 0 translation-role candidate
RS G8124-E(config-router-ospf3)# area 0 type transit
RS G8124-E(config-router-ospf3)# area 0 enable
```

4. Configure neighbor entry:

```
RS G8124-E(config-router-ospf3)# neighbor 1 address
fe80:0:0:0:dceb:ff:fe00:9
RS G8124-E(config-router-ospf3)# neighbor 1 interface 10
RS G8124-E(config-router-ospf3)# neighbor 1 priority 1
RS G8124-E(config-router-ospf3)# neighbor 1 enable
```

Chapter 28. Protocol Independent Multicast

Lenovo Networking OS supports Protocol Independent Multicast (PIM) in Sparse Mode (PIM-SM) and Dense Mode (PIM-DM).

Note: Lenovo N/OS 8.3 does not support IPv6 for PIM.

The following sections discuss PIM support for the RackSwitch G8124-E:

- "PIM Overview" on page 408
- "Supported PIM Modes and Features" on page 409
- "Basic PIM Settings" on page 410
- "Additional Sparse Mode Settings" on page 413
- "Using PIM with Other Features" on page 416
- "PIM Configuration Examples" on page 417

PIM Overview

PIM is designed for efficiently routing multicast traffic across one or more IPv4 domains. This has benefits for application such as IP television, collaboration, education, and software delivery, where a single source must deliver content (a multicast) to a group of receivers that span both wide-area and inter-domain networks.

Instead of sending a separate copy of content to each receiver, a multicast derives efficiency by sending only a single copy of content toward its intended receivers. This single copy only becomes duplicated when it reaches the target domain that includes multiple receivers, or when it reaches a necessary bifurcation point leading to different receiver domains.

PIM is used by multicast source stations, client receivers, and intermediary routers and switches, to build and maintain efficient multicast routing trees. PIM is protocol independent; It collects routing information using the existing unicast routing functions underlying the IPv4 network, but does not rely on any particular unicast protocol. For PIM to function, a Layer 3 routing protocol (such as BGP, OSPF, RIP, or static routes) must first be configured on the switch.

PIM-SM is a reverse-path routing mechanism. Client receiver stations advertise their willingness to join a multicast group. The local routing and switching devices collect multicast routing information and forward the request toward the station that will provide the multicast content. When the join requests reach the sending station, the multicast data is sent toward the receivers, flowing in the opposite direction of the original join requests.

Some routing and switching devices perform special PIM-SM functions. Within each receiver domain, one router is elected as the Designated Router (DR) for handling multicasts for the domain. DRs forward information to a similar device, the Rendezvous Point (RP), which holds the root tree for the particular multicast group.

Receiver join requests as well as sender multicast content initially converge at the RP, which generates and distributes multicast routing data for the DRs along the delivery path. As the multicast content flows, DRs use the routing tree information obtained from the RP to optimize the paths both to and from send and receive stations, bypassing the RP for the remainder of content transactions if a more efficient route is available.

DRs continue to share routing information with the RP, modifying the multicast routing tree when new receivers join, or pruning the tree when all the receivers in any particular domain are no longer part of the multicast group.

Supported PIM Modes and Features

For each interface attached to a PIM network component, PIM can be configured to operate either in PIM Sparse Mode (PIM-SM) or PIM Dense Mode (PIM-DM).

- PIM-SM is used in networks where multicast senders and receivers comprise a relatively small (sparse) portion of the overall network. PIM-SM uses a more complex process than PIM-DM for collecting and optimizing multicast routes, but minimizes impact on other IP services and is more commonly used.

- PIM-DM is used where multicast devices are a relatively large (dense) portion of the network, with very frequent (or constant) multicast traffic. PIM-DM requires less configuration on the switch than PIM-SM, but uses broadcasts that can consume more bandwidth in establishing and optimizing routes.

The following PIM modes and features are *not* currently supported in N/OS 8.3:

- Hybrid Sparse-Dense Mode (PIM-SM/DM). Sparse Mode and Dense Mode may be configured on separate IP interfaces on the switch, but are not currently supported simultaneously on the same IP interface.
- PIM Source-Specific Multicast (PIM-SSM)
- Anycast RP
- PIM RP filters
- Only configuration via the switch ISCLI is supported. PIM configuration is currently not available using the menu-based CLI, the BBI, or via SNMP.

Basic PIM Settings

To use PIM the following is required:

- The PIM feature must be enabled globally on the switch.
- PIM network components and PIM modes must be defined.
- IP interfaces must be configured for each PIM component.
- PIM neighbor filters may be defined (optional).
- If PIM-SM is used, define additional parameters:
    - Rendezvous Point
    - Designated Router preferences (optional)
    - Bootstrap Router preferences (optional)

Each of these tasks is covered in the following sections.

Note: In N/OS 8.3, PIM can be configured through the ISCLI only. PIM configuration and information are not available using the menu-based CLI, the BBI, or via SNMP.

Globally Enabling or Disabling the PIM Feature

By default, PIM is disabled on the switch. PIM can be globally enabled or disabled using the following commands:

```
RS G8124-E(config)# [no] ip pim enable
```

Defining a PIM Network Component

The G8124-E can be attached to a maximum of two independent PIM network components. Each component represents a different PIM network, and can be defined for either PIM-SM or PIM-DM operation. Basic PIM component configuration is performed using the following commands:

```
RS G8124-E(config)# ip pim component <1-2>
RS G8124-E(config-ip-pim-comp)# mode {sparse|dense}
RS G8124-E(config-ip-pim-comp)# exit
```

The sparse option will place the component in Sparse Mode (PIM-SM). The dense option will place the component in Dense Mode (PIM-DM). By default, PIM component 1 is configured for Sparse Mode. PIM component 2 is unconfigured by default.

Note: A component using PIM-SM must also be configured with a dynamic or static Rendezvous Point (see "Specifying the Rendezvous Point" on page 413).

Defining an IP Interface for PIM Use

Each network attached to an IP interface on the switch may be assigned one of the available PIM components. The same PIM component can be assigned to multiple IP interfaces. The interfaces may belong to the same VLAN, and they may also belong to different VLANs as long as their member IP addresses do not overlap.

To define an IP interface for use with PIM, first configured the interface with an IPv4 address and VLAN as follows:

```
RS G8124-E(config)# interface ip <Interface number>
RS G8124-E(config-ip-if)# ip address <IPv4 address> <IPv4 mask>
RS G8124-E(config-ip-if)# vlan <VLAN number>
RS G8124-E(config-ip-if)# enable
```

Note: The PIM feature currently supports only one VLAN for each IP interface. Configurations where different interfaces on different VLANs share IP addresses are not supported.

Next, PIM must be enabled on the interface, and the PIM network component ID must be specified:

```
RS G8124-E(config-ip-if)# ip pim enable
RS G8124-E(config-ip-if)# ip pim component-id <1-2>
RS G8124-E(config-ip-if)# exit
```

By default, PIM component 1 is automatically assigned when PIM is enabled on the IP interface.

Note: While PIM is enabled on the interface, the interface VLAN cannot be changed. To change the VLAN, first disable PIM on the interface.

PIM Neighbor Filters

The G8124-E accepts connection to up to 72 PIM interfaces. By default, the switch accepts all PIM neighbors attached to the PIM-enabled interfaces, up to the maximum number. Once the maximum is reached, the switch will deny further PIM neighbors.

To ensure that only the appropriate PIM neighbors are accepted by the switch, the administrator can use PIM neighbor filters to specify which PIM neighbors may be accepted or denied on a per-interface basis.

To turn PIM neighbor filtering on or off for a particular IP interface, use the following commands:

```
RS G8124-E(config)# interface ip <Interface number>
RS G8124-E(config-ip-if)# [no] ip pim neighbor-filter
```

When filtering is enabled, all PIM neighbor requests on the specified IP interface will be denied by default. To allow a specific PIM neighbor, use the following command:

```
RS G8124-E(config-ip-if)# ip pim neighbor-addr <neighbor IPv4 address> allow
```

To remove a PIM neighbor from the accepted list, use the following command.

```
RS G8124-E(config-ip-if)# ip pim neighbor-addr <neighbor IPv4 address> deny
RS G8124-E(config-ip-if)# exit
```

You can view configured PIM neighbor filters globally or for a specific IP interface using the following commands:

```
RS G8124-E(config)# show ip pim neighbor-filters
RS G8124-E(config)# show ip pim interface <Interface number> neighbor-filters
```

Additional Sparse Mode Settings

Specifying the Rendezvous Point

Using PIM-SM, at least one PIM-capable router must be a candidate for use as a Rendezvous Point (RP) for any given multicast group. If desired, the G8124-E can act as an RP candidate. To assign a configured switch IP interface as a candidate, use the following procedure.

1. Select the PIM component that will represent the RP candidate:

```
RS G8124-E(config)# ip pim component <1-2>
```

2. Configure the IPv4 address of the switch interface which will be advertised as a candidate RP for the specified multicast group:

```
RS G8124-E(config-ip-pim-comp)# rp-candidate rp-address <group address>
    <group address mask> <candidate IPv4 address>
```

The switch interface will participate in the election of the RP that occurs on the Bootstrap Router, or BSR (see "Specifying a Bootstrap Router" on page 414).

3. If using dynamic RP candidates, configure the amount of time that the elected interface will remain the RP for the group before a re-election is performed:

```
RS G8124-E(config-ip-pim-comp)# rp-candidate holdtime <0-255>
RS G8124-E(config-ip-pim-comp)# exit
```

Static RP

If RP no election is desired, the switch can provide a static RP. Use the following commands:

1. Enable static RP configuration.

```
RS G8124-E(config)# ip pim static-rp enable
```

2. Select the PIM component that will represent the RP candidate:

```
RS G8124-E(config)# ip pim component <1-2>
```

3. Configure the static IPv4 address.

```
RS G8124-E(config-ip-pim-comp)# rp-static rp-address <group address>
    <group address mask> <static IPv4 address>
```

Influencing the Designated Router Selection

Using PIM-SM, All PIM-enabled IP interfaces are considered as potential Designate Routers (DR) for their domain. By default, the interface with the highest IP address on the domain is selected. However, if an interface is configured with a

DR priority value, it overrides the IP address selection process. If more than one interface on a domain is configured with a DR priority, the one with the highest number is selected.

Use the following commands to configure the DR priority value (Interface IP mode):

```
RS G8124-E(config)# interface ip <Interface number>
RS G8124-E(config-ip-if)# ip pim dr-priority <value (0-4294967294)>
RS G8124-E(config-ip-if)# exit
```

Note: A value of 0 (zero) specifies that the G8124-E will not act as the DR. This setting requires the G8124-E to be connected to a peer that has a DR priority setting of 1 or higher to ensure that a DR will be present in the network.

Specifying a Bootstrap Router

Using PIM-SM, a Bootstrap Router (BSR) is a PIM-capable router that hosts the election of the RP from available candidate routers. For each PIM-enabled IP interface, the administrator can set the preference level for which the local interface becomes the BSR:

```
RS G8124-E(config)# interface ip <Interface number>
RS G8124-E(config-ip-if)# ip pim cbsr-preference <0 to 255>
RS G8124-E(config-ip-if)# exit
```

A value of 255 highly prefers the local interface as a BSR. A value of -1 indicates that the PIM CBSR preference is not configured on the switch.

Configuring a Loopback Interface

Loopback interfaces can be used in PIM Sparse Mode for Rendezvous Points (RPs) and Bootstrap Routers (BSRs). For example:

● As a static RP

```
interface loopback 1
   ip address 55.55.1.1 255.255.255.0
   enable
   exit

ip pim static-rp enable

ip pim component 1
   rp-static rp-address 224.0.0.0 240.0.0.0 55.55.1.1

interface loopback 1
   ip pim enable
   exit
```

- As a candidate RP

```
interface loopback 1
   ip address 55.55.1.1 255.255.255.0
   enable
   exit

ip pim component 1
   rp-candidate holdtime 60
   rp-candidate rp-address 224.0.0.0 240.0.0.0 55.55.1.1

interface loopback 1
   ip pim enable
   exit
```

- As a BSR

```
interface loopback 1
   ip address 55.55.1.1 255.255.255.0
   enable
   exit

interface loopback 1
   ip pim enable
   ip pim cbsr-preference 2
   exit
```

Using PIM with Other Features

PIM with ACLs or VMAPs

> If using ACLs or VMAPs, be sure to permit traffic for local hosts and routers.

PIM with IGMP

> If using IGMP (see Chapter 24, "Internet Group Management Protocol"):
>
> ● IGMP static joins can be configured with a PIM-SM or PIM-DM multicast group IPv4 address. Using the ISCLI:

```
RS G8124-E(config)# ip mroute <multicast group IPv4 address> <VLAN> <port>
```

> ● IGMP Querier is disabled by default. If IGMP Querier is needed with PIM, be sure to enable the IGMP Query feature globally, as well as on each VLAN where it is needed.
>
> ● If the switch is connected to multicast receivers and/or hosts, be sure to enable IGMP snooping globally, as well as on each VLAN where PIM receivers are attached.

PIM with VLAG

> If using VLAG, see "VLAG with PIM" on page 185.

PIM Configuration Examples

Example 1: PIM-SM with Dynamic RP

This example configures PIM Sparse Mode for one IP interface, with the switch acting as a candidate for dynamic Rendezvous Point (RP) selection.

1. Globally enable the PIM feature:

```
RS G8124-E(config)# ip pim enable
```

2. Configure a PIM network component with dynamic RP settings, and set it for PIM Sparse Mode:

```
RS G8124-E(config)# ip pim component 1
RS G8124-E(config-ip-pim-comp)# mode sparse
RS G8124-E(config-ip-pim-comp)# rp-candidate rp-address 225.1.0.0
    255.255.0.0 10.10.1.1
RS G8124-E(config-ip-pim-comp)# rp-candidate holdtime <0-255>
RS G8124-E(config-ip-pim-comp)# exit
```

Where 225.1.0.0 is the multicast group base IP address, 255.255.0.0 is the multicast group address mask, and 10.10.1.1 is the switch RP candidate address.

Note: Because, Sparse Mode is set by default for PIM component 1, the mode command is needed only if the mode has been previously changed.

3. Define an IP interface for use with PIM:

```
RS G8124-E(config)# interface ip 111
RS G8124-E(config-ip-if)# ip address 10.10.1.1 255.255.255.255
RS G8124-E(config-ip-if)# vlan 11
RS G8124-E(config-ip-if)# enable
```

The IP interface represents the PIM network being connected to the switch. The IPv4 addresses in the defined range must not be included in another IP interface on the switch under a different VLAN.

4. Enable PIM on the IP interface and assign the PIM component:

```
RS G8124-E(config-ip-if)# ip pim enable
RS G8124-E(config-ip-if)# ip pim component-id 1
```

Note: Because, PIM component 1 is assigned to the interface by default, the component-id command is needed only if the setting has been previously changed.

5. Set the Bootstrap Router (BSR) preference:

```
RS G8124-E(config-ip-if)# ip pim cbsr-preference 135
RS G8124-E(config-ip-if)# exit
```

Example 2: PIM-SM with Static RP

The following commands can be used to modify the prior example configuration to use a static RP:

```
RS G8124-E(config)# ip pim static-rp enable
RS G8124-E(config)# ip pim component 1
RS G8124-E(config-ip-pim-comp)# rp-static rp-address 225.1.0.0
    255.255.0.0 10.10.1.1
RS G8124-E(config-ip-pim-comp)# exit
```

Where 225.1.0.0 255.255.0.0 is the multicast group base address and mask, and 10.10.1.1 is the static RP address.

Note: The same static RP address must be configured for all switches in the group.

Example 3: PIM-DM

This example configures PIM Dense Mode (PIM-DM) on one IP interface. PIM-DM can be configured independently, or it can be combined with the prior PIM-SM examples (which are configured on a different PIM component) as shown in Figure 48.

Figure 48. Network with both PIM-DM and PIM-SM Components
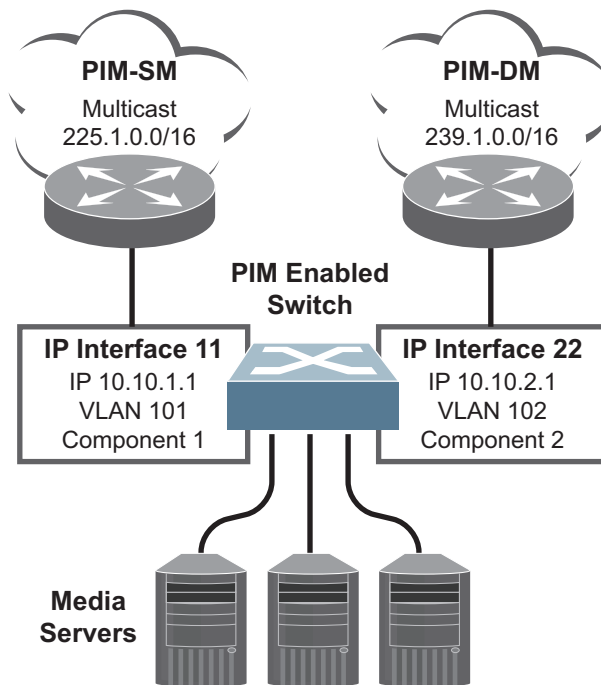
1. Configure the PIM-SM component as shown in the prior examples, or if using PIM-DM independently, enable the PIM feature.

```
RS G8124-E(config)# ip pim enable
```

2. Configure a PIM component and set the PIM mode:

```
RS G8124-E(config)# ip pim component 2
RS G8124-E(config-ip-pim-comp)# mode dense
RS G8124-E(config-ip-pim-comp)# exit
```

3. Define an IP interface for use with PIM:

```
RS G8124-E(config)# interface ip 22
RS G8124-E(config-ip-if)# ip address 10.10.2.1 255.255.255.255
RS G8124-E(config-ip-if)# vlan 102
RS G8124-E(config-ip-if)# enable
```

4. Enable PIM on the IP interface and assign the PIM component:

```
RS G8124-E(config-ip-if)# ip pim enable
RS G8124-E(config-ip-if)# ip pim component-id 2
RS G8124-E(config-ip-if)# exit
```

5. (Optional) Configure PIM border router if the IPMC traffic is flowing between PIM domains:

```
RS G8124-E(config)# ip pim pmbr enable
RS G8124-E(config)# interface ip 22
RS G8124-E(config-ip-if)# ip pim border-bit
RS G8124-E(config-ip-if)# exit
```

Note: For PIM Dense Mode, the DR, RP, and BSR settings do not apply.

Part 6: High Availability Fundamentals

Internet traffic consists of myriad services and applications which use the Internet Protocol (IP) for data delivery. However, IP is not optimized for all the various applications. High Availability goes beyond IP and makes intelligent switching decisions to provide redundant network configurations.

Chapter 29. Basic Redundancy

Lenovo Networking OS 8.3 includes various features for providing basic link or device redundancy:

-
-
-

Aggregating for Link Redundancy

Multiple switch ports can be combined together to form robust, high-bandwidth LAGs to other devices. Since LAGs are comprised of multiple physical links, the LAG is inherently fault tolerant. As long as one connection between the switches is available, the LAG remains active.

In Figure 49, four ports are aggregated together between the switch and the enterprise routing device. Connectivity is maintained as long as one of the links remain active. The links to the server are also aggregated, allowing the secondary NIC to take over in the event that the primary NIC link fails.

Figure 49. Aggregating Ports for Link Redundancy
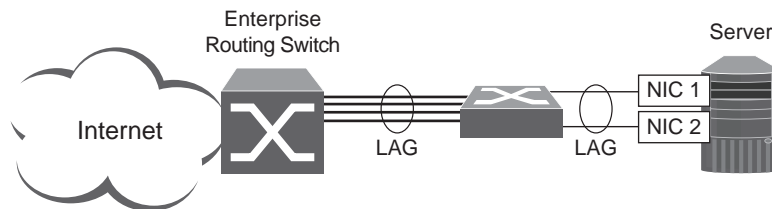
For more information on aggregation, see Chapter 8, "Ports and Link Aggregation."

Virtual Link Aggregation

Using the VLAG feature, switches can be paired as VLAG peers. The peer switches appear to the connecting device as a single virtual entity for the purpose of establishing a multi-port LAG. The VLAG-capable switches synchronize their logical view of the access layer port structure and internally prevent implicit loops. The VLAG topology also responds more quickly to link failure and does not result in unnecessary MAC flooding.

VLAGs are useful in multi-layer environments for both uplink and downlink redundancy to any regular LAG-capable device. They can also be used in for active-active VRRP connections.

For more information on VLAGs, see Chapter 10, "Virtual Link Aggregation Groups."

Hot Links

Hot Links provides basic link redundancy with fast recovery.

Hot Links consists of up to 25 triggers. A trigger consists of a pair of layer 2 interfaces, each containing an individual port, LAG, or LACP adminkey. One interface is the Master, and the other is a Backup. While the Master interface is set to the active state and forwards traffic, the Backup interface is set to the standby state and blocks traffic until the Master interface fails. If the Master interface fails, the Backup interface is set to active and forwards traffic. Once the Master interface is restored, it regains its active state and forwards traffic. The Backup transitions to the standby state.

You may select a physical port, static LAG, or an LACP adminkey as a Hot Link interface.

Forward Delay

The Forward Delay timer allows Hot Links to monitor the Master and Backup interfaces for link stability before selecting one interface to transition to the active state. Before the transition occurs, the interface must maintain a stable link for the duration of the Forward Delay interval.

For example, if you set the Forward delay timer to 10 seconds, the switch will select an interface to become active only if a link remained stable for the duration of the Forward Delay period. If the link is unstable, the Forward Delay period starts again.

Preemption

You can configure the Master interface to resume the active state whenever it becomes available. With Hot Links preemption enabled, the Master interface transitions to the active state immediately upon recovery. The Backup interface immediately transitions to the standby state. If Forward Delay is enabled, the transition occurs when an interface has maintained link stability for the duration of the Forward Delay period.

FDB Update

Use the FDB update option to notify other devices on the network about updates to the Forwarding Database (FDB). When you enable FDB update, the switch sends multicasts of addresses in the forwarding database (FDB) over the active interface, so that other devices on the network can learn the new path. The Hot Links FBD update option uses the station update rate to determine the rate at which to send FDB packets.

Configuration Guidelines

The following configuration guidelines apply to Hot links:

- When Hot Links is turned on, STP must be disabled on the hotlink ports.

- A port that is a member of the Master interface cannot be a member of the Backup interface. A port that is a member of one Hot Links trigger cannot be a member of another Hot Links trigger.
- An individual port that is configured as a Hot Link interface cannot be a member of a LAG.

Configuring Hot Links

Spanning Tree Protocol must be disabled on Hot Links ports. For instance, if Hot Links ports 1 and 2 belong to STGs 1 and 23, use the following commands to disable the STGs:

```
RS G8124-E(config)# no span stp 1 enable
RS G8124-E(config)# no span stp 23 enable
```

Use the following commands to configure Hot Links.

```
RS G8124-E(config)# [no] hotlinks bpdu    (Enable or disable Hot Links BPDU flood)
RS G8124-E(config)# [no] hotlinks enable    (Enable or disable Hot Links globally)
RS G8124-E(config)# [no] hotlinks fdb-update    (Enable or disable Hot Links FDB
update)
RS G8124-E(config)# hotlinks fdb-update-rate    (Configure FDB update rate in packets
per second)
RS G8124-E(config)# hotlinks trigger 1 enable    (Enable Hot Links Trigger 1)
RS G8124-E(config)# hotlinks trigger 1 master port 1 (Add port to Master interface)
RS G8124-E(config)# hotlinks trigger 1 backup port 2 (Add port to Backup interface)
RS G8124-E(config)# hotlinks enable      (Turn on Hot Links)
```

Chapter 30. Layer 2 Failover

The primary application for Layer 2 Failover is to support Network Adapter Teaming. With Network Adapter Teaming, all the NICs on each server share the same IP address, and are configured into a team. One NIC is the primary link, and the other is a standby link. For more details, refer to the documentation for your Ethernet adapter.

Note: Only two links per server can be used for Layer 2 LAG Failover (one primary and one backup). Network Adapter Teaming allows only one backup NIC for each server blade.

Monitoring LAG Links

Layer 2 Failover can be enabled on any LAG in the G8124-E, including LACP LAGs. LAGs can be added to failover trigger groups. Then, if some specified number of monitor links fail, the switch disables all the control ports in the switch. When the control ports are disabled, it causes the NIC team on the affected servers to failover from the primary to the backup NIC. This process is called a failover event.

When the appropriate number of links in a monitor group return to service, the switch enables the control ports. This causes the NIC team on the affected servers to fail back to the primary switch (unless Auto-Fallback is disabled on the NIC team). The backup switch processes traffic until the primary switch's control links come up, which can take up to five seconds.

Figure 50 is a simple example of Layer 2 Failover. One G8124-E is the primary, and the other is used as a backup. In this example, all ports on the primary switch belong to a single LAG, with Layer 2 Failover enabled, and Failover Limit set to 2. If two or fewer links in trigger 1 remain active, the switch temporarily disables all control ports. This action causes a failover event on Server 1 and Server 2.

Figure 50. Basic Layer 2 Failover
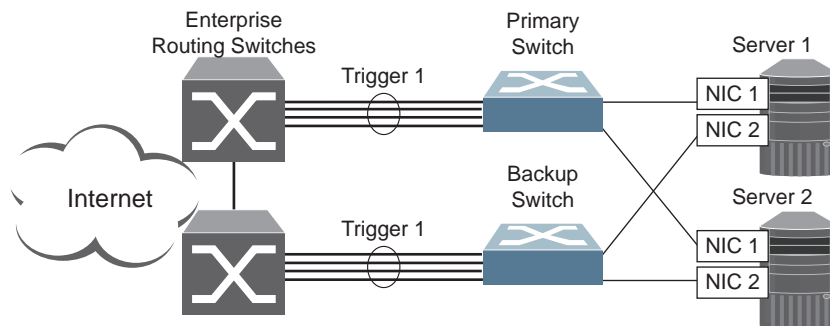
Setting the Failover Limit

The failover limit lets you specify the minimum number of operational links required within each trigger before the trigger initiates a failover event. For example, if the limit is two, a failover event occurs when the number of operational links in the trigger is two or fewer. When you set the limit to zero, the switch triggers a failover event only when no links in the trigger are operational.

Manually Monitoring Port Links

The Manual Monitor allows you to configure a set of ports and LAGs to monitor for link failures (a monitor list), and another set of ports and LAGs to disable when the trigger limit is reached (a control list). When the switch detects a link failure on the monitor list, it automatically disables the items in control list. When server ports are disabled, the corresponding server's network adapter can detect the disabled link, and trigger a network-adapter failover to another port or LAG on the switch, or another switch.

The switch automatically enables the control list items when the monitor list items return to service.

Monitor Port State

A monitor port is considered operational as long as the following conditions are true:

● The port must be in the `Link Up` state.
● If STP is enabled, the port must be in the `Forwarding` state.
● If the port is part of an LACP LAG, the port must be in the `Aggregated` state.

If any of these conditions is false, the monitor port is considered to have failed.

Control Port State

A control port is considered Operational if the monitor trigger is up. As long as the trigger is up, the port is considered operational from a teaming perspective, even if the port itself is actually in the `Down` state, `Blocking` state (if STP is enabled on the port), or `Not Aggregated` state (if part of an LACP LAG).

A control port is considered to have failed only if the monitor trigger is in the `Down` state.

To view the state of any port, use one of the following commands:

```
>> # show interface link                              (View port link status)
>> # show interface port <x> spanning-tree stp <x>    (View port STP status)
>> # show lacp information                             (View port LACP status)
```

L2 Failover with Other Features

L2 Failover works together with Link Aggregation Control Protocol (LACP) and with Spanning Tree Protocol (STP), as described in the next sections.

LACP

Link Aggregation Control Protocol allows the switch to form dynamic LAGs. You can use the *admin key* to add up to LACP LAGs to a failover trigger using monitoring. When you add an *admin key* to a trigger, any LACP LAG with that *admin key* becomes a member of the trigger.

Spanning Tree Protocol

If Spanning Tree Protocol (STP) is enabled on the ports in a failover trigger, the switch monitors the port STP state rather than the link state. A port failure results when STP is not in a Forwarding state (such as Learning, Discarding, or No Link) in all the Spanning Tree Groups (STGs) to which the port belongs. The switch automatically disables the appropriate control ports.

When the switch determines that ports in the trigger are in STP Forwarding state in any one of the STGs it belongs to, then it automatically enables the appropriate control ports. The switch *fails back* to normal operation.

For example, if a monitor port is a member of STG1, STG2, and STG3, a failover will be triggered only if the port is not in a forwarding state in all the three STGs. When the port state in any of the three STGs changes to forwarding, then the control port is enabled and normal switch operation is resumed.

Configuration Guidelines

This section provides important information about configuring Layer 2 Failover.

- Any specific failover trigger can monitor ports only, static LAGs only, or LACP LAGs only. The different types cannot be combined in the same trigger.

- A maximum of 24 LACP keys can be added per trigger.

- Management ports cannot be monitored.

- Control ports for different triggers must not overlap. Monitor ports may overlap.

Configuring Layer 2 Failover

Use the following procedure to configure a Layer 2 Failover Manual Monitor.

1. Specify the links to monitor.

```
>> # failover trigger 1 mmon monitor member 1-5
```

2. Specify the links to disable when the failover limit is reached.

```
>> # failover trigger 1 mmon control member 6-10
```

3. Configure general Failover parameters.

```
>> # failover enable
>> # failover trigger 1 enable
>> # failover trigger 1 limit 2
```

Chapter 31. Virtual Router Redundancy Protocol

The Lenovo RackSwitch G8124-E (G8124-E) supports IPv4 high-availability network topologies through an enhanced implementation of the Virtual Router Redundancy Protocol (VRRP).

Note: Lenovo Networking OS 8.3 does not support IPv6 for VRRP.

The following topics are discussed in this chapter:

- "VRRP Overview" on page 434. This section discusses VRRP operation and Lenovo N/OS redundancy configurations.

- "Failover Methods" on page 437. This section describes the three modes of high availability.

- "Lenovo N/OS Extensions to VRRP" on page 438. This section describes VRRP enhancements implemented in N/OS.

- "Virtual Router Deployment Considerations" on page 439. This section describes issues to consider when deploying virtual routers.

- "High Availability Configurations" on page 440. This section discusses the more useful and easily deployed redundant configurations.

VRRP Overview

In a high-availability network topology, no device can create a single point-of-failure for the network or force a single point-of-failure to any other part of the network. This means that your network will remain in service despite the failure of any single device. To achieve this usually requires redundancy for all vital network components.

VRRP enables redundant router configurations within a LAN, providing alternate router paths for a host to eliminate single points-of-failure within a network. Each participating VRRP-capable routing device is configured with the same virtual router IPv4 address and ID number. One of the virtual routers is elected as the master, based on a number of priority criteria, and assumes control of the shared virtual router IPv4 address. If the master fails, one of the backup virtual routers will take control of the virtual router IPv4 address and actively process traffic addressed to it.

With VRRP, Virtual Interface Routers (VIR) allow two VRRP routers to share an IP interface across the routers. VIRs provide a single Destination IPv4 (DIP) address for upstream routers to reach various servers, and provide a virtual default Gateway for the servers.

VRRP Components

Each physical router running VRRP is known as a *VRRP router*.

Virtual Router

Two or more VRRP routers can be configured to form a *virtual router* (RFC 2338). Each VRRP router may participate in one or more virtual routers. Each virtual router consists of a user-configured *virtual router identifier* (VRID) and an IPv4 address.

Virtual Router MAC Address

The VRID is used to build the *virtual router MAC Address*. The five highest-order octets of the virtual router MAC Address are the standard MAC prefix (00-00-5E-00-01) defined in RFC 2338. The VRID is used to form the lowest-order octet.

Owners and Renters

Only one of the VRRP routers in a virtual router may be configured as the IPv4 address owner. This router has the virtual router's IPv4 address as its real interface address. This router responds to packets addressed to the virtual router's IPv4 address for ICMP pings, TCP connections, and so on.

There is no requirement for any VRRP router to be the IPv4 address owner. Most VRRP installations choose not to implement an IPv4 address owner. For the purposes of this chapter, VRRP routers that are not the IPv4 address owner are called *renters*.

Master and Backup Virtual Router

Within each virtual router, one VRRP router is selected to be the virtual router master. See "Selecting the Master VRRP Router" on page 436 for an explanation of the selection process.

Note: If the IPv4 address owner is available, it will always become the virtual router master.

The virtual router master forwards packets sent to the virtual router. It also responds to Address Resolution Protocol (ARP) requests sent to the virtual router's IPv4 address. Finally, the virtual router master sends out periodic advertisements to let other VRRP routers know it is alive and its priority.

Within a virtual router, the VRRP routers not selected to be the master are known as virtual router backups. If the virtual router master fails, one of the virtual router backups becomes the master and assumes its responsibilities.

Virtual Interface Router

At Layer 3, a Virtual Interface Router (VIR) allows two VRRP routers to share an IP interface across the routers. VIRs provide a single Destination IPv4 (DIP) address for upstream routers to reach various destination networks, and provide a virtual default Gateway.

Note: Every VIR must be assigned to an IP interface, and every IP interface must be assigned to a VLAN. If no port in a VLAN has link up, the IP interface of that VLAN is down, and if the IP interface of a VIR is down, that VIR goes into INIT state.

VRRP Operation

Only the virtual router master responds to ARP requests. Therefore, the upstream routers only forward packets destined to the master. The master also responds to ICMP ping requests. The backup does not forward any traffic, nor does it respond to ARP requests.

If the master is not available, the backup becomes the master and takes over responsibility for packet forwarding and responding to ARP requests.

Selecting the Master VRRP Router

Each VRRP router is configured with a priority between 1–254. A bidding process determines which VRRP router is or becomes the master—the VRRP router with the highest priority.

The master periodically sends advertisements to an IPv4 multicast address. As long as the backups receive these advertisements, they remain in the backup state. If a backup does not receive an advertisement for three advertisement intervals, it initiates a bidding process to determine which VRRP router has the highest priority and takes over as master. In addition to the three advertisement intervals, a manually set holdoff time can further delay the backups from assuming the master status.

If, at any time, a backup determines that it has higher priority than the current master does, it can preempt the master and become the master itself, unless configured not to do so. In preemption, the backup assumes the role of master and begins to send its own advertisements. The current master sees that the backup has higher priority and will stop functioning as the master.

A backup router can stop receiving advertisements for one of two reasons—the master can be down, or all communications links between the master and the backup can be down. If the master has failed, it is clearly desirable for the backup (or one of the backups, if there is more than one) to become the master.

Note: If the master is healthy but communication between the master and the backup has failed, there will then be two masters within the virtual router. To prevent this from happening, configure redundant links to be used between the switches that form a virtual router.

Failover Methods

With service availability becoming a major concern on the Internet, service providers are increasingly deploying Internet traffic control devices, such as application switches, in redundant configurations. N/OS high availability configurations are based on VRRP. The N/OS implementation of VRRP includes proprietary extensions.

Active-Active Redundancy

In an active-active configuration, shown in Figure 51, two switches provide redundancy for each other, with both active at the same time. Each switch processes traffic on a different subnet. When a failure occurs, the remaining switch can process traffic on all subnets.

For a configuration example, see "High Availability Configurations" on page 440.

Figure 51. Active-Active Redundancy
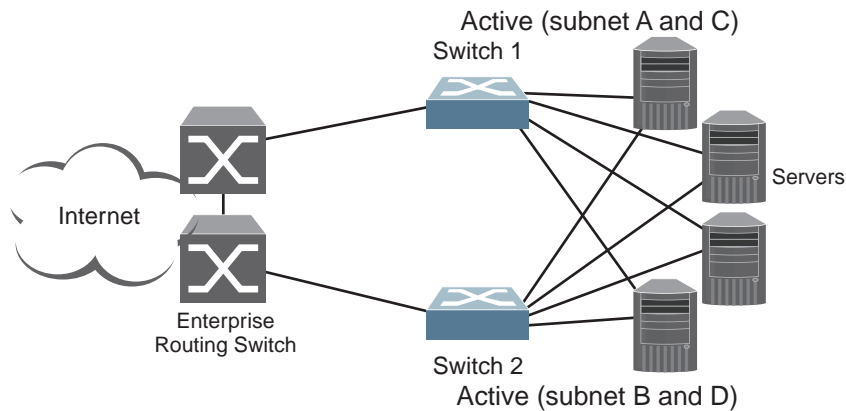
Virtual Router Group

The virtual router group ties all virtual routers on the switch together as a single entity. As members of a group, all virtual routers on the switch (and therefore the switch itself), are in either a master or standby state.

A VRRP group has the following characteristics:

- When enabled, all virtual routers behave as one entity, and all group settings override any individual virtual router settings.

- All individual virtual routers, once the VRRP group is enabled, assume the group's tracking and priority.

- When one member of a VRRP group fails, the priority of the group decreases, and the state of the entire switch changes from Master to Standby.

- When VRRP group restrictions are enabled, advertisements for the group are sent and received only on VRRP group interface.

Each VRRP advertisement can include up to 16 addresses. All virtual routers are advertised within the same packet, conserving processing and buffering resources.

Lenovo N/OS Extensions to VRRP

This section describes VRRP enhancements that are implemented in N/OS.

N/OS supports a tracking function that dynamically modifies the priority of a VRRP router, based on its current state. The objective of tracking is to have, whenever possible, the master bidding processes for various virtual routers in a LAN converge on the same switch. Tracking ensures that the selected switch is the one that offers optimal network performance. For tracking to have any effect on virtual router operation, preemption must be enabled.

N/OS can track the attributes listed in Table 31 (Router VRRP mode):

Table 31. *VRRP Tracking Parameters*

Parameter	Description
Number of IP interfaces on the switch that are active ("up") `tracking-priority-increment interfaces`	Helps elect the virtual routers with the most available routes as the master. (An IP interface is considered active when there is at least one active port on the same VLAN.) This parameter influences the VRRP router's priority in virtual interface routers.
Number of active ports on the same VLAN `tracking-priority-increment ports`	Helps elect the virtual routers with the most available ports as the master. This parameter influences the VRRP router's priority in virtual interface routers.
Number of virtual routers in master mode on the switch `tracking-priority-increment virtual-routers`	Useful for ensuring that traffic for any particular client/server pair is handled by the same switch, increasing routing efficiency. This parameter influences the VRRP router's priority in virtual interface routers.

Each tracked parameter has a user-configurable weight associated with it. As the count associated with each tracked item increases (or decreases), so does the VRRP router's priority, subject to the weighting associated with each tracked item. If the priority level of a standby is greater than that of the current master, then the standby can assume the role of the master.

See "Configuring the Switch for Tracking" on page 439 for an example on how to configure the switch for tracking VRRP priority.

Virtual Router Deployment Considerations

Assigning VRRP Virtual Router ID

During the software upgrade process, VRRP virtual router IDs will be automatically assigned if failover is enabled on the switch. When configuring virtual routers at any point after upgrade, virtual router ID numbers must be assigned. The virtual router ID may be configured as any number between 1 and 255. Use the following command to configure the virtual router ID:

```
RS G8124-E(config)# router vrrp
RS G8124-E(config-vrrp)# virtual-router 1 virtual-router-id <1-255>
```

Configuring the Switch for Tracking

Tracking configuration largely depends on user preferences and network environment. Consider the configuration shown in Figure 51 on page 437. Assume the following behavior on the network:

- Switch 1 is the master router upon initialization.

- If switch 1 is the master and it has one fewer active servers than switch 2, then switch 1 remains the master.

  This behavior is preferred because running one server down is less disruptive than bringing a new master online and severing all active connections in the process.

- If switch 1 is the master and it has two or more active servers fewer than switch 2, then switch 2 becomes the master.

- If switch 2 is the master, it remains the master even if servers are restored on switch 1 such that it has one fewer or an equal number of servers.

- If switch 2 is the master and it has one active server fewer than switch 1, then switch 1 becomes the master.

You can implement this behavior by configuring the switch for tracking as follows:

1. Set the priority for switch 1 to 101.

2. Leave the priority for switch 2 at the default value of 100.

3. On both switches, enable tracking based on ports, interfacesor virtual routers. You can choose any combination of tracking parameters, based on your network configuration.

   Note: There is no shortcut to setting tracking parameters. The goals must first be set and the outcomes of various configurations and scenarios analyzed to find settings that meet the goals.

High Availability Configurations

VRRP High-Availability Using Multiple VIRs

Figure 52 shows an example configuration where two G8124-Es are used as VRRP routers in an active-active configuration. In this configuration, both switches respond to packets.

Figure 52. Active-Active Configuration using VRRP
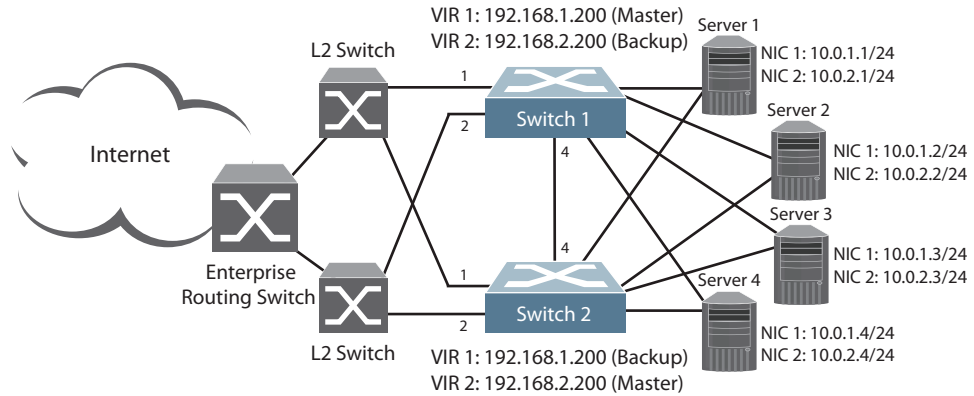
Although this example shows only two switches, there is no limit on the number of switches used in a redundant configuration. It is possible to implement an active-active configuration across all the VRRP-capable switches in a LAN.

Each VRRP-capable switch in an active-active configuration is autonomous. Switches in a virtual router need not be identically configured.

In the scenario illustrated in Figure 52, traffic destined for IPv4 address 10.0.1.1 is forwarded through the Layer 2 switch at the top of the drawing, and ingresses G8124-E 1 on port 1. Return traffic uses default gateway 1 (192.168.1.1).

If the link between G8124-E 1 and the Layer 2 switch fails, G8124-E 2 becomes the Master because it has a higher priority. Traffic is forwarded to G8124-E 2, which forwards it to G8124-E 1 through port 4. Return traffic uses default gateway 2 (192.168.2.1), and is forwarded through the Layer 2 switch at the bottom of the drawing.

To implement the active-active example, perform the following switch configuration.

Task 1: Configure G8124-E 1

1. Configure client and server interfaces.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip address 192.168.1.100 255.255.255.0
RS G8124-E(config-ip-if)# vlan 10
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip address 192.168.2.101 255.255.255.0
RS G8124-E(config-ip-if)# vlan 20
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 3
RS G8124-E(config-ip-if)# ip address 10.0.1.100 255.255.255.0
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 4
RS G8124-E(config-ip-if)# ip address 10.0.2.101 255.255.255.0
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

2. Configure the default gateways. Each default gateway points to a Layer 3 router.

```
RS G8124-E(config)# ip gateway 1 address 192.168.1.1
RS G8124-E(config)# ip gateway 1 enable
RS G8124-E(config)# ip gateway 2 address 192.168.2.1
RS G8124-E(config)# ip gateway 2 enable
```

3. Turn on VRRP and configure two Virtual Interface Routers.

```
RS G8124-E(config)# router vrrp
RS G8124-E(config-vrrp)# enable
RS G8124-E(config-vrrp)# virtual-router 1 virtual-router-id 1
RS G8124-E(config-vrrp)# virtual-router 1 interface 1
RS G8124-E(config-vrrp)# virtual-router 1 address 192.168.1.200
RS G8124-E(config-vrrp)# virtual-router 1 enable
RS G8124-E(config-vrrp)# virtual-router 2 virtual-router-id 2
RS G8124-E(config-vrrp)# virtual-router 2 interface 2
RS G8124-E(config-vrrp)# virtual-router 2 address 192.168.2.200
RS G8124-E(config-vrrp)# virtual-router 2 enable
```

4. Enable tracking on ports. Set the priority of Virtual Router 1 to 101, so that it becomes the Master.

```
RS G8124-E(config-vrrp)# virtual-router 1 track ports
RS G8124-E(config-vrrp)# virtual-router 1 priority 101
RS G8124-E(config-vrrp)# virtual-router 2 track ports
RS G8124-E(config-vrrp)# exit
```

5. Configure ports.

```
RS G8124-E(config)# vlan 10
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 10
RS G8124-E(config-if)# exit

RS G8124-E(config)# vlan 20
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 2
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 20
RS G8124-E(config-if)# exit
```

6. Turn off Spanning Tree Protocol globally.

```
RS G8124-E(config)# no spanning-tree stp 1
```

Task 2: Configure G8124-E 2

1. Configure client and server interfaces.

```
RS G8124-E(config)# interface ip 1
RS G8124-E(config-ip-if)# ip address 192.168.1.101 255.255.255.0
RS G8124-E(config-ip-if)# vlan 10
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 2
RS G8124-E(config-ip-if)# ip address 192.168.2.100 255.255.255.0
RS G8124-E(config-ip-if)# vlan 20
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 3
RS G8124-E(config-ip-if)# ip address 10.0.1.101 255.255.255.0
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
RS G8124-E(config)# interface ip 4
RS G8124-E(config-ip-if)# ip address 10.0.2.100 255.255.255.0
RS G8124-E(config-ip-if)# enable
RS G8124-E(config-ip-if)# exit
```

2. Configure the default gateways. Each default gateway points to a Layer 3 router.

```
RS G8124-E(config)# ip gateway 1 address 192.168.2.1
RS G8124-E(config)# ip gateway 1 enable
RS G8124-E(config)# ip gateway 2 address 192.168.1.1
RS G8124-E(config)# ip gateway 2 enable
```

3. Turn on VRRP and configure two Virtual Interface Routers.

```
RS G8124-E(config)# router vrrp
RS G8124-E(config-vrrp)# enable
RS G8124-E(config-vrrp)# virtual-router 1 virtual-router-id 1
RS G8124-E(config-vrrp)# virtual-router 1 interface 1
RS G8124-E(config-vrrp)# virtual-router 1 address 192.168.1.200
RS G8124-E(config-vrrp)# virtual-router 1 enable
RS G8124-E(config-vrrp)# virtual-router 2 virtual-router-id 2
RS G8124-E(config-vrrp)# virtual-router 2 interface 2
RS G8124-E(config-vrrp)# virtual-router 2 address 192.168.2.200
RS G8124-E(config-vrrp)# virtual-router 2 enable
```

4. Enable tracking on ports. Set the priority of Virtual Router 2 to 101, so that it becomes the Master.

```
RS G8124-E(config-vrrp)# virtual-router 1 track ports
RS G8124-E(config-vrrp)# virtual-router 2 track ports
RS G8124-E(config-vrrp)# virtual-router 2 priority 101
RS G8124-E(config-vrrp)# exit
```

5. Configure ports.

```
RS G8124-E(config)# vlan 10
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 10
RS G8124-E(config-if)# exit

RS G8124-E(config)# vlan 20
RS G8124-E(config-vlan)# exit
RS G8124-E(config)# interface port 2
RS G8124-E(config-if)# switchport mode trunk
RS G8124-E(config-if)# switchport trunk allowed vlan 20
RS G8124-E(config-if)# exit
```

6. Turn off Spanning Tree Protocol globally.

```
RS G8124-E(config)# no spanning-tree stp 1
```

VRRP High-Availability Using VLAGs

VRRP can be used in conjunction with VLAGs and LACP-capable servers and switches to provide seamless redundancy.

Figure 53. Active-Active Configuration using VRRP and VLAGs
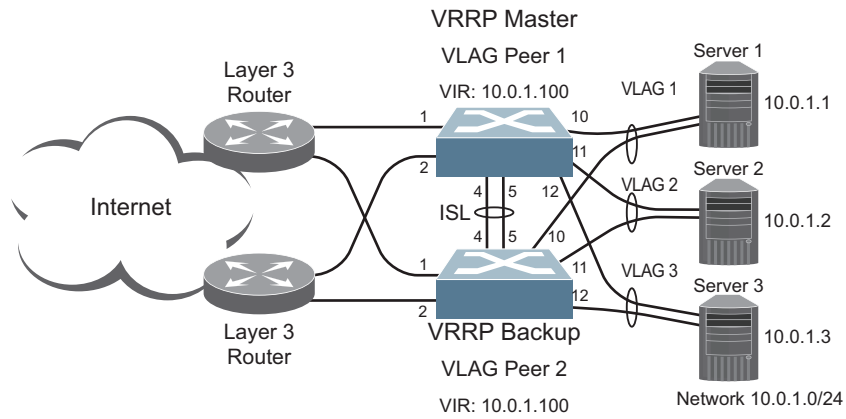
See "VLAGs with VRRP" on page 177 for a detailed configuration example.

Part 7: Network Management

Chapter 32. Link Layer Discovery Protocol

The Lenovo Networking OS software support Link Layer Discovery Protocol (LLDP). This chapter discusses the use and configuration of LLDP on the switch:

- "LLDP Overview" on page 448
- "Enabling or Disabling LLDP" on page 449
- "LLDP Transmit Features" on page 450
- "LLDP Receive Features" on page 454
- "LLDP Example Configuration" on page 458

LLDP Overview

Link Layer Discovery Protocol (LLDP) is an IEEE 802.1AB-2005 standard for discovering and managing network devices. LLDP uses Layer 2 (the data link layer), and allows network management applications to extend their awareness of the network by discovering devices that are direct neighbors of already known devices.

With LLDP, the G8124-E can advertise the presence of its ports, their major capabilities, and their current status to other LLDP stations in the same LAN. LLDP transmissions occur on ports at regular intervals or whenever there is a relevant change to their status. The switch can also receive LLDP information advertised from adjacent LLDP-capable network devices.

In addition to discovery of network resources, and notification of network changes, LLDP can help administrators quickly recognize a variety of common network configuration problems, such as unintended VLAN exclusions or mis-matched port aggregation membership.

The LLDP transmit function and receive function can be independently configured on a per-port basis. The administrator can allow any given port to transmit only, receive only, or both transmit and receive LLDP information.

The LLDP information to be distributed by the G8124-E ports, and that which has been collected from other LLDP stations, is stored in the switch's Management Information Base (MIB). Network Management Systems (NMS) can use Simple Network Management Protocol (SNMP) to access this MIB information. LLDP-related MIB information is read-only.

Changes, either to the local switch LLDP information or to the remotely received LLDP information, are flagged within the MIB for convenient tracking by SNMP-based management systems.

For LLDP to provide expected benefits, all network devices that support LLDP must be consistent in their LLDP configuration.

Enabling or Disabling LLDP

Global LLDP Setting

By default, LLDP is enabled on the G8124-E. To turn LLDP on or off, use the following command:

```
RS G8124-E(config)# [no] lldp enable   (Turn LLDP on or off globally)
```

Transmit and Receive Control

The G8124-E can also be configured to transmit or receive LLDP information on a port-by-port basis. By default, when LLDP is globally enabled on the switch, G8124-E ports transmit and receive LLDP information (see the tx_rx option in the following example). To change the LLDP transmit and receive state, the following commands are available:

```
RS G8124-E(config)# interface port 1   (Select a switch port)
RS G8124-E(config-if)# lldp admin-status tx_rx(Transmit and receive LLDP)
RS G8124-E(config-if)# lldp admin-status tx_only(Only transmit LLDP)
RS G8124-E(config-if)# lldp admin-status rx_only(Only receive LLDP)
RS G8124-E(config-if)# no lldp admin-status(Do not participate in LLDP)
RS G8124-E(config-if)# exit              (Exit port mode)
```

To view the LLDP transmit and receive status, use the following commands:

```
RS G8124-E(config)# show lldp port       (status of all ports)
RS G8124-E(config)# show interface port <n> lldp(status of selected port)
```

LLDP Transmit Features

Numerous LLDP transmit options are available, including scheduled and minimum transmit interval, expiration on remote systems, SNMP trap notification, and the types of information permitted to be shared.

Scheduled Interval

The G8124-E can be configured to transmit LLDP information to neighboring devices once each 5 to 32768 seconds. The scheduled interval is global; the same interval value applies to all LLDP transmit-enabled ports. However, to help balance LLDP transmissions and keep them from being sent simultaneously on all ports, each port maintains its own interval clock, based on its own initialization or reset time. This allows switch-wide LLDP transmissions to be spread out over time, though individual ports comply with the configured interval.

The global transmit interval can be configured using the following command:

```
RS G8124-E(config)# lldp refresh-interval <interval>
```

where *interval* is the number of seconds between LLDP transmissions. The range is 5 to 32768. The default is 30 seconds.

Minimum Interval

In addition to sending LLDP information at scheduled intervals, LLDP information is also sent when the G8124-E detects relevant changes to its configuration or status (such as when ports are enabled or disabled). To prevent the G8124-E from sending multiple LLDP packets in rapid succession when port status is in flux, a transmit delay timer can be configured.

The transmit delay timer represents the minimum time permitted between successive LLDP transmissions on a port. Any interval-driven or change-driven updates will be consolidated until the configured transmit delay expires.

The minimum transmit interval can be configured using the following command:

```
RS G8124-E(config)# lldp transmission-delay <interval>
```

where *interval* is the minimum number of seconds permitted between successive LLDP transmissions on any port. The range is 1 to one-quarter of the scheduled transmit interval (lldp refresh-interval *<value>*), up to 8192. The default is 2 seconds.

Time-to-Live for Transmitted Information

The transmitted LLDP information is held by remote systems for a limited time. A time-to-live parameter allows the switch to determine how long the transmitted data is held before it expires. The hold time is configured as a multiple of the configured transmission interval.

```
RS G8124-E(config)# lldp holdtime-multiplier <multiplier>
```

where *multiplier* is a value between 2 and 10. The default value is 4, meaning that remote systems will hold the port's LLDP information for 4 **x** the 30-second `msgtxint` value, or 120 seconds, before removing it from their MIB.

Trap Notifications

If SNMP is enabled on the G8124-E (see "Using Simple Network Management Protocol" on page 37), each port can be configured to send SNMP trap notifications whenever LLDP transmissions are sent. By default, trap notification is disabled for each port. The trap notification state can be changed using the following commands (Interface Port mode):

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# [no] lldp trap-notification
RS G8124-E(config-if)# exit
```

In addition to sending LLDP information at scheduled intervals, LLDP information is also sent when the G8124-E detects relevant changes to its configuration or status (such as when ports are enabled or disabled). To prevent the G8124-E from sending multiple trap notifications in rapid succession when port status is in flux, a global trap delay timer can be configured.

The trap delay timer represents the minimum time permitted between successive trap notifications on any port. Any interval-driven or change-driven trap notices from the port will be consolidated until the configured trap delay expires.

The minimum trap notification interval can be configured using the following command:

```
RS G8124-E(config)# lldp trap-notification-interval <interval>
```

where *interval* is the minimum number of seconds permitted between successive LLDP transmissions on any port. The range is 1 to 3600. The default is 5 seconds.

If SNMP trap notification is enabled, the notification messages can also appear in the system log. This is enabled by default. To change whether the SNMP trap notifications for LLDP events appear in the system log, use the following command:

```
RS G8124-E(config)# [no] logging log lldp
```

Changing the LLDP Transmit State

When the port is disabled, or when LLDP transmit is turned off for the port using the LLDP admin-status command options (see "Transmit and Receive Control" on page 449), a final LLDP packet is transmitted with a time-to-live value of 0. Neighbors that receive this packet will remove the LLDP information associated with the G8124-E port from their MIB.

In addition, if LLDP is fully disabled on a port and then later re-enabled, the G8124-E will temporarily delay resuming LLDP transmissions on the port to allow the port LLDP information to stabilize. The reinitialization delay interval can be globally configured for all ports using the following command:

```
RS G8124-E(config)# lldp reinit-delay <interval>
```

where *interval* is the number of seconds to wait before resuming LLDP transmissions. The range is between 1 and 10. The default is 2 seconds.

Types of Information Transmitted

When LLDP transmission is permitted on the port (see "Enabling or Disabling LLDP" on page 449), the port advertises the following required information in type/length/value (TLV) format:

- Chassis ID
- Port ID
- LLDP Time-to-Live

LLDP transmissions can also be configured to enable or disable inclusion of optional information, using the following command (Interface Port mode):

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# [no] lldp tlv <type>
RS G8124-E(config-if)# exit
```

where *type* is an LLDP information option from Table 32:

Table 32. *LLDP Optional Information Types*

Type	Description	Default
portdesc	Port Description	Enabled
sysname	System Name	Enabled
sysdescr	System Description	Enabled
syscap	System Capabilities	Enabled
mgmtaddr	Management Address	Enabled
portvid	IEEE 802.1 Port VLAN ID	Disabled
portprot	IEEE 802.1 Port and Protocol VLAN ID	Disabled
vlanname	IEEE 802.1 VLAN Name	Disabled

Table 32. *LLDP Optional Information Types (continued)*

Type	Description	Default
protid	IEEE 802.1 Protocol Identity	Disabled
macphy	IEEE 802.3 MAC/PHY Configuration/Status, including the auto-negotiation, duplex, and speed status of the port.	Disabled
powermdi	IEEE 802.3 Power via MDI, indicating the capabilities and status of devices that require or provide power over twisted-pair copper links.	Disabled
linkaggr	IEEE 802.3 Link Aggregation status for the port.	Disabled
framesz	IEEE 802.3 Maximum Frame Size for the port.	Disabled
dcbx	Data Center Bridging Capability Exchange Protocol (DCBX) for the port.	Enabled
all	Select all optional LLDP information for inclusion or exclusion.	Disabled

LLDP Receive Features

Types of Information Received

When the LLDP receive option is enabled on a port (see "Enabling or Disabling LLDP" on page 449), the port may receive the following information from LLDP-capable remote systems:

- Chassis Information
- Port Information
- LLDP Time-to-Live
- Port Description
- System Name
- System Description
- System Capabilities Supported/Enabled
- Remote Management Address

The G8124-E stores the collected LLDP information in the MIB. Each remote LLDP-capable device is responsible for transmitting regular LLDP updates. If the received updates contain LLDP information changes (to port state, configuration, LLDP MIB structures, deletion), the switch will set a change flag within the MIB for convenient notification to SNMP-based management systems.

Viewing Remote Device Information

LLDP information collected from neighboring systems can be viewed in numerous ways:

- Using a centrally-connected LLDP analysis server
- Using an SNMP agent to examine the G8124-E MIB
- Using the G8124-E Browser-Based Interface (BBI)
- Using CLI or isCLI commands on the G8124-E

Using the isCLI the following command displays remote LLDP information:

```
RS G8124-E(config)# show lldp remote-device [<index number>]
```

To view a summary of remote information, omit the *Index number* parameter. For example:

```
RS G8124-E(config)# show lldp remote-device
LLDP Remote Devices Information
Legend(possible values in DMAC column) :
NB   - Nearest Bridge         - 01-80-C2-00-00-0E
NnTB - Nearest non-TPMR Bridge - 01-80-C2-00-00-03
NCB  - Nearest Customer Bridge - 01-80-C2-00-00-00
Total number of current entries: 1


LocalPort|Index|Remote Chassis ID |Remote Port|Remote System Name|DMAC
---------|-----|----------------- |-----------|------------------|-----
    MGT  |  1  | 00 00 00 00 11 00|     5     |        C12        | NB
    20   |  2  | 74 99 75 1c 75 00|    44     |        C3+        | NB
```

To view detailed information for a remote device, specify the *Index number* as found in the summary. For example, in keeping with the sample summary, to list details for the first remote device (with an Index value of 1), use the following command:

```
RS G8124-E(config)# show lldp remote-device 1
Local Port Alias: 3
        Remote Device Index    : 1
        Remote Device TTL      : 99
        Remote Device RxChanges : false
        Chassis Type           : Mac Address
        Chassis Id             : 00-18-b1-33-1d-00
        Port Type              : Locally Assigned
        Port Id                : 23
        Port Description       : 7

        System Name       :
        System Description: Lenovo RackSwitch G8124-E, Lenovo
Networking OS: version 8.3, Boot Image: version 8.3

        System Capabilities Supported : bridge, router
        System Capabilities Enabled   : bridge, router

        Remote Management Address:
                Subtype            : IPv4
                Address            : 10.100.120.181
                Interface Subtype  : ifIndex
                Interface Number   : 128
                Object Identifier  :
```

Note: Received LLDP information can change very quickly. When using show commands, it is possible that flags for some expected events may be too short-lived to be observed in the output.

To view detailed information of all remote devices, use the following command:

```
RS G8124-E# show lldp remote-device detail

Local Port Alias: MGTA
        Remote Device Index        : 1
        Remote Device TTL          : 4678
        Remote Device RxChanges    : false
        Chassis Type               : Mac Address
        Chassis Id                 : 08-17-f4-a1-db-00
        Port Type                  : Locally Assigned
        Port Id                    : 25
        Port Description           : MGTA

        System Name                :
        System Description         : Lenovo RackSwitch G8124-E, Lenovo
Networking OS: version 8.3, Boot Image: version 6.9.1.14
        System Capabilities Supported : bridge, router
        System Capabilities Enabled   : bridge, router

        Remote Management Address:
                Subtype            : IPv4
                Address            : 10.38.22.23
                Interface Subtype  : ifIndex
                Interface Number   : 127
                Object Identifier  :

Local Port Alias: 2
        Remote Device Index        : 2
        Remote Device TTL          : 4651
        Remote Device RxChanges    : false
        Chassis Type               : Mac Address
        Chassis Id                 : 08-17-f4-a1-db-00
        Port Type                  : Locally Assigned
        Port Id                    : 2
        Port Description           : 2

        System Name                :
        System Description         : Lenovo RackSwitch G8124-E, Lenovo
Networking OS: version 8.3, Boot Image: version 6.9.1.14
        System Capabilities Supported : bridge, router
        System Capabilities Enabled   : bridge, router

        Remote Management Address:
                Subtype            : IPv4
                Address            : 10.38.22.23
                Interface Subtype  : ifIndex
                Interface Number   : 127
                Object Identifier  :

Total entries displayed:  2
```

Time-to-Live for Received Information

Each remote device LLDP packet includes an expiration time. If the switch port does not receive an LLDP update from the remote device before the time-to-live clock expires, the switch will consider the remote information to be invalid, and will remove all associated information from the MIB.

Remote devices can also intentionally set their LLDP time-to-live to 0, indicating to the switch that the LLDP information is invalid and must be immediately removed.

LLDP Example Configuration

1. Turn LLDP on globally.

```
RS G8124-E(config)# lldp enable
```

2. Set the global LLDP timer features.

```
RS G8124-E(config)# lldp refresh-interval 30(Transmit each 30 seconds)
RS G8124-E(config)# lldp transmission-delay 2(No more often than 2 sec.)
RS G8124-E(config)# lldp holdtime-multiplier 4(Remote hold 4 intervals)
RS G8124-E(config)# lldp reinit-delay 2(Wait 2 sec. after reinit.)
RS G8124-E(config)# lldp trap-notification-interval 5(Minimum 5 sec. between)
```

3. Set LLDP options for each port.

```
RS G8124-E(config)# interface port <n> (Select a switch port)
RS G8124-E(config-if)# lldp admin-status tx_rx(Transmit and receive LLDP)
RS G8124-E(config-if)# lldp trap-notification(Enable SNMP trap notifications)
RS G8124-E(config-if)# lldp tlv all     (Transmit all optional information)
RS G8124-E(config-if)# exit
```

4. Enable syslog reporting.

```
RS G8124-E(config)# logging log lldp
```

5. Verify the configuration settings:

```
RS G8124-E(config)# show lldp
```

6. View remote device information as needed.

```
RS G8124-E(config)# show lldp remote-device
   or
RS G8124-E(config)# show lldp remote-device <index number>
   or
RS G8124-E(config)# show lldp remote-devices detail
```

Chapter 33. Simple Network Management Protocol

Lenovo Networking OS provides Simple Network Management Protocol (SNMP) version 1, version 2, and version 3 support for access through any network management software, such as IBM Director or HP-OpenView.

Note: SNMP read and write functions are enabled by default. For best security practices, if SNMP is not needed for your network, it is recommended that you disable these functions prior to connecting the switch to the network.

SNMP Version 1 & Version 2

To access the SNMP agent on the G8124-E, the read and write community strings on the SNMP manager must be configured to match those on the switch. The default read community string on the switch is public and the default write community string is private.

The read and write community strings on the switch can be changed using the following commands on the CLI:

```
RS G8124-E(config)# snmp-server read-community <1-32 characters>
    -and-
RS G8124-E(config)# snmp-server write-community <1-32 characters>
```

The SNMP manager must be able to reach the management interface or any one of the IP interfaces on the switch.

For the SNMP manager to receive the SNMPv1 traps sent out by the SNMP agent on the switch, configure the trap host on the switch with the following command:

```
RS G8124-E(config)# snmp-server trap-source <trap source IP interface>
RS G8124-E(config)# snmp-server host <IPv4 address> <trap host community string>
```

Note: You can use a loopback interface to set the source IP address for SNMP traps. Use the following command to apply a configured loopback interface:
RS G8124-E(config)# **snmp-server trap-source loopback** *<1-5>*

SNMP Version 3

SNMP version 3 (SNMPv3) is an enhanced version of the Simple Network Management Protocol, approved by the Internet Engineering Steering Group in March, 2002. SNMPv3 contains additional security and authentication features that provide data origin authentication, data integrity checks, timeliness indicators and encryption to protect against threats such as masquerade, modification of information, message stream modification and disclosure.

SNMPv3 allows clients to query the MIBs securely.

SNMPv3 configuration is managed using the following command path menu:

```
RS G8124-E(config)# snmp-server ?
```

For more information on SNMP MIBs and the commands used to configure SNMP on the switch, see the *Lenovo Networking OS 8.3 Command Reference*.

Default Configuration

Lenovo N/OS has three SNMPv3 users by default. All the three users have access to all the MIBs supported by the switch:

- User 1 name is adminmd5 (password adminmd5). Authentication used is MD5. Privacy protocol used is DES.
- User 2 name is adminsha (password adminsha). Authentication used is SHA. Privacy protocol used is DES.
- User 3 name is adminshaaes (password Edpq132x!#9Zpx432w). Authentication used is SHA. Privacy protocol used is AES-128.

In boot strict mode (See "Boot Strict Mode" on page 47), Lenovo N/OS has one SNMPv3 user:

- User 1 name is adminshaaes (password Edpq132x!#9Zpx432w). Authentication used is SHA. Privacy protocol used is AES-128.

Up to 17 SNMP users can be configured on the switch. To modify an SNMP user, enter the following commands:

```
RS G8124-E(config)# snmp-server user <1-17> name <1-32 characters>
```

Users can be configured to use the authentication/privacy options. The G8124-E support two authentication algorithms: MD5 and SHA, as specified in the following command:

```
RS G8124-E(config)# snmp-server user <1-17> authentication-protocol
    {md5|sha} authentication-password

  -or-

RS G8124-E(config)# snmp-server user <1-17> authentication-protocol none
```

User Configuration Example

1. To configure a user with name "admin," authentication type MD5, and authentication password of "admin," privacy option DES with privacy password of "admin," use the following CLI commands.

```
RS G8124-E(config)# snmp-server user 5 name admin
RS G8124-E(config)# snmp-server user 5 authentication-protocol md5
    authentication-password
Changing authentication password; validation required:
Enter current admin password:          <admin. password>
Enter new authentication password:
                                       <auth. password; Should be more than 18 characters >
Re-enter new authentication password: <auth. password>
New authentication password accepted.

RS G8124-E(config)# snmp-server user 5 privacy-protocol des
    privacy-password
Changing privacy password; validation required:
Enter current admin password:          <admin. password>
Enter new privacy password:
                                       <privacy password; Should be more than 18 characters>
Re-enter new privacy password:         <privacy password>
New privacy password accepted.
```

2. Configure a user access group, along with the views the group may access. Use the access table to configure the group's access level.

```
RS G8124-E(config)# snmp-server access 5 name admingrp
RS G8124-E(config)# snmp-server access 5 level authpriv
RS G8124-E(config)# snmp-server access 5 read-view iso
RS G8124-E(config)# snmp-server access 5 write-view iso
RS G8124-E(config)# snmp-server access 5 notify-view iso
```

Because the read view, write view, and notify view are all set to "iso," the user type has access to all private and public MIBs.

3. Assign the user to the user group. Use the group table to link the user to a particular access group.

```
RS G8124-E(config)# snmp-server group 5 user-name admin
RS G8124-E(config)# snmp-server group 5 group-name admingrp
```

Configuring SNMP Trap Hosts

SNMPv1 Trap Host

1. Configure a user with no authentication and password.

```
RS G8124-E(config)# snmp-server user 10 name v1trap
```

2. Configure an access group and group table entries for the user. Use the following menu to specify which traps can be received by the user:

```
RS G8124-E(config)# snmp-server access <user number>
```

In the following example the user will receive the traps sent by the switch.

```
RS G8124-E(config)# snmp-server access 10 (Access group to view SNMPv1 traps)
    name v1trap
    security snmpv1
    notify-view iso
RS G8124-E(config)# snmp-server group 10 (Assign user to the access group)
    security snmpv1
    user-name v1trap
    group-name v1trap
```

3. Configure an entry in the notify table.

```
RS G8124-E(config)# snmp-server notify 10 name v1trap
RS G8124-E(config)# snmp-server notify 10 tag v1trap
```

4. Specify the IPv4 address and other trap parameters in the targetAddr and targetParam tables. Use the following commands to specify the user name associated with the targetParam table:

```
RS G8124-E(config)# snmp-server target-address 10 name v1trap address
    10.70.70.190
RS G8124-E(config)# snmp-server target-address 10 parameters-name v1param
RS G8124-E(config)# snmp-server target-address 10 taglist v1param
RS G8124-E(config)# snmp-server target-parameters 10 name v1param
RS G8124-E(config)# snmp-server target-parameters 10 user-name v1only
RS G8124-E(config)# snmp-server target-parameters 10 message snmpv1
```

Note: N/OS 8.3 supports only IPv4 addresses for SNMPv1 and SNMPv2 trap hosts.

5. Use the community table to specify which community string is used in the trap.

```
RS G8124-E(config)# snmp-server community 10(        Define the community string)
    index v1trap
    name public
    user-name v1trap
```

SNMPv2 Trap Host Configuration

The SNMPv2 trap host configuration is similar to the SNMPv1 trap host configuration. Wherever you specify the model, use snmpv2 instead of snmpv1.

```
RS G8124-E(config)# snmp-server user 10 name v2trap

RS G8124-E(config)# snmp-server group 10 security snmpv2
RS G8124-E(config)# snmp-server group 10 user-name v2trap
RS G8124-E(config)# snmp-server group 10 group-name v2trap
RS G8124-E(config)# snmp-server access 10 name v2trap
RS G8124-E(config)# snmp-server access 10 security snmpv2
RS G8124-E(config)# snmp-server access 10 notify-view iso

RS G8124-E(config)# snmp-server notify 10 name v2trap
RS G8124-E(config)# snmp-server notify 10 tag v2trap

RS G8124-E(config)# snmp-server target-address 10 name v2trap
    address 100.10.2.1
RS G8124-E(config)# snmp-server target-address 10 taglist v2trap
RS G8124-E(config)# snmp-server target-address 10 parameters-name
    v2param
RS G8124-E(config)# snmp-server target-parameters 10 name v2param
RS G8124-E(config)# snmp-server target-parameters 10 message snmpv2c
RS G8124-E(config)# snmp-server target-parameters 10 user-name v2trap
RS G8124-E(config)# snmp-server target-parameters 10 security snmpv2

RS G8124-E(config)# snmp-server community 10 index v2trap
RS G8124-E(config)# snmp-server community 10 user-name v2trap
```

Note: N/OS 8.3 supports only IPv4 addresses for SNMPv1 and SNMPv2 trap hosts.

SNMPv3 Trap Host Configuration

To configure a user for SNMPv3 traps, you can choose to send the traps with both privacy and authentication, with authentication only, or without privacy or authentication.

This is configured in the access table using the following commands:

```
RS G8124-E(config)# snmp-server access <1-32> level
RS G8124-E(config)# snmp-server target-parameters <1-16> level
```

Configure the user in the user table accordingly.

It is not necessary to configure the community table for SNMPv3 traps because the community string is not used by SNMPv3.

The following example shows how to configure a SNMPv3 user v3trap with authentication only:

```
RS G8124-E(config)# snmp-server user 11 name v3trap
RS G8124-E(config)# snmp-server user 11 authentication-protocol md5
    authentication-password
Changing authentication password; validation required:
Enter current admin password:          <admin. password>
Enter new authentication password:     <auth. password>
Re-enter new authentication password:  <auth. password>
New authentication password accepted.
RS G8124-E(config)# snmp-server access 11 notify-view iso
RS G8124-E(config)# snmp-server access 11 level authNopriv
RS G8124-E(config)# snmp-server group 11 user-name v3trap
RS G8124-E(config)# snmp-server group 11 group-name v3trap
RS G8124-E(config)# snmp-server notify 11 name v3trap
RS G8124-E(config)# snmp-server notify 11 tag v3trap
RS G8124-E(config)# snmp-server target-address 11 name v3trap address
    47.81.25.66
RS G8124-E(config)# snmp-server target-address 11 taglist v3trap
RS G8124-E(config)# snmp-server target-address 11 parameters-name v3param
RS G8124-E(config)# snmp-server target-parameters 11 name v3param
RS G8124-E(config)# snmp-server target-parameters 11 user-name v3trap
RS G8124-E(config)# snmp-server target-parameters 11 level authNoPriv
```

SNMP MIBs

The N/OS SNMP agent supports SNMP version 3. Security is provided through SNMP community strings. The default community strings are "`public`" for `SNMP GET` operation and "`private`" for `SNMP SET` operation. The community string can be modified only through the Industry Standard Command Line Interface (ISCLI). Detailed SNMP MIBs and trap definitions of the N/OS SNMP agent are contained in the N/OS enterprise MIB document.

The N/OS SNMP agent supports the following standard MIBs:

- `dot1x.mib`
- `ieee8021ab.mib`
- `ieee8023ad.mib`
- `lldpxdcbx.mib`
- `rfc1213.mib`
- `rfc1215.mib`
- `rfc1493.mib`
- `rfc1573.mib`
- `rfc1643.mib`
- `rfc1657.mib`
- `rfc1757.mib`
- `rfc1850.mib`
- `rfc1907.mib`
- `rfc2037.mib`
- `rfc2233.mib`
- `rfc2465.mib`
- `rfc2571.mib`
- `rfc2572.mib`
- `rfc2573.mib`
- `rfc2574.mib`
- `rfc2575.mib`
- `rfc2576.mib`
- `rfc2790.mib`
- `rfc3176.mib`
- `rfc4133.mib`
- `rfc4363.mib`

The N/OS SNMP agent supports the following generic traps as defined in RFC 1215:

- ColdStart
- WarmStart
- LinkDown
- LinkUp
- AuthenticationFailure

The SNMP agent also supports two Spanning Tree traps as defined in RFC 1493:

- NewRoot
- TopologyChange

The following are the enterprise SNMP traps supported in N/OS:

Table 33. *Lenovo N/OS-Supported Enterprise SNMP Traps*

Trap Name	Description
swLoginFailure	Signifies that someone failed to enter a valid username/password combination. swTrapDisplayString specifies whether the login attempt was from CONSOLE or TELNET. In case of TELNET login it also specifies the IP address of the host from which the attempt was made.
swValidLogin	Signifies that a user login has occurred.
swApplyComplete	Signifies that new configuration has been applied.
swSaveComplete	Signifies that new configuration has been saved.
swFwDownloadSucess	Signifies that firmware has been downloaded to [image1|image2|boot image].
swFwDownloadFailure	Signifies that firmware downloaded failed to [image1|image2|boot image].
swValidLogout	Signifies that a user logout has occurred.
swPrimaryPowerSupplyFailure	Signifies that the primary power supply has failed.
swPrimaryPowerSupplyFixed	Signifies that the primary power supply has been restored.
swDefAdminDisable	Signifies that the default admin account has been disabled.
swAcntStrngPswdNotMet	Signifies that the configured password does not match strong password complexity.
swAcntLocked	Signifies that account has been locked.
swAcntUnlocked	Signifies that account has been unlocked.
swStgNewRoot	Signifies that the bridge has become the new root of the STG.

Table 33. *Lenovo N/OS-Supported Enterprise SNMP Traps (continued)*

Trap Name	Description
swCistNewRoot	Signifies that the bridge has become the new root of the CIST.
swStgTopologyChanged	Signifies that there was a STG topology change.
swCistTopologyChanged	Signifies that there was a CIST topology change.
swHotlinksMasterUp	Signifies that the Master interface is active.
swHotlinksMasterDn	Signifies that the Master interface is not active.
swHotlinksBackupUp	Signifies that the Backup interface is active.
swHotlinksBackupDn	Signifies that the Backup interface is not active.
swHotlinksNone	Signifies that there are no active interfaces.
swTeamingCtrlUp	Signifies that the teaming is up.
swTeamingCtrlDown	Signifies that the teaming control is down.
swTeamingCtrlDownTearDownBlked	Signifies that the teaming control is down but teardown is blocked.
swTeamingCtrlError	Signifies error, action is undefined.
swLACPPortBlocked	Signifies that LACP is operationally down on a port, and traffic is blocked on the port.
swLACPPortUnblocked	Signifies that LACP is operationally up on a port, and traffic is no longer blocked on the port.
swLFDPortErrdisabled	Signifies that a port is error-disabled due to excessive link flaps.
swVlagInstanceUp	Signifies that VLAG instance is up identified in the trap message.
swVlagInstanceRemoteUp	Signifies that VLAG is down but instance on the remote instance is up.
swVlagInstanceLocalUp	Signifies that VLAG is down but local instance is up.

Table 33. *Lenovo N/OS-Supported Enterprise SNMP Traps (continued)*

Trap Name	Description
swVlagInstanceDown	Signifies that VLAG instance is down identified in the trap message.
swVlagIslUp	Signifies that connection between VLAG switches is up.
swVlagIslDown	Signifies that connection between VLAG switches is down.
sw8021x	Signifies 802.1x feature is enabled.
swDefGwUp	Signifies that the default gateway is alive. ipCurCfgGwIndex is the index of the Gateway in ipCurCfgGwTable. The range for ipCurCfgGwIndex is from 1 to ipGatewayTableMax. ipCurCfgGwAddr is the IP address of the default gateway.
swDefGwDown	Signifies that the default gateway is down. ipCurCfgGwIndex is the index of the Gateway in ipCurCfgGwTable. The range for ipCurCfgGwIndex is from 1 to ipGatewayTableMax. ipCurCfgGwAddr is the IP address of the default gateway.
swDefGwInService	Signifies that the default gateway is up and in service. ipCurCfgGwIndex is the index of the Gateway in ipCurCfgGwTable. The range for ipCurCfgGwIndex is from 1 to ipGatewayTableMax. ipCurCfgGwAddr is the IP address of the default gateway.
swDefGwNotInService	Signifies that the default gateway is alive but not in service. ipCurCfgGwIndex is the index of the Gateway in ipCurCfgGwTable. The range for ipCurCfgGwIndex is from 1 to ipGatewayTableMax. ipCurCfgGwAddr is the IP address of the default gateway.

Table 33. *Lenovo N/OS-Supported Enterprise SNMP Traps (continued)*

Trap Name	Description
swVrrpNewMaster	Indicates that the sending agent has transitioned to "Master" state. vrrpCurCfgVirtRtrIndx is the VRRP virtual router table index referenced in vrrpCurCfgVirtRtrTable. The range is from 1 to vrrpVirtRtrTableMaxSize. vrrpCurCfgVirtRtrAddr is the VRRP virtual router IP address.
swVrrpNewBackup	Indicates that the sending agent has transitioned to "Backup" state. vrrpCurCfgVirtRtrIndx is the VRRP virtual router table index referenced in vrrpCurCfgVirtRtrTable. The range is from 1 to vrrpVirtRtrTableMaxSize. vrrpCurCfgVirtRtrAddr is the VRRP virtual router IP address.
swVrrpAuthFailure	Signifies that a packet has been received from a router whose authentication key or authentication type conflicts with this router's authentication key or authentication type. Implementation of this trap is optional. vrrpCurCfgIfIndx is the VRRP interface index. This is equivalent to ifIndex in RFC 1213 mib. The range is from 1 to vrrpIfTableMaxSize. vrrpCurCfgIfPasswd is the password for authentication. It is a DisplayString of 0 to 7 characters.

Table 33. *Lenovo N/OS-Supported Enterprise SNMP Traps (continued)*

Trap Name	Description
swTcpHoldDown	Signifies that new TCP connection requests from a particular client will be blocked for a pre-determined amount of time since the rate of new TCP connections from that client has reached a pre-determined threshold. The fltCurCfgSrcIp is the client source IP address for which new TCP connection requests will be blocked. The swTrapRate specifies the amount of time in minutes that the particular client will be blocked.
swNTPSyncFailed	Signifies that synchronization with the NTP server has failed.
swNTPUpdateClock	Signifies that the system clock is updated with NTP server.
swECMPGatewayUp	Signifies that the ECMP gateway is up.
swECMPGatewayDown	Signifies that the ECMP gateway is down.
swTempExceedThreshold	Signifies that the switch temperature has exceeded maximum safety limits.
swFanFailure	Signifies that fan failure has been detected.
swFanFailureFixed	Signifies that the fan failure has been fixed.
swTempReturnThreshold	Signifies that the switch temperature has returned to under maximum safety limits.
swVMGroupVMotion	Signifies that a virtual machine has moved from a port to another.
swVMGroupVMOnline	Signifies that an advance provisioned virtual machine has came online.
swVMGroupVMVlanChange	Signifies that a virtual machine has entered a VLAN, or changed the VLAN.
swVmCheckSpoofedvm	Signifies that a spoofed VM MAC was found.

Switch Images and Configuration Files

This section describes how to use MIB calls to work with switch images and configuration files. You can use a standard SNMP tool to perform the actions, using the MIBs listed in the table below.

The table lists the MIBS used to perform operations associated with the Switch Image and Configuration files.

Table 34. *MIBs for Switch Image and Configuration Files*

MIB Name	MIB OID
agTransferServer	1.3.6.1.4.26543.2.5.1.1.7.1.0
agTransferImage	1.3.6.1.4.26543.2.5.1.1.7.2.0
agTransferImageFileName	1.3.6.1.4.26543.2.5.1.1.7.3.0
agTransferCfgFileName	1.3.6.1.4.26543.2.5.1.1.7.4.0
agTransferDumpFileName	1.3.6.1.4.26543.2.5.1.1.7.5.0
agTransferAction	1.3.6.1.4.26543.2.5.1.1.7.6.0
agTransferLastActionStatus	1.3.6.1.4.26543.2.5.1.1.7.7.0
agTransferUserName	1.3.6.1.4.26543.2.5.1.1.7.9.0
agTransferPassword	1.3.6.1.4.1.26543.2.5.1.1.7.10.0
agTransferTSDumpFileName	1.3.6.1.4.1.26543.2.5.1.1.7.11.0

The following SNMP actions can be performed using the MIBs listed in the table:

● Load a new Switch image (boot or running) from an FTP, SFTP, or TFTP server

● Load a previously saved switch configuration from an FTP, SFTP, or TFTP server

● Save the switch configuration to an FTP, SFTP, or TFTP server

● Save a switch dump to an FTP, SFTP, or TFTP server

Loading a New Switch Image

To load a new switch image with the name "`MyNewImage-1.img`" into `image2`, follow these steps. This example shows an FTP or TFTP server at IPv4 address 192.168.10.10, though IPv6 is also supported.

1. Set the FTP or TFTP server address where the switch image resides:

   ```
   Set agTransferServer.0 "192.168.10.10"
   ```

2. Set the area where the new image will be loaded:

   ```
   Set agTransferImage.0 "image2"
   ```

3. Set the name of the image:

   ```
   Set agTransferImageFileName.0 "MyNewImage-1.img"
   ```

4. If you are using an FTP server, enter a username:

   ```
   Set agTransferUserName.0 "MyName"
   ```

5. If you are using an FTP server, enter a password:

   ```
   Set agTransferPassword.0 "MyPassword"
   ```

6. Initiate the transfer. To transfer a switch image, enter 2 (gtimg):

   ```
   Set agTransferAction.0 "2"
   ```

Loading a Saved Switch Configuration

To load a saved switch configuration with the name "MyRunningConfig.cfg" into the switch, follow these steps. This example shows an FTP, SFTP, or TFTP server at IPv4 address 192.168.10.10, though IPv6 is also supported.

1. Set the FTP, SFTP, or TFTP server address where the switch Configuration File resides:

   ```
   Set agTransferServer.0 "192.168.10.10"
   ```

2. Set the name of the configuration file:

   ```
   Set agTransferCfgFileName.0 "MyRunningConfig.cfg"
   ```

3. If you are using an FTP or SFTP server, enter a username:

   ```
   Set agTransferUserName.0 "MyName"
   ```

4. If you are using an FTP or SFTP server, enter a password:

   ```
   Set agTransferPassword.0 "MyPassword"
   ```

5. Initiate the transfer. To restore a running configuration, enter 3:

   ```
   Set agTransferAction.0 "3"
   ```

Saving the Switch Configuration

To save the switch configuration to a file server, follow these steps. This example shows an FTP, SFTP, or TFTP server at IPv4 address 192.168.10.10, though IPv6 is also supported.

1. Set the FTP, SFTP, or TFTP server address where the configuration file is saved:

   ```
   Set agTransferServer.0 "192.168.10.10"
   ```

2. Set the name of the configuration file:

   ```
   Set agTransferCfgFileName.0 "MyRunningConfig.cfg"
   ```

3. If you are using an FTP or SFTP server, enter a username:

   ```
   Set agTransferUserName.0 "MyName"
   ```

4. If you are using an FTP or SFTP server, enter a password:

   ```
   Set agTransferPassword.0 "MyPassword"
   ```

5. Initiate the transfer. To save a running configuration file, enter 4:

   ```
   Set agTransferAction.0 "4"
   ```

Saving a Switch Dump

To save a switch dump to a file server, follow these steps. This example shows an FTP, SFTP, or TFTP server at 192.168.10.10, though IPv6 is also supported.

1. Set the FTP, SFTP, or TFTP server address where the configuration will be saved:

   ```
   Set agTransferServer.0 "192.168.10.10"
   ```

2. Set the name of dump file:

   ```
   Set agTransferDumpFileName.0 "MyDumpFile.dmp"
   ```

3. If you are using an FTP or SFTP server, enter a username:

   ```
   Set agTransferUserName.0 "MyName"
   ```

4. If you are using an FTP or SFTP server, enter a password:

   ```
   Set agTransferPassword.0 "MyPassword"
   ```

5. Initiate the transfer. To save a dump file, enter 5:

   ```
   Set agTransferAction.0 "5"
   ```

Chapter 34. NETCONF

The Network Configuration Protocol (NETCONF) provides a mechanism to manage the G8124-E, retrieve or modify existing configuration data, and upload new configuration data. See RFC 4741 for details on NETCONF.

NETCONF operates in a client/server model. The NETCONF client establishes a session with the switch (acting as a NETCONF server) using a Remote Procedure Call (RPC). NETCONF is based on the Extensible Markup Language (XML) for encoding data and for exchanging configuration and protocol messages.

The following topics are discussed in this section:

NETCONF Overview

NETCONF provides a method to quickly configure the switch. It also allows you to implement a configuration across multiple switches, thereby saving time and reducing the chances of configuration errors.

The NETCONF protocol defines basic operations that are equivalent to the switch ISCLI commands.

Note: The current implementation of NETCONF supports only ISCLI commands.

NETCONF is a connection-oriented protocol. See Figure 54 for an overview of NETCONF operation.

Figure 54. NETCONF Operations Procedure
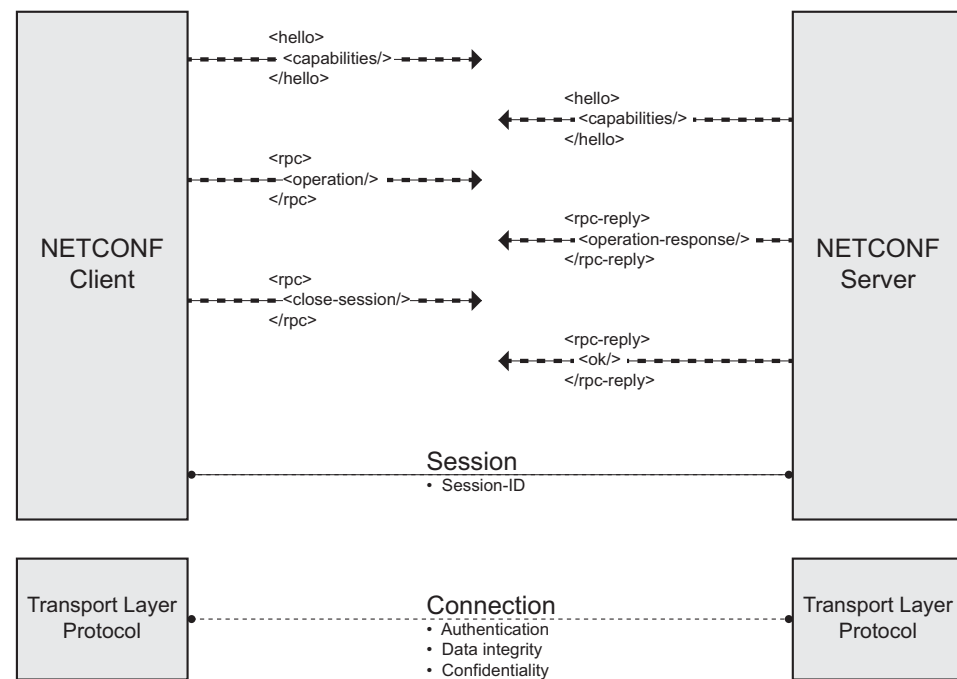
1. The client establishes a transport layer connection to the switch (acting as a NETCONF server).

2. The client and switch exchange `hello` messages to declare their capabilities.

3. The client sends a request via `rpc` message to the switch.

4. The switch sends a response via `rpc-reply` message to the client.

   Note: Steps 3 and 4 must be repeated for each request that the client sends to the switch.

5. The client sends a `close-session` message to the switch to end the NETCONF session and the transport layer connection.

6. The switch sends an `ok` response.

XML Requirements

XML is the encoding format used within NETCONF. When using XML for NETCONF:

- All NETCONF protocol elements are defined in the following namespace:

  `urn:ietf:params:xml:ns:netconf:base:1.0`

- NETCONF capability names must be Uniform Resource Identifiers (URIs):

  `urn:ietf:params:netconf:capability:`*{name}*`:1.0`

  where *{name}* is the name of the capability.

- Document type declarations must not appear in the NETCONF content.

- For Secure Shell (SSH), you must use a special message termination sequence of six characters to provide message framing:

  `]]>]]>`

Installing the NETCONF Client

You can download the required NETCONF Client installation files from www.ibm.com. Select **Support & downloads > Fixes, updates and drivers**. Follow instructions on the IBM Support Portal page to find the files.

Before installing the NETCONF client, ensure you have completed the following tasks:

- Install a supported version of Python (Python 2.6 or higher, up to but not including Python 3.0) in the folder `C:\`.

- Install the PyCrypto application appropriate to the Python version you are using.

Note: The following steps are for the Windows operating systems.

Follow these steps to install the Blade NETCONF Python Client (BNClient):

1. Extract the file `blade-netconf-python-client-v0.1.zip` to the following folder:

   `C:\`

   Under the root folder `C:\blade-netconf-python-client-v0.1`, you will see two folders:

   - `blade-netconf-python-client`

   - `python-library`

   Note: Make sure Paramiko version 1.7.4 or higher is in the folder `C:\blade-netconf-python-client-v0.1\python-ssh-library\`

2. Open the command prompt (**Select Start > Run > cmd**).

3. Enter the following command to install the SSH library:

   ```
   python C:\blade-netconf-python-client-v0.1\python-ssh-library\
   paramiko-1.7.6\setup.py install
   ```

   Note: If the `python` command does not work from the command prompt, you may need to add a system variable to set the path to the directory where you have installed Python. You can add the system variable at the following location: **My Computer > Properties > Advanced > Environment Variables**

4. Follow these steps to install BNClient:

As a python script:

a. Enter the following command for help:

```
python C:\blade-netconf-python-client-v0.1\blade-netconf-python-
client\bnclient\bnclient.py –h
```

b. Enter the following command to establish a NETCONF session:

o Using SSH to connect to default port 830:

```
python C:\blade-netconf-python-client-v0.1\blade-netconf-python-
    client\bnclient\bnclient.py {switch IP address} –u admin –p admin –o
    get
```

o Using SSH to connect to port 22:

```
python C:\blade-netconf-python-client-v0.1\blade-netconf-python-
client\bnclient\bnclient.py {switch IP address}:22 –u admin –p admin –o
get
```

As a python library:

a. Open the file
C:\blade-netconf-python-client-v0.1\blade-netconf-
python-client\example\get.py in a Python editor (For example, IDLE).

b. Change the IP address in the *hostname* field to the switch IP address, and save
the file.

c. Enter the following command to establish a session:

```
python C:\blade-netconf-python-client-v0.1\blade-netconf-python-
client\setup.py install
```

d. Enter the following command to get the running configuration:

```
python C:\blade-netconf-python-client-v0.1\blade-netconf-python-
client\example\get.py
```

Note: get.py is an example of a NETCONF operation python script. You can edit
the script or write a new script as per your requirements.

Using Juniper Perl Client

You can use Juniper Perl client instead of BNClient to communicate with the NETCONF feature on the switch. Follow these steps to use the Juniper Perl client.

Note: You must use the Linux operating system for the Juniper Perl client.

1. Extract the file `juniper-netconf-perl-client.zip` to the folder:

    `/home/user/`

    You will see two folders:

    - `juniper-netconf-perl-client`
    - `blade-netconf-perl-scripts`

2. Follow these steps to install the Juniper Perl client:

    As a Perl library:

    a. Change to the following directory:

      `/home/user/juniper-netconf-perl-client`

    b. Extract the following file:

      `netconf-perl-10.0R2.10.tar.gz`

    c. Change to the following directory:

      `/home/user/juniper-netconf-perl-client/netconf-perl-10.0R2.10`

    d. Install the client as per the instructions in the `README` file.

    Note: If the prerequisites package installation fails, manually install each file in `/home/user/juniper-netconf-perl-client\netconf-perl-prereqs-patch`.

    As a Perl script:

    a. Change to the following directory:

      `/home/user/blade-netconf-perl-scripts/`

    b. Enter the following command:

    ```
    perl get/get.pl -l admin -p admin {swich IP address}
    ```

    Note: `get.pl` is an example of a NETCONF operation Perl script. You can edit the script or write a new script as per your requirement.

Establishing a NETCONF Session

SSH is the widely used protocol for NETCONF sessions. The default SSH port for NETCONF is 830. The client may also connect to the switch through SSH port 22.

Follow these steps to establish a NETCONF session. Enter commands in the client Linux Shell.

Note: You can open a maximum of four simultaneous sessions.

1. Enter the following command to open an SSH connection:

```
ssh admin@{switch IP address} -p 830 -s netconf
```

2. Type or paste the following `hello` message:

```
<hello>
  <capabilities>
    <capability>urn:ietf:params:netconf:base:1.0</capability>
  </capabilities>
</hello>
]]>]]>
```

The switch returns a `hello` message:

```
<hello xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
  <capabilities>
    <capability>urn:ietf:params:netconf:base:1.0</capability>
    <capability>urn:ietf:params:netconf:capability:writable-running:1.0</capab
    ility>
    <capability>urn:ietf:params:netconf:capability:rollback-on-error:1.0</capa
    bility>
    <capability>urn:ietf:params:netconf:capability:startup:1.0</capability>
  </capabilities>
  <session-id>102</session-id>
</hello>
]]>]]>
```

3. Type or paste the following `rpc` message. The `get` operation is used as an example.

```
<rpc message-id="100">
  <get>
    <filter type="subtree">
      <configuration-text/>
    </filter>
  </get>
</rpc>
]]>]]>
```

The switch sends an rpc-reply message:

```
<rpc-reply message-id="100">
  <data>
     <configuration-text
     xmlns="http://www.ibm.com/netconf/1.0/config-text"> version "8.3.1"
        switch-type "Lenovo RackSwitch G8124-E"
        !
        !
        no system dhcp mgta
        !
        !
        interface ip 127
        ip address 172.31.36.51
        enable
        exit
        !
        ip gateway 3 address 172.31.1.1
        ip gateway 3 enable
        !
        !
        end
     </configuration-text>
  </data>
</rpc-reply>
]]>]]>
```

Note: Repeat Step 3 for each request you need to send to the switch.

4. Type or paste the following close-session message to close the NETCONF session and terminate the SSH connection.

```
<rpc message-id="101">
  <close-session/>
</rpc>
]]>]]>
```

The switch sends the following response:

```
<rpc-reply message-id="101">
  <ok/>
</rpc-reply>
]]>]]>
```

NETCONF Operations

The NETCONF protocol provides a set of operations to manage and retrieve switch configuration. Table 35 provides a list of protocol operations supported by the switch.

Table 35. *Protocol Operations*

Operation	Description
get-config	Retrieve all or part of the running or startup configuration.
edit-config	Load all or part of a specified configuration to the running or startup configuration.
copy-config	Replace the target running or startup configuration with a source running or startup configuration.
delete-config	Delete startup configuration.
lock	Lock the running configuration to prevent other users (via another NETCONF session) from changing it.
unlock	Release a locked running configuration.
get	Retrieve running configuration and device state information.
close-session	Request graceful termination of a NETCONF session.
kill-session	Force the termination of a NETCONF session.
get-configuration	Retrieve configuration data from the switch.
get-interface-information	Retrieve interface status information.

Protocol Operations Examples

Following are examples of the NETCONF protocol operations supported by the G8124-E.

<get-config>

Usage:

```
<rpc message-id="101" xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <get-config>
     <source>
        <running/>
     </source>
     <filter type="subtree">
        <configuration-text
        xmlns="http://www.ibm.com/netconf/1.0/config-text"/>
     </filter>
   </get-config>
</rpc>
```

Response from the switch:

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <data>
      <configuration-text
      xmlns="http://www.ibm.com/netconf/1.0/config-text">
        <!-- configuration text... -->
      </configuration-text>
   </data>
</rpc-reply>
```

See Table 36 for the tag elements and their values.

Table 36. `get-config` *Tag Element Values*

Tag Element	Description	Value
source	The configuration text you want to retrieve.	running/ or startup/
filter type="subtree"	The filter type.	subtree
!--configuration text...--	Contains the running configuration in ISCLI format.	

<edit-config>

Usage:

```
<rpc message-id="101" xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
  <edit-config>
    <target>
      <running/>
    </target>
    <default-operation>
      <merge/>
    </default-operation>
    <error-option>
      <stop-on-error/>
    </error-option>
    <config-text xmlns="http://www.ibm.com/netconf/1.0/config-text">
      <configuration-text>hostname Router</configuration-text>
    </config-text>
</edit-config>
</rpc>
```

Response from the switch:

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
  <ok/>
</rpc-reply>
```

See Table 37 for the tag elements and their values.

Table 37. `edit-config` *Tag Element Values*

Tag Element	Description	Value
target	The configuration you want to edit.	running/ or startup/

Table 37. `edit-config` *Tag Element Values*

Tag Element	Description	Value
`default-operation`	Set the default operation for the `edit-config` request.	• `merge`: The new configuration is merged with the target configuration at the corresponding level. • `replace`: The new configuration replaces the target configuration. • `none`: The target configuration does not change unless the configuration data in the `configuration-text` parameter uses the `operation` attribute to request a different operation.
`error-option`	Set the option to handle configuration error.	• `stop-on-error`: Abort the `edit-config` operation on first error. This is the default error-option. • `continue-on-error`: Continue to process configuration data on error. • `rollback-on-error`: Abort the `edit-config` operation on first error and discard the requested configuration changes.

<copy-config>

Usage:

```
<rpc message-id="101" xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <copy-config>
     <target>
       <startup/>
     </target>
     <source>
       <running/>
     </source>
   </copy-config>
</rpc>
```

Response from the switch:

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <ok/>
</rpc-reply>
```

See Table 38 for the tag elements and their values.

Table 38. `copy-config` *Tag Element Values*

Tag Element	Description	Value
target	Configuration that needs to be changed.	running/ or startup/
source	Source configuration.	running/ or startup/

<delete-config>

Usage:

```
<rpc message-id="101" xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <delete-config>
     <target>
       <startup/>
     </target>
   </delete-config>
</rpc>
```

Response from the switch:

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <ok/>
</rpc-reply>
```

See Table 39 for the tag elements and their values.

Table 39. delete-config *Tag Element Values*

Tag Element	Description	Value
target	Configuration that needs to be deleted.	startup/

<lock>

Usage:

```
<rpc message-id="101" xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <lock>
     <target>
       <running/>
     </target>
   </lock>
</rpc>
```

Response from the switch:

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <ok/>
</rpc-reply>
```

See Table 40 for the tag elements and their values.

Table 40. lock *Tag Element Values*

Tag Element	Description	Value
target	Configuration that needs to be edited.	running/

<unlock>

Usage:

```
<rpc message-id="101" xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
  <unlock>
    <target>
      <running/>
    </target>
  </unlock>
</rpc>
```

Response from the switch:

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
  <ok/>
</rpc-reply>
```

See Table 41 for the tag elements and their values.

Table 41. unlock *Tag Element Values*

Tag Element	Description	Value
target	Configuration being edited.	running/

<get>

Usage:

```
<rpc message-id="101" xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <get>
     <filter type="subtree">
       <!-- request a text version of the configuration -->
       <configuration-text
       xmlns="http://www.ibm.com/netconf/1.0/config-text"/>
     </filter>
   </get>
</rpc>
```

Response from the switch:

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <data>
     <configuration-text
     xmlns="http://www.ibm.com/netconf/1.0/config-text">
       <!-- configuration text... -->
     </configuration -text>
   </data>
</rpc-reply>
```

See Table 42 for the tag elements and their values.

Table 42. get *Tag Element Values*

Tag Element	Description	Value
filter	Filter type.	subtree
configuration-text	Configuration in ISCLI format.	

<close-session>

Usage:

```
<rpc message-id="101" xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <close-session/>
</rpc>
```

Response from the switch:

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <ok/>
</rpc-reply>
```

<kill-session>

Usage:

```
<rpc message-id="101" xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <kill-session>
     <session-id>4</session-id>
   </kill-session>
</rpc>
```

Response from the switch:

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <ok/>
</rpc-reply>
```

See Table 43 for the tag elements and their values.

Table 43. kill-session *Tag Element Values*

Tag Element	Description
session-id	ID number of the session to be killed

<get-configuration>

Usage:

```
<rpc message-id="101" xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <get-configuration database="commited" format="text"/>
</rpc>
```

Response from the switch:

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <data>
     <configuration-text
     xmlns="http://www.ibm.com/netconf/1.0/config-text">
      <!-- configuration text... -->
     </configuration -text>
   </data>
</rpc-reply>
```

See Table 44 for the tag elements and their values.

Table 44. get-configuration *Tag Element Values*

Tag Element	Description	Attributes
get-configuration	Retrieve the configuration.	database - supports only committed format - supports only text

<get-interface-information>

Usage:

```
<rpc message-id="101">
  <get-interface-information>
    <interface-name> port xx </interface-name>
    <brief/>
  </get-interface-information>
</rpc>
```

Response from switch:

● Port detail information

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <interface-information>
     <physical-interface>
         <name></name>
         <admin-status></admin-status>
         <oper-status></oper-status>
         <local-index></local-index>
         <if-type></if-type>
         <link-level-type></link-level-type>
         <mtu></mtu>
         <speed></speed>
         <link-type></link-type>
         <traffic-statistics>
             <input-bytes></input-bytes>
             <output-bytes></output-bytes>
             <input-packets></input-packets>
             <output-packets></output-packets>
         </traffic-statistics>
         <input-error-list>
             <input-errors></input-errors>
             <framing-errors></framing-errors>
             <input-giants></input-giants>
             <input-discards></input-discards>
          </input-error-list>
         <output-error-list>
             <output-collisions></output-collisions>
             <output-errors></output-errors>
             <output-drops></output-drops>
         </output-error-list>
     </physical-interface>
   </interface-information>
</rpc-reply>
```

- IP detail information

```
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
   <interface-information>
     <physical-interface>
       <logical-interface>
         <name></name>
         <local-index></local-index>
         <address-family>
         <address-family-name></address-family-name>
         <mtu></mtu>
         <interface-address>
         <ifa-destination></ifa-destination>
         <ifa-local></ifa-local>
         <ifa-broadcast></ifa-broadcast>
         </interface-address>
         </address-family>
       </logical-interface>
     </physical-interface>
   </interface-information>
</rpc-reply>
```

See Table 45 for the tag elements and their values.

Table 45. `get-interface-information` *Tag Element Values*

Tag Element	Description
interface-name	Interface name or number. You can use the tags brief/ or detail/ to specify the amount of information you need.
name	Name of the port or IP interface.
admin-status	Administration status of port interface; shutdown or no shutdown.
oper-status	Operational status of port interface; link-up or link-down.
local-index	Local index of port.
if-type	Type of port; GE, XGE.
link-level-type	Ethernet
mtu	9216 for port; 1500 for IP interface.
speed	Speed of port; 1000M, 10000M.
link-type	Type of duplex port; full, half.
input-bytes	Number of bytes received at the port.
output-bytes	Number of bytes sent from the port.
input-packets	Number of frames received at port.
output-packets	Number of frames sent out from the port.

Table 45. `get-interface-information` *Tag Element Values*

Tag Element	Description
`input-errors`	Sum of discarded frames and FCS Errors.
`framing-errors`	Number of failed frames received.
`input-giants`	Number of frames that are too long.
`input-discards`	Number of frames in discarding state.
`output-collisions`	Number of Ethernet collisions.
`output-errors`	Sum of the outgoing frame aborts and FCS errors.
`output-drops`	Number of frames dropped.
`address-family-name`	Name of the address family; must be set to either "`inet`" or "`inet6`".
`ifa-destination`	Protocol network address of the interface.
`ifa-local`	Protocol host address on the interface.
`ifa-broadcast`	Network broadcast address.

Part 8: Monitoring

The ability to monitor traffic passing through the G8124-E can be invaluable for troubleshooting some types of networking problems. This sections cover the following monitoring features:

- Remote Monitoring (RMON)
- sFlow
- Port Mirroring

Chapter 35. Remote Monitoring

Remote Monitoring (RMON) allows network devices to exchange network monitoring data.

RMON allows the switch to perform the following functions:

- Track events and trigger alarms when a threshold is reached.
- Notify administrators by issuing a syslog message or SNMP trap.

RMON Overview

The RMON MIB provides an interface between the RMON agent on the switch and an RMON management application. The RMON MIB is described in RFC 1757.

The RMON standard defines objects that are suitable for the management of Ethernet networks. The RMON agent continuously collects statistics and proactively monitors switch performance. RMON allows you to monitor traffic flowing through the switch.

The switch supports the following RMON Groups, as described in RFC 1757:

- Group 1: Statistics
- Group 2: History
- Group 3: Alarms
- Group 9: Events

RMON Group 1—Statistics

The switch supports collection of Ethernet statistics as outlined in the RMON statistics MIB, in reference to etherStatsTable. You can configure RMON statistics on a per-port basis.

RMON statistics are sampled every second, and new data overwrites any old data on a given port.

Note: RMON port statistics must be enabled for the port before you can view RMON statistics.

Example Configuration

1. Enable RMON on a port.

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# rmon
```

2. View RMON statistics for the port.

```
RS G8124-E(config-if)# show interface port 1 rmon-counters
-----------------------------------------------------------------
RMON statistics for port 1:
etherStatsDropEvents:                        NA
etherStatsOctets:                       7305626
etherStatsPkts:                           48686
etherStatsBroadcastPkts:                   4380
etherStatsMulticastPkts:                   6612
etherStatsCRCAlignErrors:                    22
etherStatsUndersizePkts:                      0
etherStatsOversizePkts:                       0
etherStatsFragments:                          2
etherStatsJabbers:                            0
etherStatsCollisions:                         0
etherStatsPkts64Octets:                   27445
etherStatsPkts65to127Octets:              12253
etherStatsPkts128to255Octets:              1046
etherStatsPkts256to511Octets:               619
etherStatsPkts512to1023Octets:             7283
etherStatsPkts1024to1518Octets:              38
```

RMON Group 2—History

The RMON History Group allows you to sample and archive Ethernet statistics for a specific interface during a specific time interval. History sampling is done per port.

Note: RMON port statistics must be enabled for the port before an RMON History Group can monitor the port.

Data is stored in *buckets*, which store data gathered during discreet sampling intervals. At each configured interval, the History index takes a sample of the current Ethernet statistics, and places them into a bucket. History data buckets reside in dynamic memory. When the switch is re-booted, the buckets are emptied.

Requested buckets are the number of buckets, or data slots, requested by the user for each History Group. Granted buckets are the number of buckets granted by the system, based on the amount of system memory available. The system grants a maximum of 50 buckets.

You can use an SNMP browser to view History samples.

History MIB Object ID

The type of data that can be sampled must be of an `ifIndex` object type, as described in RFC 1213 and RFC 1573. The most common data type for the History sample is as follows:

```
1.3.6.1.2.1.2.2.1.1.<x>
```

The last digit (*x*) represents the number of the port to monitor.

Configuring RMON History

Perform the following steps to configure RMON History on a port.

1. Enable RMON on a port.

```
RS G8124-E(config)# interface port 1
RS G8124-E(config-if)# rmon
RS G8124-E(config-if)# exit
```

2. Configure the RMON History parameters for a port.

```
RS G8124-E(config)# rmon history 1 interface-oid 1.3.6.1.2.1.2.2.1.1.<x>
RS G8124-E(config)# rmon history 1 requested-buckets 30
RS G8124-E(config)# rmon history 1 polling-interval 120
RS G8124-E(config)# rmon history 1 owner "rmon port 1 history"
```

where <x> is the number of the port to monitor. For example, the full OID for port 1 would be:

```
1.3.6.1.2.1.2.2.1.1.1
```

3. View RMON history for the port.

```
RS G8124-E(config)# show rmon history
RMON History group configuration:

Index          IFOID              Interval   Rbnum  Gbnum
-----  ----------------------  --------  -----  -----
    1  1.3.6.1.2.1.2.2.1.1.1        120     30     30

Index                         Owner
-----  ----------------------------------------------
    1  rmon port 1 history
```

RMON Group 3—Alarms

The RMON Alarm Group allows you to define a set of thresholds used to determine network performance. When a configured threshold is crossed, an alarm is generated. For example, you can configure the switch to issue an alarm if more than 1,000 CRC errors occur during a 10-minute time interval.

Each Alarm index consists of a variable to monitor, a sampling time interval, and parameters for rising and falling thresholds. The Alarm Group can be used to track rising or falling values for a MIB object. The object must be a counter, gauge, integer, or time interval.

Use one of the following commands to correlate an Alarm index to an Event index:

```
RS G8124-E(config)# rmon alarm <alarm number> rising-crossing-index
        <event number>
RS G8124-E(config)# rmon alarm <alarm number> falling-crossing-index
        <event number>
```

When the alarm threshold is reached, the corresponding event is triggered.

Alarm MIB objects

The most common data types used for alarm monitoring are ifStats: errors, drops, bad CRCs, and so on. These MIB Object Identifiers (OIDs) correlate to the ones tracked by the History Group. An example statistic follows:

```
1.3.6.1.2.1.5.1.0 – mgmt.icmp.icmpInMsgs
```

This value represents the alarm's MIB OID, as a string. Note that for non-tables, you must supply a .0 to specify end node.

Configuring RMON Alarms

Configure the RMON Alarm parameters to track ICMP messages.

```
RS G8124-E(config)# rmon alarm 1 oid 1.3.6.1.2.1.5.8.0
RS G8124-E(config)# rmon alarm 1 alarm-type rising
RS G8124-E(config)# rmon alarm 1 rising-crossing-index 110
RS G8124-E(config)# rmon alarm 1 interval-time 60
RS G8124-E(config)# rmon alarm 1 rising-limit 200
RS G8124-E(config)# rmon alarm 1 sample delta
RS G8124-E(config)# rmon alarm 1 owner "Alarm for icmpInEchos"
```

This configuration creates an RMON alarm that checks icmpInEchos on the switch once every minute. If the statistic exceeds 200 within a 60 second interval, an alarm is generated that triggers event index 110.

RMON Group 9—Events

The RMON Event Group allows you to define events that are triggered by alarms. An event can be a log message, an SNMP trap, or both.

When an alarm is generated, it triggers a corresponding event notification. Use the following commands to correlate an Event index to an alarm:

```
RS G8124-E(config)# rmon alarm <alarm number> rising-crossing-index
      <event number>
RS G8124-E(config)# rmon alarm <alarm number> falling-crossing-index
      <event number>
```

RMON events use SNMP and syslogs to send notifications. Therefore, an SNMP trap host must be configured for trap event notification to work properly.

RMON uses a syslog host to send syslog messages. Therefore, an existing syslog host must be configured for event log notification to work properly. Each log event generates a syslog of type RMON that corresponds to the event.

For example, to configure the RMON event parameters.

```
RS G8124-E(config)# rmon event 110 type log
RS G8124-E(config)# rmon event 110 description "SYSLOG_this_alarm"
RS G8124-E(config)# rmon event 110 owner "log icmpInEchos alarm"
```

This configuration creates an RMON event that sends a syslog message each time it is triggered by an alarm.

Chapter 36. sFlow

The G8124-E supports sFlow technology for monitoring traffic in data networks. The switch includes an embedded sFlow agent which can be configured to provide continuous monitoring information of IPv4 traffic to a central sFlow analyzer.

The switch is responsible only for forwarding sFlow information. A separate sFlow analyzer is required elsewhere on the network to interpret sFlow data.

Note: Lenovo Networking OS 8.3 does not support IPv6 for sFlow.

sFlow Statistical Counters

The G8124-E can be configured to send network statistics to an sFlow analyzer at regular intervals. For each port, a polling interval of 5 to 60 seconds can be configured, or 0 (the default) to disable this feature.

When polling is enabled, at the end of each configured polling interval, the G8124-E reports general port statistics and port Ethernet statistics.

sFlow Network Sampling

In addition to statistical counters, the G8124-E can be configured to collect periodic samples of the traffic data received on each port. For each sample, 128 bytes are copied, UDP-encapsulated, and sent to the configured sFlow analyzer.

For each port, the sFlow sampling rate can be configured to occur once each 256 to 65536 packets, or 0 to disable (the default). A sampling rate of 256 means that one sample will be taken for approximately every 256 packets received on the port. The sampling rate is statistical, however. It is possible to have slightly more or fewer samples sent to the analyzer for any specific group of packets (especially under low traffic conditions). The actual sample rate becomes most accurate over time, and under higher traffic flow.

sFlow sampling has the following restrictions:

- Sample Rate—The fastest sFlow sample rate is 1 out of every 256 packets.

- ACLs—sFlow sampling is performed before ACLs are processed. For ports configured both with sFlow sampling and one or more ACLs, sampling will occur regardless of the action of the ACL.

- Port Mirroring—sFlow sampling will not occur on mirrored traffic. If sFlow sampling is enabled on a port that is configured as a port monitor, the mirrored traffic will not be sampled.

- Egress traffic—sFlow sampling will not occur on egress traffic.

Note: Although sFlow sampling is not generally a CPU-intensive operation, configuring fast sampling rates (such as once every 256 packets) on ports under heavy traffic loads can cause switch CPU utilization to reach maximum. Use larger rate values for ports that experience heavy traffic.

sFlow Example Configuration

1. Specify the location of the sFlow analyzer (the server and optional port to which the sFlow information will be sent):

```
RS G8124-E(config)# sflow server <IPv4 address>(sFlow server address)
RS G8124-E(config)# sflow port <service port>(Set the optional service port)
RS G8124-E(config)# sflow enable      (Enable sFlow features)
```

By default, the switch uses established sFlow service port 6343.

To disable sFlow features across all ports, use the no sflow enable command.

2. On a per-port basis, define the statistics polling rate:

```
RS G8124-E(config)# interface port <port>
RS G8124-E(config-if)# sflow polling <polling rate>(Statistics polling rate)
```

Specify a polling rate between 5 and 60 seconds, or 0 to disable. By default, polling is 0 (disabled) for each port.

3. On a per-port basis, define the data sampling rate:

```
RS G8124-E(config-if)# sflow sampling <sampling rate>(Data sampling rate)
```

Specify a sampling rate between 256 and 65536 packets, or 0 to disable. By default, the sampling rate is 0 (disabled) for each port.

4. Save the configuration.

Chapter 37. Port Mirroring

The Lenovo Networking OS port mirroring feature allows you to mirror (copy) the packets of a target port, and forward them to a monitoring port. Port mirroring functions for all layer 2 and layer 3 traffic on a port. This feature can be used as a troubleshooting tool or to enhance the security of your network. For example, an IDS server or other traffic sniffer device or analyzer can be connected to the monitoring port to detect intruders attacking the network.

The G8124-E supports a "many to one" mirroring model. As shown in Figure 55, selected traffic for ports 1 and 2 is being monitored by port 3. In the example, both ingress traffic and egress traffic on port 2 are copied and forwarded to the monitor. However, port 1 mirroring is configured so that only ingress traffic is copied and forwarded to the monitor. A device attached to port 3 can analyze the resulting mirrored traffic.

Figure 55. Mirroring Ports
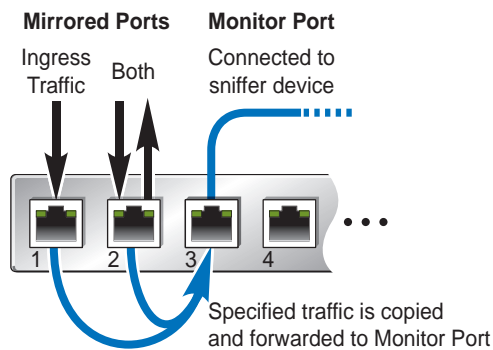
The G8124-E supports four monitor ports. Each monitor port can receive mirrored traffic from any number of target ports.

Lenovo N/OS does not support "one to many" or "many to many" mirroring models where traffic from a specific port traffic is copied to multiple monitor ports. For example, port 1 traffic cannot be monitored by both port 3 and 4 at the same time, nor can port 2 ingress traffic be monitored by a different port than its egress traffic.

Ingress and egress traffic is duplicated and sent to the monitor port after processing.

Configuring Port Mirroring

The following procedure may be used to configure port mirroring for the example shown in :

1. Specify the monitoring port, the mirroring port(s), and the port-mirror direction.

```
RS G8124-E(config)# port-mirroring monitor-port 3 mirroring-port 1 in
RS G8124-E(config)# port-mirroring monitor-port 3 mirroring-port 2 both
```

2. Enable port mirroring.

```
RS G8124-E(config)# port-mirroring enable
```

3. View the current configuration.

```
RS G8124-E# show port-mirroring

Port Monitoring : Enabled

Monitoring Ports      Mirrored Ports
3                         1, in
                          2, both
```

Part 9: Appendices

- Getting help and technical assistance
- Notices

Appendix A. Getting help and technical assistance

If you need help, service, or technical assistance or just want more information about Lenovo products, you will find a wide variety of sources available from Lenovo to assist you.

Use this information to obtain additional information about Lenovo and Lenovo products, and determine what to do if you experience a problem with your Lenovo system or optional device.

Note: This section includes references to IBM web sites and information about obtaining service. IBM is Lenovo's preferred service provider for the System x, Flex System, and NeXtScale System products.

Before you call, make sure that you have taken these steps to try to solve the problem yourself.

If you believe that you require warranty service for your Lenovo product, the service technicians will be able to assist you more efficiently if you prepare before you call.

- Check all cables to make sure that they are connected.

- Check the power switches to make sure that the system and any optional devices are turned on.

- Check for updated software, firmware, and operating-system device drivers for your Lenovo product. The Lenovo Warranty terms and conditions state that you, the owner of the Lenovo product, are responsible for maintaining and updating all software and firmware for the product (unless it is covered by an additional maintenance contract). Your service technician will request that you upgrade your software and firmware if the problem has a documented solution within a software upgrade.

- If you have installed new hardware or software in your environment, check the IBM ServerProven website to make sure that the hardware and software is supported by your product.

- Go to the IBM Support portal to check for information to help you solve the problem.

- Gather the following information to provide to the service technician. This data will help the service technician quickly provide a solution to your problem and ensure that you receive the level of service for which you might have contracted.

  - Hardware and Software Maintenance agreement contract numbers, if applicable
  - Machine type number (if applicable–Lenovo 4-digit machine identifier)
  - Model number
  - Serial number
  - Current system UEFI and firmware levels
  - Other pertinent information such as error messages and logs

- Start the process of determining a solution to your problem by making the pertinent information available to the service technicians. The IBM service technicians can start working on your solution as soon as you have completed and submitted an Electronic Service Request.

You can solve many problems without outside assistance by following the troubleshooting procedures that Lenovo provides in the online help or in the Lenovo product documentation. The Lenovo product documentation also describes the diagnostic tests that you can perform. The documentation for most systems, operating systems, and programs contains troubleshooting procedures and explanations of error messages and error codes. If you suspect a software problem, see the documentation for the operating system or program.

Appendix B. Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area.

Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service.

Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc.
1009 Think Place - Building One
Morrisville, NC 27560
U.S.A.

Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties.

Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary.

Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk.

513

Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Trademarks

Lenovo, the Lenovo logo, Flex System, System x, NeXtScale System, and X-Architecture are trademarks of Lenovo in the United States, other countries, or both.

Intel and Intel Xeon are trademarks of Intel Corporation in the United States, other countries, or both.

Internet Explorer, Microsoft, and Windows are trademarks of the Microsoft group of companies.

Linux is a registered trademark of Linus Torvalds.

Other company, product, or service names may be trademarks or service marks of others.

Important Notes

Processor speed indicates the internal clock speed of the microprocessor; other factors also affect application performance.

CD or DVD drive speed is the variable read rate. Actual speeds vary and are often less than the possible maximum.

When referring to processor storage, real and virtual storage, or channel volume, KB stands for 1 024 bytes, MB stands for 1 048 576 bytes, and GB stands for 1 073 741 824 bytes.

When referring to hard disk drive capacity or communications volume, MB stands for 1 000 000 bytes, and GB stands for 1 000 000 000 bytes. Total user-accessible capacity can vary depending on operating environments.

Maximum internal hard disk drive capacities assume the replacement of any standard hard disk drives and population of all hard-disk-drive bays with the largest currently supported drives that are available from Lenovo.

Maximum memory might require replacement of the standard memory with an optional memory module.

Each solid-state memory cell has an intrinsic, finite number of write cycles that the cell can incur. Therefore, a solid-state device has a maximum number of write cycles that it can be subjected to, expressed as total bytes written (TBW). A device that has exceeded this limit might fail to respond to system-generated commands or might be incapable of being written to. Lenovo is not responsible for replacement of a device that has exceeded its maximum guaranteed number of program/erase cycles, as documented in the Official Published Specifications for the device.

Lenovo makes no representations or warranties with respect to non-Lenovo products. Support (if any) for the non-Lenovo products is provided by the third party, not Lenovo.

Some software might differ from its retail version (if available) and might not include user manuals or all program functionality.

Recycling Information

Lenovo encourages owners of information technology (IT) equipment to responsibly recycle their equipment when it is no longer needed. Lenovo offers a variety of programs and services to assist equipment owners in recycling their IT products. For information on recycling Lenovo products, go to:

http://www.lenovo.com/recycling

Particulate Contamination

Attention: Airborne particulates (including metal flakes or particles) and reactive gases acting alone or in combination with other environmental factors such as humidity or temperature might pose a risk to the device that is described in this document.

Risks that are posed by the presence of excessive particulate levels or concentrations of harmful gases include damage that might cause the device to malfunction or cease functioning altogether. This specification sets forth limits for particulates and gases that are intended to avoid such damage. The limits must not be viewed or used as definitive limits, because numerous other factors, such as temperature or moisture content of the air, can influence the impact of particulates or environmental corrosives and gaseous contaminant transfer. In the absence of specific limits that are set forth in this document, you must implement practices that maintain particulate and gas levels that are consistent with the protection of human health and safety. If Lenovo determines that the levels of particulates or gases in your environment have caused damage to the device, Lenovo may condition provision of repair or replacement of devices or parts on implementation of appropriate remedial measures to mitigate such environmental contamination. Implementation of such remedial measures is a customer responsibility..

Contaminant	Limits
Particulate	• The room air must be continuously filtered with 40% atmospheric dust spot efficiency (MERV 9) according to ASHRAE Standard 52.2[1]. • Air that enters a data center must be filtered to 99.97% efficiency or greater, using high-efficiency particulate air (HEPA) filters that meet MIL-STD-282. • The deliquescent relative humidity of the particulate contamination must be more than 60%[2]. • The room must be free of conductive contamination such as zinc whiskers.
Gaseous	• Copper: Class G1 as per ANSI/ISA 71.04-1985[3] • Silver: Corrosion rate of less than 300 Å in 30 days
[1] ASHRAE 52.2-2008 - *Method of Testing General Ventilation Air-Cleaning Devices for Removal Efficiency by Particle Size.* Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. [2] The deliquescent relative humidity of particulate contamination is the relative humidity at which the dust absorbs enough water to become wet and promote ionic conduction. [3] ANSI/ISA-71.04-1985. *Environmental conditions for process measurement and control systems: Airborne contaminants.* Instrument Society of America, Research Triangle Park, North Carolina, U.S.A.	

Telecommunication Regulatory Statement

This product may not be certified in your country for connection by any means whatsoever to interfaces of public telecommunications networks. Further certification may be required by law prior to making any such connection. Contact a Lenovo representative or reseller for any questions.

Electronic Emission Notices

When you attach a monitor to the equipment, you must use the designated monitor cable and any interference suppression devices that are supplied with the monitor.

Federal Communications Commission (FCC) Statement

Note: This equipment has been tested and found to comply with the limits for a Class A digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference, in which case the user will be required to correct the interference at his own expense.

Properly shielded and grounded cables and connectors must be used to meet FCC emission limits. Lenovo is not responsible for any radio or television interference caused by using other than recommended cables and connectors or by unauthorized changes or modifications to this equipment. Unauthorized changes or modifications could void the user's authority to operate the equipment.

This device complies with Part 15 of the FCC Rules. Operation is subject to the following two conditions: (1) this device may not cause harmful interference, and (2) this device must accept any interference received, including interference that might cause undesired operation.

Industry Canada Class A Emission Compliance Statement

This Class A digital apparatus complies with Canadian ICES-003.

Avis de Conformité à la Réglementation d'Industrie Canada

Cet appareil numérique de la classe A est conforme à la norme NMB-003 du Canada.

Australia and New Zealand Class A Statement

Attention: This is a Class A product. In a domestic environment this product may cause radio interference in which case the user may be required to take adequate measures.

European Union EMC Directive Conformance Statement

This product is in conformity with the protection requirements of EU Council Directive 2004/108/EC on the approximation of the laws of the Member States relating to electromagnetic compatibility. Lenovo cannot accept responsibility for any failure to satisfy the protection requirements resulting from a non-recommended modification of the product, including the installation of option cards from other manufacturers.

This product has been tested and found to comply with the limits for Class A Information Technology Equipment according to European Standard EN 55022. The limits for Class A equipment were derived for commercial and industrial environments to provide reasonable protection against interference with licensed communication equipment.

Lenovo, Einsteinova 21, 851 01 Bratislava, Slovakia

Germany Class A Statement

Zulassungsbescheinigung laut dem Deutschen Gesetz über die elektromagnetische Verträglichkeit von Betriebsmitteln, EMVG vom 20. Juli 2007 (früher Gesetz über die elektromagnetische Verträglichkeit von Geräten), bzw. der EMV EG Richtlinie 2004/108/EC (früher 89/336/EWG), für Geräte der Klasse A.

Dieses Gerät ist berechtigt, in übereinstimmung mit dem Deutschen EMVG das EG-Konformitätszeichen - CE - zu führen. Verantwortlich für die Konformitätserklärung nach Paragraf 5 des EMVG ist die Lenovo (Deutschland) GmbH, Gropiusplatz 10, D-70563 Stuttgart.

Informationen in Hinsicht EMVG Paragraf 4 Abs. (1) 4:

Das Gerät erfüllt die Schutzanforderungen nach EN 55024 und EN 55022 Klasse A.

Nach der EN 55022: "Dies ist eine Einrichtung der Klasse A. Diese Einrichtung kann im Wohnbereich Funkstörungen verursachen; in diesem Fall kann vom Betreiber verlangt werden, angemessene Maßnahmen durchzuführen und dafür aufzukommen."

Nach dem EMVG: Dieses Produkt entspricht den Schutzanforderungen der EU-Richtlinie 2004/108/EG (früher 89/336/EWG) zur Angleichung der Rechtsvorschriften über die elektromagnetische Verträglichkeit in den EU-Mitgliedsstaaten und hält die Grenzwerte der EN 55022 Klasse A ein.

Um dieses sicherzustellen, sind die Geräte wie in den Handbüchern beschrieben zu installieren und zu betreiben. Des Weiteren dürfen auch nur von der Lenovo empfohlene Kabel angeschlossen werden. Lenovo übernimmt keine Verantwortung für die Einhaltung der Schutzanforderungen, wenn das Produkt ohne Zustimmung der Lenovo verändert bzw. wenn Erweiterungskomponenten von Fremdherstellern ohne Empfehlung der Lenovo gesteckt/eingebaut werden.

Deutschland:

Einhaltung des Gesetzes über die elektromagnetische Verträglichkeit von Betriebsmittein
Dieses Produkt entspricht dem "Gesetz über die elektromagnetische Verträglichkeit von Betriebsmitteln" EMVG (früher "Gesetz über die elektromagnetische Verträglichkeit von Geräten"). Dies ist die Umsetzung der EU-Richtlinie 2004/108/EG (früher 89/336/EWG) in der Bundesrepublik Deutschland.

Zulassungsbescheinigung laut dem Deutschen Gesetz über die elektromagnetische Verträglichkeit von Betriebsmitteln, EMVG vom 20. Juli 2007 (früher Gesetz über die elektromagnetische Verträglichkeit von Geräten), bzw. der EMV EG Richtlinie 2004/108/EC (früher 89/336/EWG), für Geräte der Klasse A.

Dieses Gerät ist berechtigt, in übereinstimmung mit dem Deutschen EMVG das EG-Konformitätszeichen - CE - zu führen. Verantwortlich für die Konformitätserklärung nach Paragraf 5 des EMVG ist die Lenovo (Deutschland) GmbH, Gropiusplatz 10, D-70563 Stuttgart.

Informationen in Hinsicht EMVG Paragraf 4 Abs. (1) 4:
Das Gerät erfüllt die Schutzanforderungen nach EN 55024 und EN 55022 Klasse A.

Nach der EN 55022: "Dies ist eine Einrichtung der Klasse A. Diese Einrichtung kann im Wohnbereich Funkstörungen verursachen; in diesem Fall kann vom Betreiber verlangt werden, angemessene Maßnahmen durchzuführen und dafür aufzukommen."

Nach dem EMVG: "Geräte dürfen an Orten, für die sie nicht ausreichend entstört sind, nur mit besonderer Genehmigung des Bundesministers für Post und Telekommunikation oder des Bundesamtes für Post und Telekommunikation betrieben werden. Die Genehmigung wird erteilt, wenn keine elektromagnetischen Störungen zu erwarten sind." (Auszug aus dem EMVG, Paragraph 3, Abs. 4). Dieses Genehmigungsverfahren ist nach Paragraph 9 EMVG in Verbindung mit der entsprechenden Kostenverordnung (Amtsblatt 14/93) kostenpflichtig.

Anmerkung: Um die Einhaltung des EMVG sicherzustellen sind die Geräte, wie in den Handbüchern angegeben, zu installieren und zu betreiben.

Japan VCCI Class A Statement

この装置は、クラス A 情報技術装置です。この装置を家庭環境で使用すると電波妨害を引き起こすことがあります。この場合には使用者が適切な対策を講ずるよう要求されることがあります。　　　　VCCI-A

This is a Class A product based on the standard of the Voluntary Control Council for Interference (VCCI). If this equipment is used in a domestic environment, radio interference may occur, in which case the user may be required to take corrective actions.

Japan Electronics and Information Technology Industries Association (JEITA) Statement

高調波ガイドライン適合品

Japan Electronics and Information Technology Industries Association (JEITA) Confirmed Harmonics Guidelines (products less than or equal to 20 A per phase)

高調波ガイドライン準用品

Japan Electronics and Information Technology Industries Association (JEITA) Confirmed Harmonics Guidelines with Modifications (products greater than 20 A per phase).

Korea Communications Commission (KCC) Statement

이 기기는 업무용(A급)으로 전자파적합기기로
서 판매자 또는 사용자는 이 점을 주의하시기
바라며, 가정외의 지역에서 사용하는 것을 목
적으로 합니다.

This is electromagnetic wave compatibility equipment for business (Type A). Sellers and users need to pay attention to it. This is for any areas other than home.

Russia Electromagnetic Interference (EMI) Class A statement

ВНИМАНИЕ! Настоящее изделие относится к классу А.
В жилых помещениях оно может создавать радиопомехи, для
снижения которых необходимы дополнительные меры

People's Republic of China Class A electronic emission statement

中华人民共和国"A类"警告声明

声 明
此为A级产品，在生活环境中，该产品可能会造成无线电干扰。在这种情况下，可能需要用户对其干扰采取切实可行的措施。

Taiwan Class A compliance statement

警告使用者:
這是甲類的資訊產品，在
居住的環境中使用時，可
能會造成射頻干擾，在這
種情況下，使用者會被要
求採取某些適當的對策。