

Lenovo Network

Application Guide

for Lenovo Cloud Network Operating System 10.1

LenovoTM

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.

First Edition (June 2016)

© Copyright Lenovo 2016
Portions © Copyright IBM Corporation 2014.

LIMITED AND RESTRICTED RIGHTS NOTICE: If data or software is delivered pursuant a General Services Administration "GSA" contract, use, reproduction, or disclosure is subject to restrictions set forth in Contract No. GS-35F-05925.

Lenovo and the Lenovo logo are trademarks of Lenovo in the United States, other countries, or both.

Contents

Preface	17
Who Should Use This Guide18
Application Guide Overview19
Additional References22
Typographic Conventions23
Part 1: Getting Started	25
Chapter 1. Switch Administration	27
Administration Interfaces28
Industry Standard Command Line Interface28
Establishing a Connection29
Using the Switch Management Interface29
Using the Switch Ethernet Ports30
Using Telnet31
Using Secure Shell.32
Using SSH with Password Authentication32
Using SSH with Public Key Authentication33
Using Simple Network Management Protocol.34
DHCP IP Address Services.35
DHCP Client Configuration35
DHCPv4 Hostname Configuration (Option 12)36
DHCPv4 Syslog Server (Option 7)36
DHCPv4 NTP Server (Option 42)37
DHCPv4 Vendor Class Identifier (Option 60)37
DHCP Relay Agent38
DHCPv4 Option 8239
Switch Login Levels.40
Ping42
Ping Configurable Parameters43
Test Interruption43
Ping Count43
Ping Packet Interval43
Ping Packet Size.44
Ping Source.44
Ping DF-Bit44
Ping Timeout45
Ping VRF.45
Ping Interactive Mode46
Traceroute47
Traceroute Configurable Parameters.48
Test Interruption48
Traceroute Source48
Traceroute VRF48
Traceroute Interactive Mode49

Network Time Protocol	50
NTP Synchronization Retry	50
NTP Client and Peer	51
NTP Authentication Field Encryption Key	52
NTP Polling Intervals	52
NTP Preference	53
Dynamic and Static NTP Servers	53
NTP Authentication	53
NTP Authentication Configuration Example	54
System Logging	55
Syslog Outputs	56
Syslog Severity Levels	57
Syslog Time Stamping	58
Syslog Rate Limit	58
Syslog Servers	59
Idle Disconnect	60
Python Scripting	61
REST API Programming	62
Chapter 2. System License Keys	63
Obtaining License Keys	64
Installing License Keys	65
Uninstalling License Keys	66
Transferring License Keys	67
ONIE License Key	68
Chapter 3. Switch Software Management	69
Installing New Software to Your Switch	70
Installing System Images from a Remote Server	70
Installing System Images from a USB Device	72
Installing U-boot from a Remote Server	73
Installing U-boot from a USB Device	74
Selecting a Software Image to Run	75
Reloading the Switch	76
Copying Configuration Files	77
Copy Configuration Files via a Remote Server	77
Copy Configuration Files to a USB Device	77
The Boot Management Menu	78
Boot Recovery Mode	79
Recover from a Failed Image Upgrade using TFTP	80
Recovering from a Failed Image Upgrade using XModem Download	82
Physical Presence	84
ONIE submenu	85

ONIE86
Installing ONIE from a Remote Server86
Installing ONIE from a USB Device87
Booting in ONIE Mode.88
Booting in ONIE Install Mode.88
Booting in ONIE Uninstall Mode89
Booting in ONIE Update Mode89
Booting in ONIE Rescue Mode89
Part 2: Securing the Switch	91
Chapter 4. Securing Administration	93
Secure Shell and Secure Copy94
SSH Encryption and Authentication95
Generating RSA/DSA Host Key for SSH Access95
SSH Integration with TACACS+ Authentication.95
Configuring SSH on the Switch96
Using SSH Client Commands.97
To Log In to the Switch.97
Using Secure Copy98
Copying a File Using SCP.98
Copying Startup Configuration Using SCP98
Copying Running Configuration Using SCP98
Copying Technical Support Using SCP98
End User Access Control.99
Considerations for Configuring End User Accounts99
Strong Passwords99
User Access Control	100
Setting up Users.	100
Defining a User's Access Level	101
Deleting a User	101
The Default User	102
Administrator Password Recovery.	102
Chapter 5. Authentication & Authorization Protocols	105
TACACS+ Authentication	106
How TACACS+ Authentication Works.	106
TACACS+ Authentication Features in Cloud NOS.	107
Authorization.	107
Accounting	107
Configuring TACACS+ Authentication on the Switch	108
Authentication, Authorization and Accounting	109
AAA Groups	110
Group Lists.	110
Configuring AAA Groups	111
Authentication	112
Configuring AAA Authentication	112
Authorization	114
Configuring AAA Authorization	114
Accounting.	115
Configuring AAA Accounting.	115

Chapter 6. Access Control Lists	117
Supported ACL Types	118
Summary of Packet Classifiers	119
Summary of ACL Actions	121
Configuring Port ACLs (PACLs)	122
Configuring Router ACLs (RACLs)	123
Configuring VLAN ACLs (VACLs)	124
VACL Configuration Example	125
Configuring Management ACLs (MACLs)	126
ACL Order of Precedence	127
Creating and Modifying ACLs	128
Creating an IPv4 ACL	129
Removing an IPv4 ACL	129
Resequencing an IPv4 ACL	129
Creating a MAC ACL	130
Removing a MAC ACL	130
Resequencing a MAC ACL	130
Creating an ARP ACL	131
Removing an ARP ACL	131
Resequencing an ARP ACL	131
Viewing ACL Rule Statistics	132
ACL Configuration Examples	133
ACL Example 1	133
ACL Example 2	133
ACL Example 3	134
ACL Example 4	134
ACL Example 5	135
ACL Example 6	135
 Part 3: Switch Basics	 137
 Chapter 7. Interface Management	 139
Interface Management Overview	140
Management Interface	141
Physical Ports	142
Port Aggregation	144
Loopback Interfaces	145
Switch Virtual Interfaces	146
Basic Interface Configuration	147
Interface Description	148
Interface Duplex	148
Interface MAC Address	149
Interface Maximum Transmission Unit	149
Interface Shutdown	150
Interface Speed	150
Flow Control	151
Storm Control	152

Chapter 8. Forwarding Database.	.153
MAC Learning	154
Static MAC addresses	155
Aging Time	156
Chapter 9. VLANs	.157
VLAN Overview	158
VLAN Configuration	159
Creating a VLAN	160
Deleting a VLAN	161
Configuring the State of a VLAN	162
Configuring the Name of a VLAN.	164
Configuring a Switch Access Port	165
Configuring the Access VLAN	165
Configuring a Switch Trunk Port	167
Configuring the Allowed VLAN List.	167
Configuring the Native VLAN	168
Native VLAN Tagging	170
Configuring Native VLAN Tagging	171
IPMC Flooding	173
VLAN Topologies and Design Considerations	174
Multiple VLANs with Trunk Mode Adapters	175
VLAN Configuration Example	177
Chapter 10. Ports and Link Aggregation	.179
G8272 Port Configuration Profiles.	180
Aggregation Overview	183
Creating a LAG	183
Static LAGs	184
Static LAG Configuration Rules	184
Configuring a Static LAG	185
Link Aggregation Control Protocol	188
Configuring LACP	188
System Priority	189
Port Priority	189
LACP Timeout	190
LACP Individual	190
LACP Configuration Example.	191
LAG Hashing	193
LAG Hashing Configuration	195
Chapter 11. Multiple Spanning Tree Protocol	.197
MSTP Overview	198
Global MSTP Control	199
Common Internal Spanning Tree	199
Port States	199

Bridge Protocol Data Units	200
Determining the Path for Forwarding BPDUs.	200
BPDU Guard	200
BPDU Filter	201
Port Path Cost	201
Bridge Priority	202
Port Priority	202
Root Guard.	202
Error Disable Recovery	203
MST Region	204
Global MST Instance Parameters	205
Hop Count.	205
Forward Delay	205
Hello Timer	206
Maximum Age Interval	206
Port Type and Link Type	207
Edge Port	207
Link Type	207
MSTP Configuration Guidelines	208
MSTP Configuration Example	208
Chapter 12. Virtual Link Aggregation Groups	211
vLAG Overview	212
vLAG Capacities	214
vLAG Benefits	214
vLAG Synchronization Mechanism	215
vLAG System MAC	215
vLAG and LACP Individual	216
vLAG and LACP System Priority	216
vLAG LACP Misconfigurations or Cabling Errors.	216
FDB Synchronization	217
vLAG and MSTP	217
vLAGs versus regular LAGs	219
Configuring vLAGs.	220
vLAG ISL	221
vLAG Role Election	221
vLAG Instance	222
FDB Refresh	223
vLAG Startup Delay	223
vLAG Auto-recovery	224
vLAG Tier ID	224
Health Check	225
Basic Health Check Configuration Example	226
Basic vLAG Configuration Example.	227
Configuring the ISL	228
Configuring the vLAG.	229
vLAG Configuration - VLANs Mapped to a MST Instance	230
Configuring the ISL	230
Configuring the vLAG.	231

Configuring vLAGs in Multiple Layers	232
Task 1: Configure Layer 2/3 Border Region	233
Configuring Border Router 1	233
Configuring Border Router 2	233
Task 2: Configure switches in the Layer 2 region	234
Configuring Switch A	234
Configuring Switch B	235
Configuring Switches C and D	237
Configuring Switch E	238
Configuring Switch F	239
Chapter 13. Quality of Service	241
QoS Overview	242
Class Maps	243
QoS Classification Types	244
Using ACL Filters	244
Using Class of Service Filters	245
Using DiffServ Code Point (DSCP) Filters.	246
Using TCP/UDP Port Filters.	248
Using Precedence Filters	248
Using Protocol Filters	249
Queuing Classification Types	250
Class Map Configuration Examples	251
QoS Class Map Configuration Example.	251
Queueing Class Map Configuration Example	251
Policy Maps	252
Ingress Policing.	252
Defining Single-Rate and Dual-Rate Policers	252
Marking	254
Queuing Policing	254
Bandwidth	254
Shaping	254
Priority.	254
Policy Map Configuration Example	255
QoS Policy Map Configuration Example	255
Queueing Policy Map Configuration Example	256
Control Plane Protection.	257
Control Plane Configuration Examples.	258
WRED	260
Configuring WRED	260
WRED Configuration Example	260
Interface Service Policy	262
Limitations.	262
Microburst Detection	263

Part 4: IP Routing.	265
Chapter 14. Basic IP Routing	267
IP Routing.	268
Direct and Indirect Routing	269
Static and Dynamic Routing	269
Default Gateway	270
Routing Information Base	271
Routes with Indirect Next-hops.	271
Bidirectional Forwarding Detection.	272
BFD Asynchronous Mode	273
BFD Echo Mode	273
BFD and BGP	274
BFD Peer Support.	274
BFD Static Routes.	274
BFD Authentication	275
Generalized TTL Security Mechanism	276
Routing Between IP Subnets	277
Example of Subnet Routing	278
Using VLANs to Segregate Broadcast Domains	279
Configuration Example	279
ECMP Static Routes.	282
RIB Support for ECMP Routes	282
ECMP Hashing.	282
Configuring ECMP Static Routes	283
Dynamic Host Configuration Protocol.	284
Internet Control Message Protocol	285
ICMP Redirects.	286
ICMP Port Unreachable	286
ICMP Unreachable (except Port)	286
Chapter 15. Routed Ports	287
Routed Ports Overview	288
Configuring a Routed Port.	290
Configuring OSPF on Routed Ports	291
OSPF Configuration Example.	291
Chapter 16. Address Resolution Protocol.	293
ARP Overview	294
ARP Aging Timer	295
ARP Inspection	296
Static ARP Entries	297
Static ARP Configuration Example	297
ARP Entry States	298
ARP Table Refresh	299

Chapter 17. Internet Protocol Version 6	301
IPv6 Address Format	302
IPv6 Address Types.	303
Unicast Address	303
Multicast.	303
Anycast	304
IPv6 Interfaces	305
Neighbor Discovery.	306
Neighbor Discovery Overview	306
Router.	307
Supported Applications	308
Configuration Guidelines	309
IPv6 Configuration Examples	310
IPv6 Example 1	310
IPv6 Example 2	310
IPv6 Limitations	311
Chapter 18. Internet Group Management Protocol	313
IGMP Terms	314
How IGMP Works	315
IGMP Capacity and Default Values	316
IGMP Snooping	317
IGMPv3 Snooping.	318
Spanning Tree Topology Change	318
IGMP Querier	319
Querier Election.	319
Multicast Router Discovery.	322
IGMP Query Messages.	322
IGMP Groups	323
IGMP Snooping Configuration Guidelines	324
IGMP Snooping Configuration Example	325
Advanced IGMP Snooping Configuration Example	327
Prerequisites	328
Configuration	329
Switch A Configuration	329
Switch B Configuration.	330
Switch C Configuration	331
Troubleshooting	332
Additional IGMP Features	335
Report Suppression	335
Robustness Variable.	335
Fast Leave	336
Static Multicast Router.	337

Chapter 19. Border Gateway Protocol	339
BGP Overview	340
BGP Router Identifier	340
Internal Routing Versus External Routing	341
Route Reflector	343
Route Reflection Configuration Example	344
Restrictions	345
Forming BGP Peer Routers	346
BGP Peers and Dynamic Peers	346
Static Peers	346
Dynamic Peers	347
Loopback Interfaces	348
What is a Route Map?	349
Next Hop Peer IP Address	350
Incoming and Outgoing Route Maps	350
Precedence	350
Configuration Overview	351
Aggregating Routes	352
Redistributing Routes	353
BGP Communities	355
BGP Community	356
BGP Extended Community	357
BGP Confederation	358
BGP Path Attributes	359
Well-Known Mandatory	359
Well-Known Discretionary	359
Optional Transitive	359
Optional Non-Transitive	360
Best Path Selection Logic	361
BGP Best Path Selection	361
BGP Weight	362
Local Preference	362
Metric (Multi-Exit Discriminator) Attribute	362
Next Hop	363
Best Path Selection Tuning	363
BGP ECMP	365
BGP Features and Functions	366
AS-Path Filter	366
BGP Capability Code	366
Administrative Distance	366
TTL-Security Check	367
Local-AS	367
BGP Authentication	368
Originate Default Route	368
IP Prefix-List Filter	369
Dynamic Capability	370
BGP Graceful Restart	370
BGP Damping	371
Soft Reconfiguration Inbound	372

BGP Route Refresh	372
BGP Multiple Address Families	373
BGP BFD.	373
BGP Next Hop Tracking	374
BGP Tuning	374
BGP Failover Configuration	375
Default Redistribution and Route Aggregation Example	377
Designing a Clos Network Using BGP.	379
Clos Network BGP Configuration Example.	380
Configure Fabric Switch SF1	381
Configure Spine Switch SP11	383
Configure Leaf Switch LP11.	386
Chapter 20. Open Shortest Path First.	389
OSPFv2 Overview	390
Types of OSPF Areas	391
Types of OSPF Routing Devices.	392
Neighbors and Adjacencies.	393
The Link-State Database	393
The Shortest Path First Tree	394
Internal Versus External Routing	394
OSPFv2 Implementation in Cloud NOS	395
Configurable Parameters.	395
Defining Areas	396
Using the Area ID to Assign the OSPF Area Number.	396
Attaching an Area to a Network	396
Interface Cost.	397
Electing the Designated Router and Backup	397
Summarizing Routes	397
Default Routes	398
Virtual Links	399
Router ID	400
Authentication	400
Configuring Plain Text OSPF Passwords	401
Configuring MD5 Authentication	402
Loopback Interfaces in OSPF	402
Graceful Restart Helper	403
OSPFv2 Configuration Examples	404
Example 1: Simple OSPF Domain	405
Example 2: Virtual Links	406
Example 3: Summarizing Routes	409
Verifying OSPF Configuration	410
Chapter 21. Route Maps	411
Route Maps Overview.	412
Permit and Deny Rules	413
Match and Apply Clauses	414
Route Maps Configuration Example.	417

Part 5: High Availability Fundamentals	419
Chapter 22. Basic Redundancy	421
Aggregating for Link Redundancy	422
Virtual Link Aggregation	423
Chapter 23. Virtual Router Redundancy Protocol	425
VRRP Overview	426
VRRP Components	427
Virtual Router	427
Virtual Router MAC Address	427
Owners and Renters	427
Master and Backup Virtual Router.	427
Virtual Interface Router	427
Assigning VRRP Virtual Router ID	428
VRRP Operation	428
Selecting the Master VRRP Router.	428
Failover Methods.	429
Active-Active Redundancy.	429
Cloud NOS Extensions to VRRP	430
VRRP Advertisement Interval and Sub-second Failover	430
Interface Tracking.	431
Switch Back Delay	431
Backward Compatibility with VRRPv2	432
VRRP Accept Mode	432
VRRP Preemption	433
VRRP Priority	433
IPv6 VRRP.	434
Virtual Router Deployment Considerations	436
Configuring the Switch for Tracking.	436
Basic VRRP Configuration.	437
High Availability Configuration	439
VRRP High-Availability Using Multiple VIRs	439
Task 1: Configure Switch 1	440
Task 2: Configure Switch 2	441
Part 6: Network Management	443
Chapter 24. Link Layer Discovery Protocol	445
LLDP Overview	446
Enabling or Disabling LLDP	447
Transmit and Receive Control	447
LLDP Transmit Features.	448
Scheduled Interval	448
Minimum Interval	448
Time-to-Live for Transmitted Information	449
Trap Notifications	449
Changing the LLDP Transmit State	450
Types of Information Transmitted.	451

LLDP Receive Features	452
Types of Information Received	452
Viewing Remote Device Information	452
Time-to-Live for Received Information	453
Debugging LLDP	454
LLDP Debugging Types	454
LLDP Example Configuration	456
Chapter 25. Simple Network Management Protocol	459
SNMP Version 1 & Version 2	460
SNMP Notifications	460
SNMP Device Contact and Location	460
One-Time Authentication for SNMP over TCP	460
SNMP Version 3	461
Default Configuration	462
Configuration Examples	463
Basic SNMP Configuration Example	463
User Configuration Example	463
Configuring SNMP Trap Hosts	464
SNMP MIBs	465
Part 7: Monitoring	467
Chapter 26. Port Mirroring	469
Port Mirroring Overview	470
SPAN Configuration	471
Sources	471
Destinations	471
Sessions	471
Configuration Example	472
ERSPAN Configuration	473
Session Types	473
Sources	474
Destinations	474
ERSPAN Source Session Configuration Example	475
ERSPAN Destination Session Configuration Example	476
Limitations	477
Part 8: Appendices	479
Appendix A. Getting help and technical assistance	481
Appendix B. Notices	483
Trademarks	485
Important Notes	486
Recycling Information	487
Particulate Contamination	488
Telecommunication Regulatory Statement	489

Electronic Emission Notices	490
Federal Communications Commission (FCC) Statement	490
Industry Canada Class A Emission Compliance Statement	490
Avis de Conformité à la Réglementation d'Industrie Canada	490
Australia and New Zealand Class A Statement	490
European Union - Compliance to the Electromagnetic Compatibility Directive	490
Germany Class A Statement	491
Japan VCCI Class A Statement	492
Japan Electronics and Information Technology Industries Association	
(JEITA) Statement.	492
Korea Communications Commission (KCC) Statement.	493
Russia Electromagnetic Interference (EMI) Class A statement	493
People's Republic of China Class A electronic emission statement	493
Taiwan Class A compliance statement	493
Index	495

Preface

This *Application Guide* describes how to configure and use the Cloud Network OS 10.1 software on the Lenovo RackSwitch G8272 (referred to as G8272 throughout this document). For documentation on installing the switch physically, see the *Lenovo RackSwitch G8272 Installation Guide*.

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.

Application Guide Overview

This guide will help you plan, implement, and administer Cloud NOS 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 CNOS products with the basics of switch management. This part includes the following chapters:

- [Chapter 1, “Switch Administration”](#), describes how to access the G8272 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 or Secure Shell.
- [Chapter 2, “System License Keys”](#), describes how to install additional features on the switch.
- [Chapter 3, “Switch Software Management”](#), describes how to update the CNOS software operating on the switch.

Part 2: Securing the Switch

This material contains information about implementing security protocols on the switch. This part includes the following chapters:

- [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 methods for remote administrators. This includes using Terminal Access Controller Access-Control System Plus (TACACS+) and Authentication, Authorization, and Accounting (AAA).
- [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

This material contains information about setting up features on the switch. This part includes the following chapters:

- [Chapter 7, “Interface Management”](#), describes how to configure the switch interfaces, like the ethernet or management ports.
- [Chapter 8, “Forwarding Database”](#), describes how a Layer 2 device can be configured to learn and store MAC addresses and their corresponding ports.
- [Chapter 9, “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.
- [Chapter 10, “Ports and Link Aggregation”](#), describes how to group multiple physical ports together to aggregate the bandwidth between large-scale network devices.

- [Chapter 11, “Multiple Spanning Tree Protocol”](#), describes how to use the Multiple Spanning Tree Protocol (MSTP) to build a loop-free network topology.
- [Chapter 12, “Virtual Link Aggregation Groups”](#), describes using Virtual Link Aggregation Groups (VLAGs) to form LAGs spanning multiple VLAG-capable aggregator switches.
- [Chapter 13, “Quality of Service”](#), discusses Quality of Service (QoS) features, including IP filtering using class maps, Differentiated Services, and IEEE 802.1p priority values.

Part 4: IP Routing

This part includes the following chapters:

- [Chapter 14, “Basic IP Routing”](#), describes how to configure the G8272 for IP routing using IP subnets, BFD, and DHCP Relay.
- [Chapter 15, “Routed Ports”](#), describes how to configure a switch port to forward Layer 3 traffic.
- [Chapter 16, “Address Resolution Protocol”](#), describes how to use the Address Resolution Protocol (ARP) protocol to map an IPv4 address to a MAC address.
- [Chapter 17, “Internet Protocol Version 6”](#), describes how to configure the G8272 to use IPv6.
- [Chapter 18, “Internet Group Management Protocol”](#), describes how CNOS implements Internet Group Management Protocol (IGMP) Snooping to conserve bandwidth in a multicast-switching environment.
- [Chapter 19, “Border Gateway Protocol”](#), describes Border Gateway Protocol (BGP) concepts and features supported in CNOS.
- [Chapter 20, “Open Shortest Path First”](#), describes key Open Shortest Path First (OSPF) concepts, and how they are implemented in CNOS, and provides examples of how to configure your switch for OSPF support.
- [Chapter 21, “Route Maps”](#), describes route maps that are used to define route policy by permitting or denying certain routes based on a configured set of rules.

Part 5: High Availability Fundamentals

This part includes the following chapters:

- [Chapter 22, “Basic Redundancy”](#), describes how the G8272 supports redundancy through LAGs and VLAGs.
- [Chapter 23, “Virtual Router Redundancy Protocol”](#), describes how the G8272 supports high-availability network topologies using Virtual Router Redundancy Protocol (VRRP).

Part 6: Network Management

This part includes the following chapters:

- [Chapter 24, “Link Layer Discovery Protocol”](#), describes how Link Layer Discovery Protocol (LLDP) helps neighboring network devices learn about each others’ ports and capabilities.
- [Chapter 25, “Simple Network Management Protocol”](#), describes how to configure the switch for management through a Simple Network Management Protocol (SNMP) client.

Part 7: Monitoring

This part includes the following chapter:

- [Chapter 26, “Port Mirroring”](#), discusses tools to copy selected port traffic to a remote monitor port for network analysis.

Part 8: Appendices

This part includes the following appendices:

- [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 G8272 is available in the following guides:

- *Lenovo RackSwitch G8272 Installation Guide*
- *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*
- *Lenovo Network Release Notes for Lenovo Cloud Network Operating System 10.1*
- *Lenovo Network Python Programming Guide for Lenovo Cloud Network Operating System 10.1*
- *Lenovo Network REST API Programming Guide for Lenovo Cloud Network Operating System 10.1*

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 <code>readme.txt</code> file. G8272#
ABC123	This bold type appears in command examples. It shows text that must be typed in exactly as shown.	G8272# ping
<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: G8272# telnet <IP address> Read your <i>User's Guide</i> thoroughly.
{ }	Command items shown inside brackets are mandatory and cannot be excluded. Do not type the brackets.	G8272# cp {ftp sftp}
[]	Command items shown inside brackets are optional and can be used or excluded as the situation demands. Do not type the brackets.	G8272# configure [device]
	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.	G8272# cp {ftp sftp}
AaBbCc123	This block type depicts menus, buttons, and other controls that appear in graphical interfaces.	Click the Save button.

Part 1: Getting Started

This section discusses the following topics:

- [“Switch Administration” on page 27](#)
- [“System License Keys” on page 63](#)
- [“Switch Software Management” on page 69](#)

Chapter 1. Switch Administration

Your RackSwitch G8272 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 Cloud Network OS for the G8272 provides a variety of options for accessing the switch to perform a variety of configurations and to view switch information and statistics.

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

Administration Interfaces

Cloud NOS 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, while others require configuration; some can be accessed by local connection to the switch, while 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 (CLI) and menu system for switch access via a serial port connection or an optional Telnet or SSH session
- SNMP support for access through network management software such as Switch Center or Lenovo XClarity

The specific interface chosen for an administrative session depends on your preferences, 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 *Lenovo RackSwitch G8272 Installation Guide*).

Industry Standard Command Line Interface

The Industry Standard Command Line Interface (ISCLI) provides a simple and 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 G8272 (this option is always available)
- Telnet connection over the network
- SSH connection over the network

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 IP address can be provided automatically to the switch using a service such as DHCP (see “[DHCP IP Address Services](#)” on page 35). An IPv6 address can also 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 Interface

To manage the switch through the management interface, you must configure it with an IP interface. Configure the IP address and network mask and default gateway address:

1. Log onto the switch.
2. Enter Global Configuration mode.

```
G8272> enable
G8272# configure device
G8272(config)#
```

3. Configure a management IP address and network mask:

- IPv4 configuration:

```
G8272(config)# interface mgmt 0
G8272(config-if)# ip address <IPv4 address>/<IPv4 network mask length>
G8272(config-if)# no shutdown
G8272(config-if)# exit
```

- IPv6 configuration:

```
G8272(config)# interface mgmt 0
G8272(config-if)# ipv6 address <IPv6 address>/<IPv6 network mask length>
G8272(config-if)# no shutdown
G8272(config-if)# exit
```

4. Configure the appropriate default gateway:

- IPv4 configuration:

```
G8272(config)# vrf context management
G8272(config-vrf)# ip route 0.0.0.0 0.0.0.0 <default gateway IPv4 address>
G8272(config-vrf)# exit
```

- IPv6 configuration:

```
G8272(config)# vrf context management
G8272(config-vrf)# ipv6 route ::/0 <default gateway IPv6 address>
G8272(config-vrf)# exit
```

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

Using the Switch Ethernet Ports

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

1. Log onto the switch.
2. Enter interface mode and configure an ethernet interface as routed port.

```
G8272> enable
G8272# configure device
G8272(config)# interface ethernet <chassis number>/<port number>
G8272(config-if)# no bridge-port
```

3. Configure the interface IP address and network mask.

- IPv4 configuration:

```
G8272(config-if)# ip address <IPv4 address>/<IPv4 prefix length>
```

- IPv6 configuration:

```
G8272(config-if)# ipv6 address <IPv6 address>/<IPv6 prefix length>
```

4. Configure the default gateway.

- IPv4 configuration:

```
G8272(config)# vrf context management
G8272(config-vrf)# ip route 0.0.0.0 0.0.0.0 <default gateway IPv4 address>
G8272(config-vrf)# exit
```

- IPv6 configuration:

```
G8272(config)# vrf context management
G8272(config-vrf)# ipv6 route ::/0 <default gateway IPv6 address>
G8272(config-vrf)# exit
```

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

The G8272 supports an industry standard command-line interface (ISCLI) that you can use to configure and control the switch over the network using a Telnet or an SSH client. 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.

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 disabled. Use the following command to enable or disable Telnet access:

```
G8272(config)# [no] feature telnet
```

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 client on your workstation, use Telnet as the protocol type and the switch’s IP address as the hostname.

You will then be prompted to enter a password as explained in [“Switch Login Levels” on page 40](#).

By default, Telnet uses TCP port 23 of the remote host to establish a connection from the switch. When initializing a Telnet session, you can specify the TCP port of the remote host by using the following command on the switch:

```
G8272# telnet <switch IPv4 address> port <1-65535>
```

Note: The specified port will be used only for the current Telnet session. Future sessions will not use the selected port.

By default, Telnet clients will connect to the local Telnet server using TCP port 23 on the switch. To configure the TCP port used by a Telnet client when establishing a connection to the switch, use the following command:

```
G8272(config)# telnet server port <1-65535>
```

Using Secure Shell

Although a remote network administrator can manage the configuration of a G8272 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.

By default, SSH access is enabled. Use the following command to enable or disable SSH access:

```
G8272(config)# [no] feature ssh
```

The switch can do only one session of key/cipher generation at a time. Thus, an SSH client will not be able to log in if the switch is doing key generation at that time. Similarly, the system will fail to do the key generation if an SSH 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, aes192-ctr, aes256-ctr, arcfour128, arcfour256
- MAC: hmac-sha1, hmac-ripemd160, umac-64@openssh.com
- User Authentication: Local password authentication, public key authentication, TACACS+

Cloud Network 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

Once the IP parameters are configured, you can access the command line interface using an SSH connection.

To establish an SSH connection with the switch, run the SSH client on your workstation, use SSH as the protocol type and the switch's IP address as the hostname.

You will then be prompted to enter a password as explained in [“Switch Login Levels” on page 40](#).

Using SSH with Public Key Authentication

SSH can also be used for switch authentication based on asymmetric cryptography. Public encryption keys can be generated on the switch and used to authenticate incoming login attempts based on the client's 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. Disable SSH:

```
G8272(config)# no feature ssh
```

Note: SSH settings can't be modified if SSH is enabled.

2. Generate an SSH key:

- DSA:

```
G8272(config)# ssh key dsa
```

- RSA:

```
G8272(config)# ssh key rsa
```

Note: You can also configure the length of the RSA key by using the following command:

```
G8272(config)# ssh key rsa length <768-2048>
```

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

```
G8272(config)# ssh login-attempts <1-10>
```

Note: The default number of failed attempts is 3.

4. Re-enable SSH:

```
G8272(config)# feature ssh
```

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

Using Simple Network Management Protocol

CNOS provides Simple Network Management Protocol (SNMP) version 1, 2, and 3 support for access through any network management software, such as Switch Center or Lenovo XClarity.

Note: The SNMP read function is enabled by default. For best security practices, if SNMP is not needed for your network, disable this function prior to connecting the switch to the network.

To access the SNMP agent on the G8272, the read and write community strings on the SNMP manager must be configured to match those on the switch.

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

- read-only access community string:

```
G8272(config)# snmp-server community <community string (1-32 characters)> ro
```

- read-write access community string:

```
G8272(config)# snmp-server community <community string (1-32 characters)> rw
```

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 command:

```
G8272(config)# snmp-server host <IP address> traps version 1 <trap host community string>
```

For more information on SNMP usage and configuration, see [Chapter 25, “Simple Network Management Protocol.”](#)

DHCP IP Address Services

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

The G8272 can function as a relay agent for DHCP. This allows clients to be assigned an IP 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 IP address request to up to five DHCP servers. Additionally, up to five domain specific DHCP servers can be configured for each of up to 10 VLANs.

When a switch receives a DHCP request from a client seeking an IP address, the switch acts as a proxy for the client. The request is forwarded as a UDP unicast MAC layer message to the DHCP servers configured for the client's VLAN or to the global DHCP servers if no domain-specific DHCP servers are configured for the client's VLAN. The servers respond to the switch with a unicast reply that contains the IP default gateway and the IP 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 G8272 is described in RFC 1542. DHCP uses User Datagram Protocol (UDP) as its transport protocol. The client sends messages to the server on port 67 and receives messages from the server on port 68.

DHCP Client Configuration

DHCP is enabled by default on the management interface and disabled on all other interfaces. You can enable DHCP only on a maximum of 10 interfaces, including the management interface.

To enable or disable DHCP on an interface (for example ethernet interface 1/12), use the following command:

- for DHCPv4:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# no bridge-port
G8272(config-if)# ip address dhcp
```

- for DHCPv6:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# no bridge-port
G8272(config-if)# ipv6 address dhcp
```

Notes:

- DHCP cannot be enabled on an interface configured as a switch port, only on routing ports.
- Manually configuring an IP address on an interface will disable DHCP for that interface.

DHCPv4 Hostname Configuration (Option 12)

The G8272 supports DHCPv4 hostname configuration as described in RFC 2132, option 12. DHCPv4 hostname configuration is enabled by default.

The switch's hostname can be manually configured using the following command:

```
G8272(config)# hostname <system network name>
```

Note: If the hostname is manually configured, the switch does not replace it with the hostname received from the DHCPv4 server.

After DHCP configures the hostname on the switch, if the DHCPv4 configuration is disabled, the switch retains the hostname.

To enable or disable DHCP hostname configuration, use the following command on an interface (in this example, ethernet port 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# [no] ip dhcp client request host-name
```

To view the system hostname use the following command:

```
G8272> display hostname
```

Note: The switch prompt also displays the hostname.

DHCPv4 Syslog Server (Option 7)

The G8272 supports the requesting of the Syslog server IP address from the DHCP server as described in RFC 2132, option 7. The DHCPv4 Syslog server request option is disabled by default.

Note: Manually configured Syslog servers take priority over the DHCPv4 Syslog server.

Up to two Syslog server addresses received from the DHCPv4 server can be used. The Syslog server addresses can be learned over the management port or an ethernet port.

To enable or disable the DHCP Syslog server request, use the following command on an interface (in this example, ethernet port 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# [no] ip dhcp client request log-server
```

To view the Syslog server address, use the following command:

```
G8272> display logging server

IPv4 Servers:
    *2.2.2.1
severity: (debugging)
facility: local7
* - Values assigned by DHCP Client.
```

DHCPv4 NTP Server (Option 42)

This option request the DHCP server to provide a list of IP addresses indicating Network Time Protocol (NTP) servers available to the client. The NTP servers are listed in order of preference. The G8272 supports the requesting of NTP servers as described in RFC 2132, option 42.

By default, the switch does not include this request in DHCPv4 messages. To enable or disable this option on an interface, use the following command (in this example, ethernet port 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# [no] ip dhcp client request ntp-server
```

Note: Any manually configured NTP server will not be overwritten by the NTP servers received via DHCPv4.

To view the list of NTP servers, use the following command:

```
G8272> display ntp peers
```

DHCPv4 Vendor Class Identifier (Option 60)

This option is used by a DHCP client to identify itself to the DHCP server. It is used to define the vendor type and functionality of the DHCP client. The DHCP client can communicate to a server that it uses a specific type of hardware or software by specifying its Vendor Class Identifier (VCI).

The G8272 supports the identifying of a TFTP server as described in RFC 2132, option 60.

Each switch interface can be configured with a different VCI.

By default, the switch will include this option in DHCPv4 packets. To enable or disable the identification of TFTP servers use the following command (in this example, ethernet port 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# [no] ip dhcp client class-id
```

Note: For the G8272, the default VCI is 'LENOVO G8272'.

DHCP Relay Agent

When DHCP clients and associated servers are not on the same physical subnet, a DHCP relay agent can transfer DHCP messages between them. When a DHCP request arrives on an interface, the relay agent forwards the packet to all DHCP server IP addresses configured on that interface. The relay agent forwards replies from all DHCP servers to the host that sent the request. If no DHCP servers are configured on that interface, the relay agent will not forward packets.

DHCP has two versions. DHCPv4 is used to configure hosts with IPv4 addresses, IPv4 prefixes, and other configuration data required to operate in an IPv4 network. DHCPv6 is used to configure hosts with IPv6 addresses, IPv6 prefixes, and other configuration data required to operate in an IPv6 network.

For DHCPv4, you can configure the relay agent to add the relay agent information (option 82) in the DHCPv4 message and then forward it to the DHCPv4 server. The reply from the server is forwarded back to the client after removing option 82.

The DHCP Relay Agent is globally enabled by default. To globally enable or disable DHCP use the following command:

- for DHCPv4:

```
G8272(config)# [no] ip dhcp relay
```

- for DHCPv6:

```
G8272(config)# [no] ipv6 dhcp relay
```

DHCP relay can be configured differently on each ethernet port or VLAN. The maximum number of DHCP servers configured on an interface is 32. To configure DHCP on an interface, use the following steps:

1. Enter the configuration menu for the desired interface (in this example, ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)#
```

2. Configure the DHCP server address:

- for DHCPv4:

```
G8272(config-if)# ip dhcp relay address <IPv4 server address>
```

- for DHCPv6:

```
G8272(config-if)# ipv6 dhcp relay address <IPv6 server address>
```

3. To view the current DHCP settings, use the following command:

- for DHCPv4:

```
G8272> display ip dhcp relay
```

- for DHCPv6:

```
G8272> display ipv6 dhcp relay
```

DHCPv4 Option 82

DHCPv4 option 82 provides a mechanism for generating IP addresses based on the location in the network of the client device. When you enable the DHCPv4 relay agent option on the switch, it inserts the relay agent information option 82 in the packet. The switch then sends a unicast DHCPv4 request packet to the DHCPv4 server. The DHCPv4 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. The DHCPv4 relay agent strips off the option 82 field in the packet and sends the packet to the DHCPv4 client.

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

To configure DHCPv4 option 82, use the following command:

```
G8272(config)# ip dhcp relay information option
```

Switch Login Levels

To enable better switch management and user accountability, two levels or *classes* of user access have been implemented on the G8272. The levels of access to CLI management functions and screens increase as needed to perform various switch management tasks. Conceptually, access classes are defined as follows:

- *Network Operators* can only make temporary changes on the G8272. These changes will be lost when the switch is reloaded or reset. Operators have access to the switch management features used for daily switch operations. Because any changes an operator makes are undone by a reload of the switch, operators cannot severely impact switch operation.
- *Network Administrators* are the only ones that may make permanent changes to the switch configuration—changes that are persistent across a reload or reset of the switch. Administrators can access switch functions to configure and troubleshoot problems on the G8272. Because administrators can also make temporary (operator-level) changes as well, they must be aware of the interactions between temporary and permanent changes.

Note: The default (predefined) access classes cannot be removed or their rules modified. Also, new access classes cannot be created.

Access to switch functions is controlled through the use of unique usernames and passwords. Once you are connected to the switch via console, Telnet, or SSH, you are prompted to enter a password. The default username and password combinations for each access level are listed in [Table 2](#).

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
oper	oper	The Operator manages all functions of the switch. The Operator can reset ports, except the management port.	Disabled
admin	admin	The Administrator has complete access to all menus, information, and configuration commands on the G8272, including the ability to change both the operator and administrator passwords.	Enabled

For more details, see [“End User Access Control” on page 99](#).

To display the current role configurations, use the following command:

```
G8272> display role

Role : network-admin
  Description: Predefined network admin role has access to all commands
  on the switch
-----
Rule   Perm   Type      Scope      Entity
-----
1      permit read-write

Role : network-operator
  Description: Predefined network operator role has access to all read
  commands on the switch
-----
Rule   Perm   Type      Scope      Entity
-----
1      permit read
```

While a network administrator has access to all of the CLI commands, a network operator has a more limited access, only being able to run commands such as:

- **display**
- **end**
- **exit**
- **logout**
- **quit**
- **terminal**
- **enable**
- **disable**
- **ping**
- **ping6**
- **tracert**
- **tracert6**
- **ssh**
- **shh6**
- **telnet**
- **telnet6**
- **where**
- **configure device**

Ping

Ping (Poll INternet Gateway) is an administration utility used to test the connectivity between two network IP devices. It also measures the length of time it takes for a packet to be sent to a remote host plus the length of time it takes for an acknowledgement of that packet to be received by the source host.

Ping functions by sending an Internet Control Message Protocol (ICMP) echo request to the specified remote host and waiting for an ICMP reply from that host.

Using this method, ping also determines the time interval between when the echo request is sent and when the echo reply is received. This interval is called round-trip time. At the end of the test, ping will display the minimum, maximum, and average round-trip times, and the standard deviation of the mean.

Besides the round-trip time, ping can also measure the rate of packet loss. This is determined by the number of received echo replies over the number of sent echo requests. It is displayed as a percentage.

The G8272 also supports ping for IPv6 addressing.

To perform a standard ping test, use the following commands:

- IPv4:

```
G8272# ping <target IPv4 address>
```

- IPv6:

```
G8272# ping6 <target IPv6 address>
```

For example:

```
G8272# ping 10.10.10.1

PING 10.10.10.1 (10.10.10.1) from 10.10.10.127: 56(84) bytes of data.
64 bytes from 10.10.10.1: icmp_seq=1 ttl=61 time=0.368 ms
64 bytes from 10.10.10.1: icmp_seq=2 ttl=61 time=0.280 ms
64 bytes from 10.10.10.1: icmp_seq=3 ttl=61 time=0.308 ms
64 bytes from 10.10.10.1: icmp_seq=4 ttl=61 time=0.291 ms
64 bytes from 10.10.10.1: icmp_seq=5 ttl=61 time=0.320 ms

--- 10.10.10.1 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 3996ms
rtt min/avg/max/mdev = 0.280/0.313/0.368/0.034 ms
```

Ping Configurable Parameters

Ping can be configured with various parameters, such as specifying the number or size of echo requests, the time interval between each transmission, or the nonresponsive timeout interval for sent packets.

Test Interruption

Ping tests can be manually stopped at any point in the process. When the interruption is detected, ping will stop sending echo requests and display the results based on the packets transmitted up to that point.

To manually terminate a ping test, press **Ctrl + C**.

Ping Count

By default, ping transmits a sequence of five echo requests. To configure the number of packets sent during the test, use the following command:

```
G8272# ping <target IPv4 address> count <1-655350>
```

Ping can also be configured to continuously send echo requests until the test is manually interrupted. To achieve this, use the following command:

```
G8272# ping <target IPv4 address> count unlimited
```

For IPv6 addressing, the commands are as follows:

```
G8272# ping6 <target IPv6 address> count <1-655350>
```

```
G8272# ping6 <target IPv6 address> count unlimited
```

Ping Packet Interval

By default, ping does not wait between consecutive echo requests. As soon as a echo reply has been received or the nonresponsive timer has expired, ping will send the next echo request.

To configure a time interval, in seconds, between the transmission of packets, use the following command:

```
G8272# ping <target IPv4 address> interval <1-60>
```

For IPv6 addressing, the command is as follows:

```
G8272# ping6 <target IPv6 address> interval <1-60>
```

Ping Packet Size

By default, ping sends echo requests with a packet size of 56 bytes. Specifying a larger size than the default can help in detecting the loss of big packets.

To configure the packet size, in bytes, use the following command:

```
G8272# ping <target IPv4 address> packet-size <1-65468>
```

For IPv6 addressing, the command is as follows:

```
G8272# ping6 <target IPv6 address> packet-size <1-65468>
```

Ping Source

By default, ping automatically chooses the outgoing interface for echo requests and sends the packets using the IP address of that interface. To check the connectivity of different paths through the network, you can specify the interface used for sending echo requests.

To use a specific interface during the ping test, use the following command:

```
G8272# ping <target IPv4 address> source <source IPv4 address>
```

Note: The *source IPv4 address* is the IP address of the desired switch interface.

You can also choose the interface used for the ping test by directly specifying the desired interface. To achieve this, use the following command (in this example, ethernet port 1/12 is used):

```
G8272# ping <target IPv4 address> interface ethernet 1/12
```

For IPv6 addressing, the commands are as follows:

```
G8272# ping6 <target IPv6 address> source <source IPv6 address>
```

```
G8272# ping6 <target IPv6 address> interface ethernet 1/12
```

Ping DF-Bit

By default, echo requests are fragmented when they are forwarded through the network. Configuring packets not to be fragmented when traversing the network can help in determining the maximum transmission unit (MTU) of the path.

To enable the non-fragmentation of echo requests, use the following command:

```
G8272# ping <target IPv4 address> df-bit
```

Note: This parameter is configurable only for IPv4 addressing.

Ping Timeout

By default, after sending an echo request, ping waits up to a maximum of two seconds for an echo reply. If this time interval expires and an echo reply is not received, ping will declare that the remote host has timed out and that the sent packet is lost.

To configure the timeout interval, in seconds, use the following command:

```
G8272# ping <target IPv4 address> timeout <1-60>
```

For IPv6 addressing, the command is as follows:

```
G8272# ping6 <target IPv6 address> timeout <1-60>
```

Ping VRF

By default, ping uses the default Virtual Routing and Forwarding (VRF) instance. To configure ping to use a different VRF instance, use the following command:

```
G8272# ping <target IPv4 address> vrf {default|management}
```

Note: You can choose only between the default or management VRF instances.

For IPv6 addressing, the command is as follows:

```
G8272# ping6 <target IPv6 address> vrf {default|management}
```

Ping Interactive Mode

To configure a custom ping test, you can choose what parameters to change by combining the previously presented commands.

Besides this option, you can customize a ping test by using Ping Interactive Mode. In this mode, you can configure additional parameters: the type of service (ToS), the hop limit or time-to-live (TTL) and the data pattern.

Note: Ping Interactive Mode is only available for IPv4 addressing.

To enter Ping Interactive Mode, use the following command:

```
G8272# ping
```

You will be prompted to specify the value of each configurable parameter. If you do not enter a value, the default will be used.

```
G8272# ping
Vrf context to use [default]: management
Protocol [ip]:
Target IP address: 10.241.1.11
Repeat count [5]: 7
Datagram size [56]: 100
Timeout in seconds [2]: 1
Sending interval in seconds [1]:
Extended commands [n]: yes
Source address or interface:
Type of service [0]:
Set DF bit in IP header? [no]: yes
Data pattern [0xABCD]:
PATTERN: 0xabcd
PING 10.241.1.11 (10.241.1.11) 100(128) bytes of data.
108 bytes from 10.241.1.11: icmp_seq=1 ttl=61 time=0.337 ms
108 bytes from 10.241.1.11: icmp_seq=2 ttl=61 time=0.288 ms
108 bytes from 10.241.1.11: icmp_seq=3 ttl=61 time=0.311 ms
108 bytes from 10.241.1.11: icmp_seq=4 ttl=61 time=0.288 ms
108 bytes from 10.241.1.11: icmp_seq=5 ttl=61 time=0.317 ms
108 bytes from 10.241.1.11: icmp_seq=6 ttl=61 time=0.288 ms
108 bytes from 10.241.1.11: icmp_seq=7 ttl=61 time=0.315 ms

--- 10.241.1.11 ping statistics ---
7 packets transmitted, 7 received, 0% packet loss, time 5997ms
rtt min/avg/max/mdev = 0.288/0.306/0.337/0.022 ms
```

Traceroute

Traceroute is a diagnostic tool used to determine the network route between the switch and a remote device. It displays the network nodes (routers or gateway devices) crossed by a packet until it arrives at the specified destination.

Traceroute sends a sequence of User Datagram Protocol (UDP) packets addressed to a remote device. To determine the intermediate routers between the source and the destination devices, traceroute adjusts the time-to-live (TTL) value, also known as hop limit, of each sequence of sent packets. When a packet crosses a router, its hop limit is decreased by one. If a router detects a hop limit of zero, it discards the packet and sends the source host an Internet Control Message Protocol (ICMP) error message - Time Exceeded.

Traceroute configures the starting sequence of packets with a hop limit of one. The packets reach the first router and their hop limit is reduced from one to zero. The router will not forward the packets, but will instead discard them. Then, it sends an ICMP error message to the source host.

Traceroute sends the next set of packets with a hop limit of two. This time, the first router forwards the packets, reducing their TTL value from two to one. The packets reach the second router, which updates their hop limit to zero and discards them. Then, the second router will send the source host an ICMP error message.

Traceroute continues to send packets with increasing hop limit until the targeted remote device receives the packets and returns an ICMP echo reply.

After receiving the echo reply, traceroute uses the returned ICMP messages to create a list of the routers crossed by the packets. It uses the time interval between transmission and reception of packets as the delay (or latency) value for each node.

The G8272 also supports traceroute for IPv6 addressing.

To perform a traceroute test, use the following commands:

- IPv4:

```
G8272# traceroute <target IPv4 address>
```

- IPv6:

```
G8272# traceroute6 <target IPv6 address>
```

For example:

```
G8272# traceroute 10.241.1.11  
  
traceroute to 10.241.1.11 (10.241.1.11), 30 hops max, 56 byte packets  
 1  10.241.41.1 (10.241.41.1)  1.988 ms  2.117 ms  2.299 ms  
 2  10.241.4.254 (10.241.4.254)  1.903 ms  1.914 ms  2.649 ms  
 3  10.241.1.33 (10.241.1.33)  1.138 ms  1.195 ms  1.242 ms  
 4  10.241.1.11 (10.241.1.11)  1.085 ms !X  1.079 ms !X  1.087 ms !X
```

Traceroute Configurable Parameters

Traceroute is less customizable than ping, providing options only for choosing the outgoing interface or Virtual Routing and Forwarding (VRF) instance.

Test Interruption

Traceroute tests can be manually stopped at any point in the process. When the interruption is detected, traceroute will stop sending UDP packets and display the results based on the packets transmitted up to that point.

To manually terminate a Traceroute test, press **Ctrl + C**.

Traceroute Source

By default, traceroute automatically chooses the outgoing interface for sending UDP packets and transmits the packets using the IP address of that interface. To check the connectivity of different paths through the network, you can specify the interface used for sending packets.

To use a certain interface during a traceroute test, use the following command:

```
G8272# traceroute <target IPv4 address> source <source IPv4 address>
```

Note: The *source IPv4 address* is the IP address of the desired switch interface.

For IPv6 addressing, the command is as follows:

```
G8272# traceroute6 <target IPv6 address> source <source IPv6 address>
```

In the case of IPv6 addressing, you can also choose the interface used for the traceroute test by directly specifying the desired interface. To achieve this, use the following command (in this example, ethernet port 1/12 is used):

```
G8272# traceroute6 <target IPv6 address> interface ethernet 1/12
```

Traceroute VRF

By default, traceroute uses the default Virtual Routing and Forwarding (VRF) instance. To configure traceroute to use a different VRF instance, use the following command:

```
G8272# traceroute <target IPv4 address> vrf {default|management}
```

Note: You can choose only between the default or management VRF instances.

For IPv6 addressing, the command is as follows:

```
G8272# traceroute6 <target IPv6 address> vrf {default|management}
```


Traceroute Interactive Mode

To customize a traceroute test, you can also use Traceroute Interactive Mode. In this mode, you can configure additional parameters, such as the timeout interval, the hop limit or time-to-live (TTL), or the destination UDP port number.

Note: Traceroute Interactive Mode is only available for IPv4 addressing.

To enter Traceroute Interactive Mode, use the following command:

```
G8272# traceroute
```

You will be prompted to specify the value of each configurable parameter. If you do not enter a value, the default will be used.

```
G8272# traceroute

Vrf context to use [default]: management
Protocol [ip]:
Target IP address: 10.241.1.11
Source address: 10.241.41.27
Numeric display [n]: yes
Timeout in seconds [2]: 1
Probe count [3]: 5
Maximum time to live [30]: 20
Port Number [33434]: 15477
traceroute to 10.241.1.11 (10.241.1.11), 20 hops max, 56 byte packets
 1 10.241.41.1 0.665 ms 0.870 ms 1.102 ms 1.375 ms 1.611 ms
 2 10.241.4.254 1.459 ms 1.530 ms 1.568 ms 1.607 ms 1.646 ms
 3 10.241.1.33 0.790 ms 0.854 ms 1.654 ms 1.701 ms 1.755 ms
 4 10.241.1.11 0.287 ms !X 0.148 ms !X 0.096 ms !X 0.120 ms !X 0.130 ms !X
```

Network Time Protocol

The Network Time Protocol (NTP) is used to synchronize the internal clocks of network devices. This is helpful for troubleshooting network problems by correlating events on different network devices using, for example, syslog messages.

NTP provides the switch with a mechanism to accurately update its clock to be consistent with the clocks of other network devices within a precision of one millisecond.

NTP uses User Datagram Protocol (UDP) to communicate across the network.

By default, the NTP service is enabled on the switch. To enable or disable the NTP service, use the following command:

```
G8272(config)# [no] feature ntp
```

This can also be achieved by using the following command:

```
G8272(config)# [no] ntp enable
```

When NTP is enabled by using the **ntp enable** command, any configuration settings are saved in a file on the switch. If the protocol is disabled using the **no ntp enable** and then re-enabled, the switch will use that file to restore the NTP configuration. Any NTP commands issued while the protocol is disabled are ignored by the switch.

Note: If the NTP service is disabled by issuing the **no ntp feature** command, the NTP configuration is erased and it will not be restored after re-enabling the service.

By default, NTP will communicate across the interfaces configured in the management Virtual Routing and Forwarding (VRF) instance. To allow NTP to transmit and receive information using other VRF instances, use the following command:

```
G8272(config)# ntp use-vrf {default|management}
```

Note: Only the default or management VRF instances can be used.

NTP Synchronization Retry

An NTP synchronization can be manually triggered by using the following command:

```
G8272# ntp sync-retry
```

NTP Client and Peer

The switch can function as an NTP client and synchronize its internal clock with the clock of a remote NTP server. This synchronization is only one way, as the NTP server does not update its clock to match the one of the NTP client.

The G8272 supports both IPv4 or IPv6 address-based NTP.

To add or remove a remote NTP server, use the following command:

```
G8272(config)# [no] ntp server <IP address>
```

For the purpose of redundancy in scenarios where an NTP server is unreachable, NTP peers can be configured to communicate with each other. This mechanism is called symmetric-active mode and NTP peers associated in this manner can synchronize their internal clocks. The switch clock is updated using the clock of the remote NTP peer and vice-versa.

To add or remove a remote NTP peer, use the following command:

```
G8272(config)# [no] ntp peer <IP address>
```

Note: An NTP server and an NTP peer cannot share the same IP address. If you configure an NTP server using a specific IP address, you cannot use the same address to set up an NTP peer or vice-versa.

To view the list of currently configured NTP servers or peers, use the following command:

```
G8272> display ntp peers

-----
Peer IP Address                               Serv/Peer
-----
9.110.36.180                                  Server (configured)
129.36.17.25                                  Server (configured)
10.10.122.36                                  Peer (configured)
10.10.122.59                                  Peer (configured)
```

Note: The maximum number of NTP servers and peers statically or dynamically configured on the switch is 64.

To view the status of NTP servers and peers, use the following command:

```
G8272> display ntp peer-status

Total peers : 1
* - selected for sync, + - peer mode(active),
- - peer mode(passive), = - polled in client mode
remote          local          st    poll    reach    delay
-----
=9.110.36.180   9.111.86.200    16    64     0        0.00000
```

NTP Authentication Field Encryption Key

Information (packets) exchanged between the switch and NTP servers or peers can be configured to include encrypted authentication fields. These fields are encoded using a key.

By default, no encrypted fields are included in NTP packets.

To configure encryption authentication fields for NTP servers, use the following command:

```
G8272(config)# ntp server <IP address> key <1-65534>
```

To configure encryption authentication fields for NTP peers, use the following command:

```
G8272(config)# ntp peer <IP address> key <1-65534>
```

NTP Polling Intervals

The switch will periodically poll configured NTP servers or peers, and synchronize its internal clock.

Each individual NTP server or peer can be configured with different poll intervals by specifying a minimum and maximum time limit. The polling interval limits are configured in seconds, using powers of 2.

Each limit can be configured with a minimum value of 4, meaning 2^4 or 16 seconds, and a maximum value of 16, meaning 2^{16} or 65536 seconds (18.2 hours).

To configure the minimum polling interval for NTP servers, use the following command:

```
G8272(config)# ntp server <IP address> minpoll <4-16>
```

To configure the minimum polling interval for NTP peers, use the following command:

```
G8272(config)# ntp peer <IP address> minpoll <4-16>
```

By default, the minimum polling limit has a value of 4 (16 seconds).

To configure the maximum polling interval for NTP servers, use the following command:

```
G8272(config)# ntp server <IP address> maxpoll <4-16>
```

To configure the maximum polling interval for NTP peers, use the following command:

```
G8272(config)# ntp peer <IP address> maxpoll <4-16>
```

By default, the maximum polling limit has a value of 6 (64 seconds).

NTP Preference

During configuration, some NTP servers or peers can be marked as preferred over other devices. Preferred NTP servers or peers take precedence when a clock update is required.

To configure an NTP server as preferred, use the following command:

```
G8272(config)# ntp server <IP address> prefer
```

To configure an NTP peer as preferred, use the following command:

```
G8272(config)# ntp peer <IP address> prefer
```

Dynamic and Static NTP Servers

Using the Dynamic Host Configuration Protocol (DHCP) client, the switch can dynamically learn the addresses of NTP servers.

The DHCP client will add a dynamic server to the list of NTP servers and peers only if the new server isn't already statically configured on the switch, or the list hasn't reached its maximum size of 64 entries.

Static NTP servers take precedence over the ones dynamically learned through DHCP. Any dynamic NTP server entry is removed if a static entry for the same NTP server is configured on the switch.

If the list of NTP servers and peers is full, a statically configured NTP server will replace the oldest dynamic entry in the list. If there are no dynamic entries, the static server will not be added to the list.

The dynamic learning of NTP servers can only be enabled on individual switch interfaces. For more details, see [“DHCPv4 NTP Server \(Option 42\)” on page 37](#).

NTP Authentication

The switch can be configured to use authentication keys when exchanging packets with an NTP server. NTP authentication ensures that the internal clock is synchronized only using information received from a trusted NTP server.

NTP authentication uses a digital signature that is added to a packet, and it does not encrypt the data contained inside the packet. The packet and the configured authentication key are used to generate a non-reversible magic number which is appended to the packet. The receiving device does the same computation and then compares the results. If authentication succeeds (the two magic numbers match), the packet is used to synchronize the time. If authentication fails (the two magic numbers do not match), the packet is discarded.

By default, NTP authentication is disabled.

To configure an authentication key, use the following command:

```
G8272(config)# ntp authentication-key <1-65534 (key ID)> md5 <key string>
```

Note: You can configure up to 65534 NTP authentication keys. The maximum size of a *key string* is 8 characters. The *key string* is used together with the packet to generate the non-reversible magic number.

To remove an authentication key, use the following command:

```
G8272(config)# no ntp authentication-key <1-65534 (key ID)>
```

To view the configured NTP authentication keys, use the following command:

```
G8272> display ntp authentication-keys
```

Once an authentication key is created, the key can be configured as a trusted key and be used during the NTP authentication process. To enable or disable the use of a specific key in the authentication process, use the following command:

```
G8272(config)# [no] ntp trusted-key <1-65534 (key ID)>
```

To view the list of NTP trusted keys, use the following command:

```
G8272> display ntp trusted-keys
```

To enable or disable NTP authentication, use the following command:

```
G8272(config)# [no] ntp authenticate
```

To check the status of NTP authentication, use the following command:

```
G8272> display ntp authentication-status
```

NTP Authentication Configuration Example

To enable NTP authentication on the switch, use the following steps.

1. Configure an NTP authentication key:

```
G8272(config)# ntp authentication-key 10 md5 MyString
```

2. Configure the authentication key as a trusted key:

```
G8272(config)# ntp trusted-key 10
```

3. Enable NTP authentication:

```
G8272(config)# ntp authenticate
```

System Logging

System logging (syslog) is the mechanism through which messages generated by different software components, such as BGP, NTP, or Telnet, are reported by the switch applications. These messages are reported using several types of outputs including the console terminal, virtual teletype (VTY) sessions (Telnet or SSH), log files, or they are sent to remote syslog servers. Logging messages provide operational information about software components, including the status of the application, error reports, and detailed debugging data.

The G8272 uses the following format when outputting syslog messages:

```
<Timestamp> <IP/Hostname> %<Syslog ID>:[<Application Process Facility>] <Message>
```

Note: The angle brackets (< and >) are not present in the syslog message format, but they are used here to differentiate the individual components of the message.

The following parameters are used:

- **Timestamp**

The time of the message is displayed in the following format:

```
<year>-<month>-<day>-T<hour (1-24)>:<minute>:<second><timezone>
```

For example: 2015-08-01T18:39:59+00:00

- **IP/Hostname**

The IP address or hostname is displayed when configured.

For example: 1.1.1.1 or G8272

- **Syslog ID**

The syslog ID consists of three components that identify the syslog message:

```
<facility>-<severity level>-<mnemonic>
```

- **facility** - the application to which the message refers (such as BGP or NTP)
- **severity level** - the severity level of the message (a integer from 0 to 7)
- **mnemonic** - a text string that uniquely describes a syslog message

- **Application Process Facility**

The application process facility is only used when the syslog message is generated by a shared library facility. In this case, the facility parameter within the Syslog ID field identifies the shared library facility, whereas the application process facility identifies the client process, which is using the shared library.

- **Message**

The syslog message containing detailed event information.

Following are a couple of examples of syslog messages:

```
2014-08-15T04:50:33+00:00 switch %NSM-6-INIT: syslog service initialized
```

```
2015-03-23T21:19:17+00:00 G8272 %SECUREIMG-6-VALID_IMAGE_DETECTED: [IMISH]  
Valid image detected
```

Syslog Outputs

The syslog messages are stored in two types of log files: the customer log file and the platform log file.

The customer log file stores messages reported by the switch applications. This includes logs that record normal daily operations of the protocols, as well as error conditions detected. These events are primarily used for troubleshooting and auditing. The customer log file is persistent across switch reloads, and it can be modified or removed. It can be configured to log messages of a certain severity. Its maximum size can also be configured to a value between 4 kilobytes (KB) and 10 megabytes (MB). The log file consists of 8 individual parts, each with a maximum size of 1.25 MB. By default, the switch is configured with a log file that has a maximum size of 10 MB and it stores syslog messages with a severity level of 6 or higher.

The platform log file is used to store system messages generated by the switch operating system. These messages are mostly used for internal and debugging purposes, and are typically not of interest to the end users. These messages are stored in a set of files which are persistent across switch reloads. However, the log file is permanent and it can't be removed or modified. It logs any system messages regardless of their severity and it consists of 4 individual parts, each with a maximum size of 1 MB, adding to a maximum total of 4 MB.

To configure the storing of syslog messages in a customer log file on the switch, use the following command:

```
G8272(config)# logging logfile <log name> <0-7 (severity level)> size <4096-10485760 (bytes)>
```

To delete all customer log file entries, use the following command:

```
G8272# remove logging logfile
```

To disable the logging of messages to the log file, use the following command:

```
G8272(config)# no logging logfile
```

Syslog messages can also be enabled to be sent to devices connected via console or virtual teletype (VTY) sessions (Telnet, SSH), or to remote dedicated syslog servers.

By default, the logging of messages is enabled on console sessions for messages with a severity level of 2 or higher, and on VTY sessions for messages with a severity level of 5 or higher.

To enable or disable message logging on console sessions, use the following command:

```
G8272(config)# [no] logging console
```

To enable or disable message logging on VTY sessions, use the command:

```
G8272(config)# [no] logging monitor
```


Syslog Severity Levels

System Logging has eight severity levels that are attributed with a syslog message. They are commonly associated with an integer from 0 to 7, where 0 is the highest severity level and 7 the lowest.

Following are the different types of syslog severity levels:

- Emergency (0) - the system is unstable
- Alert (1) - a major error has occurred, that requires immediate action
- Critical (2) - an urgent error occurred, that requires immediate action
- Error (3) - a non-urgent error occurred
- Warning (4) - indication that an error might occur if not remedied
- Notice (5) - unusual events, but not error conditions
- Informational (6) - normal operation messages
- Debug (7) - information useful for debugging

When a severity level is configured, the switch will only log messages that have a severity equal or greater than the configured value. For example, when setting a severity level of 4, the switch will log messages that have a severity of 0, 1, 2, 3, or 4. Messages with a severity level of 5, 6, or 7 are not logged.

Each software component (such as OSPF or Telnet) or facility on the switch can be configured with a different severity level. While message logging for a component cannot be disabled, the severity level can be set to emergency (0), logging only the most severe messages for the specified component.

To configure the severity level of a facility, use the following command:

```
G8272(config)# logging level <facility> <0-7 (severity level)>
```

You can configure the severity level of all facilities by using the following command:

```
G8272(config)# logging level all <0-7 (severity level)>
```

Note: Each facility has its own default syslog severity level, which is usually different among facilities.

To view the severity level of a facility, use the following command:

```
G8272(config)# display logging level <facility>
```

To reset the severity level to its default value, use the following command:

```
G8272(config)# [no] logging level <facility>
```

For more details on syslog severity level commands, consult the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

Syslog Time Stamping

Each syslog message is generated with a timestamp that specifies the moment in time when the event of the message occurred.

The timestamp of a syslog message has the following syntax:

```
<year>-<month>-<day>-T<hour>:<minute>:<second><timezone>
```

Note: The angle brackets (< and >) are not present in the timestamp syntax, but they are used here to differentiate the individual components of the message.

For example: 2015-08-01T18:39:59+00:00

By default, syslog uses seconds as its base unit for timestamps. System logging can be configured to add fractions of a second (milliseconds or microseconds) to the timestamp of a message. To achieve this, use the following command:

```
G8272(config)# logging timestamp {microseconds|milliseconds|seconds}
```

To reset syslog to use seconds in timestamps, use the following command:

```
G8272(config)# no logging timestamp {microseconds|milliseconds|seconds}
```

Syslog Rate Limit

By default, the switch has a limit of how many messages it can log during a specific time interval. Any messages received above the maximum limit are not logged.

For system messages, the default rate limit is 512 messages every 5 seconds.

For every severity level, the default rate limit is 1024 messages every 10 seconds. Regardless of the software component that generated them, only 1024 messages with a certain severity can be logged during 10 seconds.

For every software component, the default rate limit is 512 messages every 10 seconds.

To configure the rate limit of a specific type of syslog message, use the following command:

```
G8272(config)# logging rate-limit num {<1-4096 (number of messages)>|default}  
interval {<1-600 (seconds)>|default} {facility <facility>|level <0-7 (severity level)>|  
system}
```

For more details on syslog rate limit commands, consult the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

Syslog Servers

You can configure the logging of messages to a remote dedicated syslog server. This can be statically configured, or dynamically learned via the Dynamic Host Configuration Protocol (DHCP).

By default, the logging of messages to a dedicated syslog server is disabled.

To add or remove a static remote syslog server, use the following command:

```
G8272(config)# [no] logging server <IP address> <0-7 (severity level)> [facility <facility>]
```

Static syslog servers take precedence over dynamically learned ones. If a syslog server is added using DHCP and then a static server is configured on the switch, the dynamic entry may be removed if one of the following conditions is met:

- the static server being added has the same IP address as the currently configured dynamic server
- the maximum number of servers allowed for a given type of IP addressing or address family (IPv4 or IPv6) has been reached

If a syslog server is already statically configured on the switch, it will not be dynamically added to the list.

The maximum number of configured syslog servers (static or dynamic) is three for each type of IP addressing, meaning that the switch can have up to three configured IPv4 syslog servers and up to three configured IPv6 syslog servers.

Note: Only DHCPv4 can be configured to request syslog server addresses from a DHCP server. Thus, only IPv4 syslog servers can be dynamically learned. IPv6 syslog servers can only be added statically.

The dynamic learning of syslog servers can be achieved only on individual switch interfaces. For more details, see [“DHCPv4 Syslog Server \(Option 7\)” on page 36](#).

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 35791 minutes, where 0 means the session will never timeout.

Use the following command to set the idle timeout value:

```
G8272# terminal session-timeout <0-35791>
```

Python Scripting

You can create and execute local Python scripts on switches to create small programs that allow a switch to automatically provision itself, perform fault monitoring, upgrade the image files, or auto-generate configuration files. You can use local scripts as a key part of your auto-provisioning solutions. You can also manage scripts on the Lenovo RackSwitch G8272.

You can implement version control systems, automatically generate alerts, create custom logging tools, and automate the management of network devices. Using Python scripts, you can perform many functions that can be performed through the CLI. In addition to configuration, you can notify users by sending e-mails or updating the Syslog.

See the *Lenovo Network Python Programming Guide for Lenovo Cloud Network Operating System 10.1* for details on how to create and execute Python scripts.

REST API Programming

The Lenovo REpresentational State Transfer (REST) Application Programming Interface (API) enables you to remotely configure and manage a Lenovo switch using REST and HyperText Transfer Protocol (HTTP).

The REST API is a JavaScript Object Notation-based (JSON) wrapper around Lenovo's Python Programming interface. It is a component of Configuration, Management, and Reporting (CMR) on CNOS.

See the *Lenovo REST API Programming Guide for Lenovo Cloud Network Operating System 10.1* for details on to use the Lenovo REST API.

Chapter 2. System License Keys

License keys determine the number of available features on the G8272. Each switch comes with a basic license that provides the use of a limited number of functions. On top of the basic license, optional upgrade licenses can be installed to expand the number of available features.

License keys will be installed only once per feature on a specific switch, regardless of the Lenovo Network OS used. When a new Lenovo NOS is installed on the switch, the license keys will need to be reinstalled.

The G8272 supports dual boot, meaning it can run both Lenovo Enterprise Network OS (ENOS) and Cloud Network OS (CNOS). When you install a license key using CNOS it will be also available in ENOS and vice-versa.

This section discusses the following topics:

- [“ONIE License Key” on page 68](#)
- [“Obtaining License Keys” on page 64](#)
- [“Installing License Keys” on page 65](#)
- [“Uninstalling License Keys” on page 66](#)
- [“Transferring License Keys” on page 67](#)

Obtaining License Keys

License keys or activation keys can be acquired using the *IBM System x Features on Demand* (FoD) website:

<http://www.ibm.com/systems/x/fod/>

You can also use the website to review and manage licenses, and to obtain additional help if required.

Note: An IBM ID and password are required to log into the FoD website. If you do not have an IBM ID, you can register at the website.

License keys are provided as files that must be uploaded to the G8272. To acquire an activation key, use the FoD website to purchase an Authorization Code. You will need to provide the unique ID (UID) of the specific G8272 where the key will be installed. The UID is the last 12 characters of the G8272 serial number. This serial number is located on the Part Number (PN) label and is also displayed during successful login to the device.

You can view the UID of a switch using the following command:

```
G8272> display license host-id  
System serial number: Y052MV4CR026
```

When available, download the activation key file from the FoD site.

Installing License Keys

Once an FoD license key file have been acquired, it must be installed on the G8272. The example below depicts the use of the G8272 Command Line Interface (CLI).

To install activation keys, use the following steps:

1. Log into the G8272.
2. Enter Global Configuration command mode:

```
G8272> enable
G8272# configure device
G8272(config)#
```

3. You can install the license key by copying it from a remote server or a USB device:
 - Remote FTP, SCP, SFTP, or TFTP server:

```
G8272(config)# license install {ftp|scp|sftp|tftp} <server address and file path>
vrf management
```

For example, using a TFTP server:

```
G8272(config)# license install tftp
tftp://10.120.33.12/fod-keys/fodEE3 vrf management
```

- USB device:

```
G8272(config)# license install usb1 <file path>
```

For example:

```
G8272(config)# license install usb1 fod-keys/fodEE3
```

Note: Repeat this step when installing multiple license keys.

The license key has now been successfully installed on the switch.

4. To view installed license keys, use the following command:

```
G8272> display license brief
```

Uninstalling License Keys

If you wish to remove or disable a feature that is available only through the use of a license key, you can achieve this by using the following command:

```
G8272(config)# no license install <license file>
```

For example:

```
G8272(config)# no license install fodEE3
```

Note: When removing the ONIE license key from the switch, the ONIE software image is also removed.

Transferring License Keys

License keys are based on the unique G8272 device serial number and are non-transferable.

In the event that the G8272 must be replaced, a new activation key must be acquired and installed. When the replacement is handled through IBM Service and Support, your original license will be transferred to the serial number of the replacement unit and you will be provided with a new license key.

ONIE License Key

The Open Network Install Environment (ONIE) is a small Linux-based operating system that provides an open install environment for networking devices without operating systems.

Currently, ONIE is not installed on the G8272 and it can be activated only using an FoD license key. The process of obtaining and installing an activation key is described in [Chapter 2, "System License Keys"](#).

After you have successfully installed the appropriate FoD license key, you will be able to install ONIE on the switch. For more details on ONIE and how to install and use it, see the *Lenovo ONIE Quick Start Guide*.

Chapter 3. Switch Software Management

The switch software image is the executable code running on the G8272. 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 G8272, go to the following website:

<https://www-945.ibm.com/support/fixcentral/>

To determine the software version currently used on the switch, use the following switch command in Privileged EXEC mode:

```
G8272# display boot

Current FLASH software:
  active image: version 10.1.1.0, downloaded 02:05:34 UTC Sun May 15 2016
  standby image: empty
  Uboot: version 10.1.1.0, downloaded 02:05:36 UTC Sun May 15 2016
  ONIE: empty
Currently set to boot software active image
Currently scheduled reboot time: none
Current port mode: default mode
```

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

1. Load a new software image and boot image onto an FTP, SFTP, SCP, or TFTP server on your network.
2. Transfer the new images to your switch.
3. Specify the new software image as the one which will be loaded into switch memory the next time a switch reload occurs.
4. Reload the switch.

For instructions on the typical upgrade process using the CNOS ISCLI, see [“Installing New Software to Your Switch” on page 70](#).



CAUTION:

Although the typical upgrade process is all that is necessary in most cases, upgrading from (or reverting to) some versions of Cloud Network 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.

Installing New Software to Your Switch

The G8272 can store up to two different CNOS images (called **active** and **standby**) and a special boot software named U-boot. U-Boot is an open source, primary boot loader used to package the instructions needed to boot the switch operating system kernel.

When you install a new CNOS image, it will be placed in the standby image on the switch. When you reload the switch using the standby image, it is swapped with the active image. After the switch reloads, the current active image will become the new standby image, and the current standby image will become the new active image.

Note: U-boot is included in the CNOS image.

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

- The new software image loaded on a FTP, SCP, SFTP or TFTP server on your network.
- The IP address of the FTP, SCP, SFTP, or TFTP server.
- The name of the new system image or ONIE image.

Note: Before installing an ONIE image on the switch, you need to install the appropriate license key. For more details, see [“System License Keys” on page 63](#).

When the software requirements are met, use the following procedures to download the new software to your switch using the ISCLI.

Installing System Images from a Remote Server

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

```
G8272# cp {ftp|scp|sftp|tftp} <server address and file path> system-image os vrf management
```

For example, using the TFTP protocol:

```
G8272# cp tftp tftp://10.120.33.12/nos-images/G8272-CNOS-10.1.1.0.imgs system-image os vrf management
```

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

Note: If you use the parameter **all** instead of **os** you will also install the U-boot file included in the CNOS image.

2. If required by the FTP, SCP, or SFTP server, enter the appropriate username and password.
3. You will be prompted to confirm the download of the CNOS image:

```
Confirm download operation? (y/n) [n]
```

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

```
Confirm download operation? (y/n) [n] y
Download in progress
.....
Copy Success

Install image...This takes about 90 seconds. Please wait
Valid image detected.
Check image signature succeeded
Extracting image: 100%
Installing system image to slot 2:
Installing image: 100%
Extracting image: 100%
OS image installation succeeded.

Standby image now contains Software Version 10.1.1.0
```

4. The new CNOS version is installed as the standby image. Once the installation is successful, you are asked if you want the switch to boot next time using the new image.

```
Switch is currently set to boot active image.
Do you want to change that to the standby image? (y/n) [n] y
Switch is to be booted with standby image.
```

5. Reboot the switch to run the new CNOS version:

```
G8272# reload
```

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

Note: Any unsaved configuration changes will be lost once the switch starts reloading.

Installing System Images from a USB Device

1. Use the following command to copy a CNOS image from a USB drive to the switch:

```
G8272# cp usb1 <file path> system-image os
```

In this example, a CNOS image is copied from a directory on a USB drive:

```
G8272# cp usb1 nos-images/G8272-CNOS-10.1.1.0.imgs system-image os
```

Note: If you use the parameter **all** instead of **os** you will also install the U-boot file included in the CNOS image.

2. You will be prompted to confirm the download of the CNOS image:

```
Confirm download operation? (y/n) [n]
```

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

```
Confirm download operation? (y/n) [n] y
Download in progress
.....
Copy Success

Install image...This takes about 90 seconds. Please wait
Valid image detected.
Check image signature succeeded
Extracting image: 100%
Installing system image to slot 2:
Installing image: 100%
Extracting image: 100%
OS image installation succeeded.
Standby image now contains Software Version 10.1.1.0
```

3. The new CNOS version is installed as the standby image. Once the installation is successful, you are asked if you want the switch to boot next time using the new image.

```
Switch is currently set to boot active image.
Do you want to change that to the standby image? (y/n) [n] y
Switch is to be booted with standby image.
```

4. Reboot the switch to run the new software:

```
G8272# reload
```

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

Note: Any unsaved configuration changes will be lost once the switch starts reloading.

Installing U-boot from a Remote Server

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

```
G8272# cp {ftp|scp|sftp|tftp} <server address and file path> system-image boot vrf management
```

For example, using the TFTP protocol:

```
G8272# cp tftp tftp://10.120.33.12/nos-images/G8272-CNOS-10.1.1.0.imgs system-image boot vrf management
```

Note: If you use the parameter **all** instead of **boot** you will also install the CNOS image.

The exact form of the server address and the file path of the U-boot file will vary by server. However, the file location is normally relative to the FTP, SCP, SFTP, or TFTP directory (for example, `tftpboot`).

2. If required by the FTP, SCP, or SFTP server, enter the appropriate username and password.
3. You will be prompted to confirm the download of the U-boot file:

```
Confirm download operation? (y/n) [n]
```

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

```
Confirm download operation? (y/n) [n] y
Download in progress
.....
Copy Success

Install image...This takes about 90 seconds. Please wait
Valid image detected.
Check image signature succeeded
Extracting image: 100%
Installing Boot Loader:
Updating flash: 100%
Boot image installation succeeded.
Boot loader now contains Software Version 10.1.1.0
```

On the next reload, the switch will use the new U-boot file.

Installing U-boot from a USB Device

1. Use the following command to copy a U-boot file from a USB drive to the switch:

```
G8272# cp usb1 <file path> system-image boot
```

In this example, a U-boot file is copied from a directory on the USB drive:

```
G8272# cp usb1 nos-images/G8272-CNOS-10.1.1.0.imgs system-image boot
```

Note: If you use the parameter **all** instead of **boot** you will also install the CNOS image.

2. You will be prompted to confirm the download of the software image:

```
Confirm download operation? (y/n) [n]
```

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

```
Confirm download operation? (y/n) [n] y
Download in progress
.....
Copy Success

Install image...This takes about 90 seconds. Please wait
Valid image detected.
Check image signature succeeded
Extracting image: 100%
Installing Boot Loader:
Updating flash: 100%
Boot image installation succeeded.
Boot loader now contains Software Version 10.1.1.0
```

On the next reload, the switch will use the new U-boot file.

Selecting a Software Image to Run

The switch can store up to two CNOS images. These images are referred to as the active and standby images. The active image is the image that the switch is currently using when rebooting. The standby image is the image that the switch keeps as backup, in case of a failure when booting the active image.

To see what CNOS versions are stored in each image and what image is configured to be used on the next reload, use the following command:

```
G8272# display boot

Current FLASH software:
  active image: version 10.1.1.0, downloaded 11:45:39 UTC Mon Feb 15 2016
  standby image: version 10.1.1.0, downloaded 14:44:26 UTC Tue Feb 9 2016
  Uboot: version 10.1.1.0, downloaded 11:17:04 UTC Wed Nov  4 2015
Currently set to boot software active image
Currently scheduled reboot time: none
Current port mode: default mode
```

You can choose which CNOS image the switch will use on the next reload by running the following command:

```
G8272(config)# startup image {active|standby}
```

For example:

```
G8272(config)# startup image standby

G8272# display boot

Current FLASH software:
  active image: version 10.1.1.0, downloaded 11:45:39 UTC Mon Feb 15 2016
  standby image: version 10.1.1.0, downloaded 14:44:26 UTC Tue Feb 9 2016
  Uboot: version 10.1.1.0, downloaded 11:17:04 UTC Wed Nov  4 2015
Currently set to boot software standby image
Currently scheduled reboot time: none
Current port mode: default mode
```

The switch will now reload using the standby image. After reloading, the switch interchanges the CNOS images. The standby image before the reload becomes the new active image and vice-versa.

Reloading the Switch

You can reboot the switch to make your CNOS image files changes occur. To reload the switch, use the following command:

```
G8272# reload
```

You are prompted for confirmation:

```
G8272# reload
reboot system? (y/n):
```

After confirming, the switch will halt and perform a reload. Any opened sessions will be closed. After the reload is complete, you'll have to log in the switch.

```
G8272# reload
reboot system? (y/n): y

Broadcast message from root@G8272 (Mon Feb 15 15:47:06 2016):

The system ireboot: Restarting system
...
G8272 login:
```

Note: Any unsaved configurations will be lost once the reload process starts.

Copying Configuration Files

You can also copy running and startup configuration files to or from the switch. This can be done via a remote server or a USB device.

Note: You cannot copy to the switch configuration files that exceed 4 MB in size or that have a line count larger than 100,000 lines.

Copy Configuration Files via a Remote Server

You can use the ISCLI to copy configurations files from or to a remote FTP, SCP, SFTP, or TFTP server.

Use the following command to copy the startup configuration file from a remote server:

```
G8272# cp {ftp|scp|sftp|tftp} <server address and file path> startup-config [vrf {default|management}]
```

For example, using the TFTP protocol:

```
G8272# cp tftp tftp://10.120.33.12/conf/mystartup.cfg startup-config
```

Use the following command to copy the startup or running configuration file to a remote server:

```
G8272# cp {running-config|startup-config} {ftp|scp|sftp|tftp} <server address and file path> [vrf {default|management}]
```

For example, using the TFTP protocol:

```
G8272# cp tftp tftp://10.120.33.12/conf/mystartup.cfg startup-config
```

Copy Configuration Files to a USB Device

You can insert a USB drive into the USB port on the G8272. You can copy files to the USB drive.

Use the following command to copy a configuration file to the USB drive:

```
G8272# cp {running-config|startup-config} usb1 <file path>
```

In this example, the running configuration file is copied to a directory on the USB drive:

```
G8272# cp running-config usb1 configuration/myconfig.cfg
```

The Boot Management Menu

The Boot Management menu allows you to switch the NOS image, reset the switch to its factory defaults, reset the default user password, or to recover from a failed software installation.

You can interrupt the startup process and enter the Boot Management menu from the serial console port. When the system displays the following message, press **Shift + B**. The Boot Management menu will appear.

```
The system is going down for reboot NOW!
INIT: reboot: Restarting system

...

Press shift-B for startup menu or shift-R for recovery mode: ..
Running Startup Menu

...

Boot Management Menu
  M - Change boot mode (Legacy vs NextGen)
  I - Change booting image
  R - Boot in recovery mode (tftp and xmodem download of images to
recover switch)
  P - Password reset
  O - ONIE submenu
  Q - Reboot
  E - Exit
Please choose your menu option:
```

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

- To change the startup mode, press **M** and follow the screen prompts. Legacy refers to Lenovo Enterprise Network OS (8.2 or prior) and NextGen refers to Lenovo Cloud Network OS (10.1 or later).
- To change the booting image, press **I** and follow the screen prompts.
- To boot in recovery mode, press **R**.
- To reset the default user password, press **P** and follow the screen prompts.
- To enter the ONIE submenu, press **O**.
- To reload the switch, press **Q**. The reloading process will start again.
- To exit the Boot Management menu, press **E**. The reloading process continues.

Note: The ONIE submenu will only be available if you have activated ONIE by installing an appropriate license key. For more details, see [Chapter 2, "System License Keys"](#).

Boot Recovery Mode

The Boot Recovery Mode allows you to recover from a failed software or boot image upgrade using TFTP or XModem download.

To enter Boot Recovery Mode you must select the **Startup in recovery mode** option from the Boot Management Menu.

```
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
    P) Physical presence (low security mode)
    F) Filesystem check
    R) Reboot
    E) Exit

Option? :
```

The Boot Recovery Mode menu allows you to perform the following actions:

- To recover from a failed software or boot image upgrade using TFTP, press **T** and follow the screen prompts. For more details, see [“Recover from a Failed Image Upgrade using TFTP”](#) on page 80.
- To recover from a failed software or boot image upgrade using XModem download, press **X** and follow the screen prompts. For more details, see [“Recovering from a Failed Image Upgrade using XModem Download”](#) on page 82.
- To enable the loading of a special image, press **P** and follow the screen prompts. For more details, see [“Physical Presence”](#) on page 84.
- To check if the switch is ready to run Cloud NOS software, press **F**. It performs a check to see if the filesystem is optimally partitioned and updates it accordingly.
- To restart the reload process from the beginning, press **R**.
- To exit Boot Recovery Mode menu, press **E**.

Recover from a Failed Image Upgrade using TFTP

Use the following procedure to recover from a failed image upgrade using TFTP:

1. Connect a PC to the console port of the switch.
2. Open a terminal emulator program that supports the Telnet protocol (for example, HyperTerminal, SecureCRT, or PuTTY) and input the proper hostname or IP address and the port to connect to the console interface of the switch.
3. Reload the switch and access the Boot Management menu by pressing **Shift + B** when the following message appears and the dots are being displayed.

```
Press shift-B for startup menu or shift-R for recovery mode: ...
```

4. Enter Boot Recovery Mode by selecting **R**. The Recovery Mode menu will appear.
5. To start the recovery process using TFTP, select **T**. The following message will appear:

```
Performing TFTP rescue. Please answer the following questions (enter 'q' to quit):
```

6. Enter the IP address of the management port:

```
IP addr :
```

7. Enter the network mask of the management port:

```
Netmask :
```

8. Enter the gateway of the management port:

```
Gateway :
```

9. Enter the IP address of the TFTP server:

```
Server addr :
```

10. Enter the file path and the filename of the image:

```
Image Filename:
```

After the procedure is complete, the Recovery Mode menu will be redisplayed.

Following is an example of a successful recovery procedure using TFTP:

```
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
    P) Physical presence (low security mode)
    F) Filesystem check
    R) Reboot
    E) Exit

Option? : t
Performing TFTP rescue. Please answer the following questions (enter 'q'
to quit):
IP addr :10.241.6.4
Netmask :255.255.255.128
Gateway :10.241.6.66
Server addr:10.72.97.135
Image Filename: G8272-CNOS-10.1.1.0.imgs
    Netmask : 255.255.255.128
    Gateway : 10.241.6.66
Configuring management port.....
Installing image G8272-CNOS-10.1.1.0.imgs from TFTP server 10.72.97.135

Extracting images ... Do *NOT* power cycle the switch.
Installing Application: Image signature verified.
Installing image as standby image: 100%

Standby image update succeeded
Updating install log. File G8272-CNOS-10.1.1.0.imgs installed from
10.72.97.135 at 15:29:30 on 12-3-2015
Please select one of the following options:
    T) Configure networking and tftp download an image
    X) Use xmodem 1K to serial download an image
    P) Physical presence (low security mode)
    F) Filesystem check
    R) Reboot
    E) Exit

Option? :
```

Recovering from a Failed Image Upgrade using XModem Download

Use the following procedure to recover from a failed 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, SecureCRT or PuTTY) and select the following serial port characteristics:
 - Speed: 9600 bps
 - Data Bits: 8
 - Stop Bits: 1
 - Parity: None
 - Flow Control: None
3. Reload the switch and access the Boot Management menu by pressing **Shift + B** when the following message appears and the dots are being displayed.

```
Press shift-B for startup menu or shift-R for recovery mode: ...
```

4. Enter Boot Recovery Mode by selecting **R**. The Recovery Mode menu will appear.
5. Select **X** for Xmodem download. You will see the following display:

```
Running xmodem rescue.....
```

6. When you see the following message, change the Serial Port speed to 115200 bps:

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

7. 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:

```
... Waiting for the <Enter> key to be hit before the download can start...  
CC
```

8. Select the image to download. Xmodem initiates the file transfer. When download is complete, you are asked to change the Serial Port speed back to 9600 bps:

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

9. Press **Enter** to start installing the image.

The image install will begin. After the procedure is complete, the Recovery Mode menu will be redisplayed.

```
Extracting images ... Do *NOT* power cycle the switch.
Installing Root Filesystem:
Image signature verified. 100%
Installing Kernel:
Image signature verified. 100%
Installing Device Tree:
Image signature verified. 100%
Installing Boot Loader: 100%
Updating install log. File image installed from xmodem at 18:06:02 on
13-3-2015
Please select one of the following options:
    T) Configure networking and tftp download an image
    X) Use xmodem 1K to serial download an image
    P) Physical presence (low security mode)
    F) Filesystem check
    R) Reboot
    E) Exit

Option? :
```

Boot image recovery is complete.

Physical Presence

Use the following procedure to enable the installation of special images on the switch, when a deeper troubleshooting analysis is required:

1. Connect a PC to the console port of the switch.
2. Open a terminal emulator program that supports a serial port connection 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** when the following message appears and the dots are being displayed.

```
Press shift-B for startup menu or shift-R for recovery mode: ...
```

4. Enter Boot Recovery Mode by selecting **R**. The Recovery Mode menu will appear.
5. To begin the Physical Presence procedure, select **P**. The following warning message will appear and you will be prompted for confirmation (press **y** to confirm):

```
WARNING: the following test is used to determine physical presence and if
completed will put the switch in low security mode.

Do you wish to continue y/n? y
```

6. A security test will be performed. The system location (blue) LED will blink a number of times between 1 and 12. Enter that number:

```
Hit a key to start the test. The blue location LED will blink a number of
times.

.....
How many times did the LED blink?
```

7. After entering the correct number, the Recovery Mode menu will reappear. To install a special image, use one of the following procedures:
 - TFTP (for details, see [page 80](#))
 - XModem Download (for details, see [page 82](#))

Note: You have three attempts to successfully complete the security test. After three incorrect attempts, the switch will reboot.

After the test is completed, the switch will be put in low security mode. This mode will allow you to install special images on the switch. To revert to normal security mode, you must reboot the switch or press **P** again in the Recovery Mode menu.

ONIE submenu

In the ONIE submenu you can choose from different ONIE operations that include installing and uninstalling NOS images, upgrading ONIE, and using ONIE rescue and recovery mode to troubleshoot a broken system.

For more details about ONIE, see [“ONIE” on page 86](#).

Note: The ONIE submenu will only be available if you have activated ONIE by installing an appropriate license key. For more details, see [Chapter 2, “System License Keys”](#).

To enter the ONIE submenu you must select the “ONIE submenu” option from the Boot Management Menu.

```
ONIE Menu
  I - Boot ONIE OS installer
  N - Boot NOS mode (system default)
  R - Boot ONIE rescue mode
  U - Boot ONIE self update mode
  D - Boot ONIE OS uninstaller
  E - Exit ONIE menu

Please choose your menu option:
```

The ONIE submenu allows you to perform the following actions:

- To boot the switch in ONIE install mode, press **I**. For more details, see [“Booting in ONIE Install Mode” on page 88](#).
- To boot the switch using the installed NOS image, press **N**.
- To boot the switch in ONIE rescue mode, press **R**. For more details, see [“Booting in ONIE Rescue Mode” on page 89](#).
- To boot the switch in ONIE update mode, press **U**. For more details, see [“Booting in ONIE Update Mode” on page 89](#).
- To boot the switch in ONIE uninstall mode, press **D**. For more details, see [“Booting in ONIE Uninstall Mode” on page 89](#).
- To exit the ONIE submenu and return to the Boot Management Menu, press **E**.

ONIE

The Open Network Install Environment (ONIE) is a small Linux-based operating system that provides an open install environment for networking devices without operating systems.

ONIE enables a network switch ecosystem for end-users to choose among different Network Operating Systems (NOS). Practically, ONIE boots on a switch, discovers NOS installer images available on the local network or USB drive, copies the chosen image to the switch, and provides an environment where the installer can load the NOS onto the switch.

The G8272 is compatible with ONIE. To use ONIE on the switch, you must first install a license key. For more details on how to obtain and install such a license, see [Chapter 2, “System License Keys”](#).

For more details on ONIE and how to install and use it, see the *Lenovo ONIE Quick Start Guide*.

Installing ONIE from a Remote Server

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

```
G8272# cp {ftp|scp|sftp|tftp} <server address and file path> system-image onie vrf management
```

For example, using the TFTP protocol:

```
G8272# cp tftp tftp://10.120.33.12/onie/G8272-ONIE-CNOS-10.1.1.0.imgs system-image onie vrf management
```

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

2. If required by the FTP, SCP, or SFTP server, enter the appropriate username and password.
3. You will be prompted to confirm the download of the ONIE image:

```
Confirm download operation? (y/n) [n]
```

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

```
Confirm download operation? (y/n) [n] y
Download in progress
.....
Copy Success

Install image...This takes about 90 seconds. Please wait
Valid image detected.
Check image signature succeeded
Installing image: 100%
Extracting image: 100%
ONIE image installation succeeded.
ONIE image now contains Software Version 10.1.1.0
```

You have successfully installed the ONIE image.

Installing ONIE from a USB Device

1. Use the following command to copy an ONIE image from a USB drive to the switch:

```
G8272# cp usb1 <file path> system-image onie
```

In this example, an ONIE image is copied from a directory on the USB drive:

```
G8272# cp usb1 onie-image/G8272-ONIE-CNOS-10.1.1.0.imgs system-image onie
```

2. You will be prompted to confirm the download of the ONIE image:

```
Confirm download operation? (y/n) [n]
```

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

```
Confirm download operation? (y/n) [n] y
Download in progress
.....
Copy Success

Install image...This takes about 90 seconds. Please wait
Valid image detected.
Check image signature succeeded
Installing image: 100%
Extracting image: 100%
ONIE image installation succeeded.
ONIE image now contains Software Version 10.1.1.0
```

You have successfully installed the ONIE image.

Booting in ONIE Mode

After you have activated and installed an ONIE image, you can boot the switch in ONIE mode. ONIE offers four modes for booting:

- ONIE install mode
- ONIE uninstall mode
- ONIE update mode
- ONIE rescue mode

To view the currently installed ONIE image, use the following command:

```
G8272# display boot

Current FLASH software:
  active image: version 10.1.1.0, downloaded 13:45:51 UTC Wed Feb 17 2016
  standby image: empty
  Uboot: version 10.1.1.0, downloaded 13:45:51 UTC Wed Feb 17 2016
  ONIE: version 10.1.1.0, downloaded 14:16:53 UTC Wed Feb 17 2016
  Currently set to boot software active image
  Currently scheduled reboot time: none
  Current port mode: default mode
```

For more information, see the *Lenovo ONIE Quick Start Guide*.

Booting in ONIE Install Mode

ONIE install mode boots the ONIE kernel and file-system, and executes the ONIE discovery process. The discovery process uses various mechanisms to find a NOS installer program.

Once a NOS installer is found, it is downloaded and executed. The installer loads the Network OS on the switch, updates required ONIE environment variables, and reloads the switch.

The switch will now run its newly installed Network OS. You can, at any point, reload the switch in ONIE installer mode to install another Network OS.

To boot the switch in ONIE install mode on the next reload, use the following command:

```
G8272(config)# startup image onie-image install
```

To check the next boot image, use the following command:

```
G8272# display boot

Current FLASH software:
  active image: version 10.1.1.0, downloaded 13:45:51 UTC Wed Feb 17 2016
  standby image: empty
  Uboot: version 10.1.1.0, downloaded 13:45:51 UTC Wed Feb 17 2016
  ONIE: version 10.1.1.0, downloaded 14:16:53 UTC Wed Feb 17 2016
  Currently set to boot ONIE install mode
  Currently scheduled reboot time: none
  Current port mode: default mode
```


Booting in ONIE Uninstall Mode

ONIE uninstall mode will reload the switch and delete the NOS image running on the switch. After the uninstall is complete, the ONIE discovery process is executed. For mode information, see [“Booting in ONIE Install Mode” on page 88](#).

To boot the switch in ONIE uninstall mode on the next reload, use the following command:

```
G8272(config)# startup image onie-image uninstall
```

Note: The following warning will be displayed after choosing ONIE uninstall mode. You will be prompted for confirmation (press **y** to confirm):

```
WARNING: Booting ONIE in uninstall mode will completely uninstall the
operating system!
Only the ONIE image will be preserved.
Are you sure you want the next boot to uninstall the OS? (y/n) y
```

Once you confirm, the switch will boot in ONIE uninstall mode on the next reload.

Booting in ONIE Update Mode

ONIE update mode will install a new version of ONIE on the switch, including the boot loader and the Linux kernel. It works similarly to the ONIE install mode, except the discovery mode will look for an ONIE image and not a NOS image.

Note: This will only upgrade the ONIE partition on the switch and will not affect the NOS image.

To boot the switch in ONIE update mode on the next reload, use the following command:

```
G8272(config)# startup image onie-image update
```

Note: The Lenovo U-boot file will be overwritten by the ONIE update process. Although the Lenovo NOS image is not deleted, the switch will be unable to boot using the CNOS image.

Booting in ONIE Rescue Mode

ONIE rescue and recovery mode is useful when trying to recover from a broken NOS image. It works similarly to the ONIE install mode, except the ONIE discovery process is disabled. ONIE will not try to find and install a NOS image, but instead it will allow you to troubleshoot the system.

To boot the switch in ONIE rescue mode on the next reload, use the following command:

```
G8272(config)# startup image onie-image rescue
```


Part 2: Securing the Switch

This section discusses the following switch security topics:

- [“Securing Administration” on page 93](#)
- [“Authentication & Authorization Protocols” on page 105](#)
- [“Access Control Lists” on page 117](#)

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 94](#)
- [“End User Access Control” on page 99](#)

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 25, “Simple Network Management Protocol”](#)).

Secure Shell and Secure Copy

Because using Telnet does not provide a secure connection for managing a G8272, Secure Shell (SSH) and Secure Copy (SCP) features have been included for G8272 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 G8272 over a network to execute management commands. The SSH module consists of a SSH server and a SSH client. By default, the SSH server listens on TCP port number 22. Once a SSH client establishes a connection to the SSH server, the client acts as a virtual terminal through which the remote system is accessed.

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 G8272, SCP is used to download and upload the switch configuration via secure channels.

SSH is enabled by default and using this features provides the following benefits:

- Identifying the administrator using Name/Password
- Authentication of remote administrators
- Encrypting messages between the remote administrator and switch
- Secure copy support

Cloud Network 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

SSH Encryption and Authentication

The following encryption and authentication methods are supported for SSH:

- Server Host Authentication: Client RSA authenticates the switch at the beginning of every connection
- Key Exchange: RSA, DSA
- Encryption: 3DES-CBC, DES
- User Authentication: Local password authentication, TACACS+

Generating RSA/DSA Host Key for SSH Access

To support the SSH host feature, a RSA or DSA host key is required. The host key is 2048 bits and is used to identify the G8272.

To configure a new RSA or DSA host key, first connect to the G8272 through the console port or via a Telnet session and enter the following command to generate it manually:

- To generate a new DSA key:

```
G8272(config)# ssh key dsa
```

- To generate a new RSA key:

```
G8272(config)# ssh key rsa
```

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, a 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 a SSH/SCP client is logging in at that time.

SSH Integration with TACACS+ Authentication

SSH 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.

Configuring SSH on the Switch

To configure SSH parameters, use the following procedure:

1. Configure the number of unsuccessful SSH login attempts (the default value is 3):

```
G8272(config)# ssh login-attempts <1-10>
```

2. Optionally, you can change the SSH server port (the default port is 22):

```
G8272(config)# ssh server port <1-65535>
```

Note: To change the number of SSH login attempts and server port, the SSH feature must be disabled.

3. Optionally, you can configure a DSA or RSA host key:

```
G8272(config)# ssh key {dsa|rsa}
```

4. Check the current SSH server settings:

```
G8272# display ssh server
```


Using SSH Client Commands

This section shows the format for using some client commands.

To Log In to the Switch

1. To log in to a SSH server you can connect via IPv4 or IPv6:
 - a. Connecting to the SSH server using IPv4:

```
G8272# ssh <username>@<server IPv4 address>
```

Example:

```
G8272# ssh admin@205.178.15.157
```

- b. Connecting to the SSH server using IPv6:

```
G8272# ssh6 <server IPv6 address>
```

Example:

```
G8272# ssh6 FE80::0202:B3FF:FE1E:8329
```

2. Optionally, you can specify the SSH server port used to establish the connection:

```
G8272# ssh <username>@<server IPv4 address> port <1-65535>
```

3. Optionally, you can also specify the Virtual Routing and Forwarding (VRF) instance used for the session:

```
G8272# ssh <username>@<server IPv4 address> port <1-65535> vrf {default|management}
```

Using Secure Copy

You can use SCP to copy files from or to a remote server. This includes the copying of the running configuration, startup configuration or technical support files.

Copying a File Using SCP

To copy a file to a remote server using SCP use the following command:

```
G8272# cp file <filename> scp
```

To copy a file from a remote server using SCP use the following command:

```
G8272# cp scp file <filename>
```

Copying Startup Configuration Using SCP

To copy the startup configuration file to a remote server using SCP use the following command:

```
G8272# cp startup-config scp
```

To copy the startup configuration file from a remote server using SCP use the following command:

```
G8272# cp scp startup-config
```

Copying Running Configuration Using SCP

To copy the running configuration file to a remote server using SCP use the following command:

```
G8272# cp running-config scp
```

Copying Technical Support Using SCP

To copy the technical support dump file to a remote server using SCP use the following command:

```
G8272# cp tech-support scp
```

End User Access Control

Cloud NOS 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 100 user IDs are supported on the switch.
- CNOS offers end user support for console, Telnet and SSH access to the switch.
- The length of the username must be between 2 and 28 alphanumeric characters.
- Passwords for end users must be between 8 and 80 alphanumeric characters in length for TACACS+, Telnet, SSH and console access.

Strong Passwords

Cloud NOS requires the use of Strong Passwords for users to access the G8272. Strong Passwords enhance security because they make password guessing more difficult.

The following rules must be followed when creating a password for a user:

- Minimum length: 8 characters; maximum length: 80 characters;
- Must contain at least one uppercase alphabet;
- Must contain at least one lowercase alphabet;
- Must contain at least one number;
- Cannot be same as the username;
- When changing passwords no four consecutive characters can be the same as in the old password;

A user without a password will not be allowed to log in on the switch.

User Access Control

The end-user access control commands allow you to configure end-user accounts. Any changes to a user will only take place once that user logs out and creates a new session.

To displays the currently logged in user, use the following command:

```
G8272# display users
```

Setting up Users

Up to 100 users can be configured on the switch. Use the following commands to define any username and set the user password:

1. Create the user:

```
G8272(config)# username <username>
```

Note: The username must be between 2 and 28 alphanumeric characters long.

2. Define a password for the user:

```
G8272(config)# username <username> password <password>
```

Note: The password must be between 8 and 80 alphanumeric characters long and must contain at least one uppercase alphabet and one number.

3. Check the current user accounts:

```
G8272# display user-accounts
```

The following words are reserved and cannot be used as usernames:

Table 3. Words not permitted as usernames

adm	bin	daemon	ftp	ftpuser
games	gdm	gopher	halt	lp
mail	mailnull	man	mtsuser	news
shutdown	sync	sys	uucp	xf
root	vsca	lighttpd	tcpdump	oprofile
ntp	dhcpcd	sshd	nosx	obs
proxy	backup	list	irc	gnats
tss	cyrus	messagebus		

Defining a User's Access Level

Every user account will have a certain access level or role attached. The user role defines what operations the user will be allowed to perform on the switch. The user can have one of two roles:

- **network-admin**: complete read and write access to the switch;
- **network-operator**: only read access to the switch.

To view the current settings for each role, use the following command:

```
G8272# display role
```

To change a user's role, use the following command:

```
G8272(config)# username <username> role {network-admin|network-operator}
```

Note: If a user is created without specifying an access level, the default user role will be network-operator.

To view the current user roles for each user, use the following command:

```
G8272# display user-accounts
```

Deleting a User

To delete a user, use the following steps:

1. Delete the user:

```
G8272(config)# no username <username>
```

2. Verify if the user has been removed:

```
G8272# display user-account
```

The Default User

The user 'admin' will automatically be installed on the switch by default. This is used to initially set up the switch.

Its default role is network-admin and the default password is 'admin'.

The 'admin' user account cannot be removed nor its user role changed. However, the password associated with this account can be modified by using the following command:

```
G8272(config)# username admin password <new password>
```

Administrator Password Recovery

In case you have changed the password of the 'admin' user and have forgotten it, you can follow the steps below to reset the password to the default value:

1. Connect to the switch using the console port.
2. Reload the switch:

```
G8272# reload
```

3. During the reload process, the message below will appear. Press **Shift + B** to enter Boot Management.

```
G8272# reload
reboot system? (y/n): y

Broadcast message from root@G8272 (Mon Feb 15 15:53:53 2016):

The system is going down for reboot NOW!
INIT: reboot: Restarting system

U-Boot 2014.01 (Oct 07 2015 - 10:47:24) - ONIE lenovo

CPU0: P2020, Version: 2.1, (0x80e20021)
Core: e500, Version: 5.1, (0x80211051)
Clock Configuration:
  CPU0:1200 MHz, CPU1:1200 MHz,
  CCB:600 MHz,
  DDR:400 MHz (800 MT/s data rate) (Asynchronous), LBC:37.500 MHz
L1: D-cache 32 KiB enabled
  I-cache 32 KiB enabled
Board: G8272
I2C: ready
DRAM: 4 GiB (DDR3, 64-bit, CL=6, ECC on)
Booting Lenovo NOS....
Press shift-B for startup menu or shift-R for recovery mode: ...
```

4. Select 'Password reset' from the Boot Management menu by pressing **P**:

```
Boot Management Menu
  M - Change startup mode (E-NOS vs C-NOS)
  I - Change booting image
  R - Startup in recovery mode (tftp and xmodem download of images
to recover switch)
  P - Password reset
  Q - Reboot
  E - Exit
Please choose your menu option: P
```

5. You will prompt for confirmation:

```
Confirm resetting system password: y/N? y
Password reset requested. Please reboot to complete the process
```

6. Reload the switch by selecting the 'Reboot' option. Press **Q**:

```
Boot Management Menu
  M - Change startup mode (E-NOS vs C-NOS)
  I - Change booting image
  R - Startup in recovery mode (tftp and xmodem download of images
to recover switch)
  P - Password reset
  Q - Reboot
  E - Exit
Please choose your menu option: Q
```

7. After the reload is complete, you will be able to log in the switch by using the default user 'admin' with the default password 'admin'.

8. You can change the default password by using the following command:

```
G8272(config)# username admin password <new password>
```

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:

- [“TACACS+ Authentication” on page 106](#)
- [“Authentication, Authorization and Accounting” on page 109](#)

TACACS+ Authentication

CNOS supports authentication and authorization with networks using the Terminal Access Controller Access-Control System Plus (TACACS+) protocol. The G8272 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 G8272 either through a data port or management port.

TACACS+ offers the following advantages:

- TACACS+ uses TCP-based connection-oriented transport; TCP offers a connection-oriented transport, while UDP offers best-effort delivery.
- TACACS+ offers full packet encryption in authentication requests.
- TACACS+ separates authentication, authorization and accounting.

How TACACS+ Authentication Works

TACACS+ authentication follows the steps below:

1. Remote administrator connects to the switch and provides user name and password.
2. Using the authentication/authorization protocol, the switch sends request to authentication server.
3. Authentication server checks the request against the user ID database.
4. Using the 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.

Note: A username can have a maximum length of 32 characters, that includes only lowercase letters and numbers. The username must also start with a letter. Usernames that start with a number are invalid.

Note: A password can have a maximum length of 255 characters, that includes lowercase or uppercase letters, numbers and special characters.

TACACS+ Authentication Features in Cloud NOS

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 tries to gain access to its services. CNOS supports ASCII inbound login to the device.

Note: 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 CNOS management access levels is shown in [Table 4](#). The authorization levels must be defined on the TACACS+ server.

Table 4. *Default TACACS+ Authorization Levels*

CNOS User Access Level	TACACS+ level
operator	1
admin	15 or 16

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.

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 G8272 supports the following TACACS+ accounting attributes:

- protocol (Telnet/SSH)
- start_time
- stop_time

Configuring TACACS+ Authentication on the Switch

You can configure up to 4 TACACS+ authentication servers. To set a TACACS+ server on the switch, follow the steps below:

1. Enable TACACS+ on the switch:

```
G8272(config)# feature tacacs+
```

2. Configure the IPv4 or IPv6 addresses of a TACACS+ server. Optionally, you can specify the TACACS+ server port.

```
G8272(config)# tacacs-server host <IPv4 or IPv6 address> port <1-65535>
```

Note: If you want to configure a fifth TACACS+ server, you must first delete one of the existing four servers by using the following command:

```
G8272(config)# no tacacs-server host <IPv4 or IPv6 address>
```

3. Configure the TACACS+ encryption/decryption key.

- a. You can set up a global key for all TACACS+ servers:

```
G8272(config)# tacacs-server key [0|7] <1-63 character key>
```

- b. Or you can set up a key for each individual TACACS+ server:

```
G8272(config)# tacacs-server host <IPv4 or IPv6 address> key [0|7] <1-63 character key>
```

Note: Setting up a key for a specific TACACS+ server will overwrite the globally configured key.

Note: By default, the TACACS+ key is not encrypted. To specify the encryption level of the key, use the following parameters:

- o **0** - the key is unencrypted (clear text);
- o **7** - the key is encrypted.

4. Check the current TACACS+ settings and configured servers:

```
G8272> display tacacs-server
```

Authentication, Authorization and Accounting

Authentication, Authorization and Accounting (AAA) allows the switch to use secure protocols to access the switch. The AAA mechanism uses TACACS+ as an underlying protocol. For AAA to work properly, you must configure the TACACS+ protocol and a TACACS+ server. Once these are set up, all services can be configured through AAA. For more information about configuring TACACS+, see [“TACACS+ Authentication” on page 106](#).

AAA uses TACACS+ to provide the following services: authentication, authorization and accounting. After setting up a TACACS+ server, the switch must be configured through AAA to use the respective server.

Authentication is the service that AAA provides to authenticate users that can log in on the switch. Authentication is supported at login for either a remote protocol connection like SSH/Telnet (named default) or for the console.

Authorization is the service that AAA provides to give specific users specific rights. Some users may have administrator privileges, other users may execute only basic level commands. All this can be configured in the server that is used by AAA, which has to be set up to query the server for this information.

Accounting is the service that AAA provides to log all user login and logout actions on the switch. This is useful for administrators that want to keep a clear log of anyone logging in or attempting to do so.

AAA Groups

You can put already defined TACACS+ together under a common group using AAA. You can configure all the servers inside a group using a single set of commands. AAA groups have specific parameters that are applied for the feature using a certain group and for all the servers configured inside that group.

There are a few restrictions when configuring AAA groups:

- Each group must contain only servers that were previously defined using the same protocol.
- Each server can be a member of only one AAA group. If you add a server to a second group, it will be removed from the first group.
- Only three AAA groups can be used at the same time. Each group corresponds to one service: authentication, authorization or accounting.

When multiple servers are configured in an AAA group, the first one will be queried. If the query fails (there is no response from the first server), the next server is interrogated. Any answer (positive or negative) is considered authoritative for the query and it is final.

Group Lists

When configuring AAA services (authentication, authorization and accounting), you can set up a list of AAA groups for each service. The *group list* can have up to a maximum of 8 AAA groups, each separated by space. The groups are queried one after the other in the order they were specified. If the query fails (no answer from the AAA group), the next group is interrogated.

Group lists can end with 'local' or 'none' or both. If all other groups fail, then, if 'local' was specified, local authentication will be used, or, if 'none' was used, no authentication will be required.

Note: A group list must contain at least one other group besides 'local' or 'none'.

Configuring AAA Groups

To create and configure an AAA group follow the steps below:

1. Create the AAA group:

```
G8272(config)# aaa group server tacacs+ <group name>
G8272(config-tacacs)#
```

Note: After creating the group, you will enter AAA Group configuration mode.

2. Add the desired TACACS+ servers to the current group:

```
G8272(config-tacacs)# server <IPv4 or IPv6 address>
```

Note: Only servers that were already defined using TACACS+ can be added to the group. To configure a TACACS+ server, see [“Configuring TACACS+ Authentication on the Switch” on page 108](#).

3. Optionally, you can set the Virtual Routing and Forwarding (VRF) instance for the current group:

```
G8272(config-tacacs)# use-vrf <VRF instance>
```

4. Check the currently configured AAA groups:

```
G8272> display aaa groups
```

Authentication

Authentication can be configured for the two options separately. You can choose to use either the pluggable authentication module (PAM) by specifying a group of servers (see [“AAA Groups” on page 110](#)) or local authentication (local database on the switch). When using PAM, only the users defined on the authentication server can log in with the privilege levels and security configuration given by the server. The user database is kept remote and when the user tries to log in, AAA will query the defined server and provide access.

Role-based authentication is supported. When the application does not want to authorize each command separately, the switch will send a privilege level request to the AAA server. Then, with respect to the privileged level obtained, it will allow or disallow shell commands for the user.

For telnet or console connections the administrator doesn't need to create local users on the switch for them to log in. The user will be created on demand and will be kept indefinitely.

The administrator needs to create a username on the switch if he or she wants that user to log in using Secure Shell (SSH) via TACACS+. The user should be created without a password. The user can also be created by a previous login with telnet on the switch, if the user doesn't exist. SSH does additional checks compared to telnet and if a user does not exist, SSH denies access to that user.

Also a maximum number of failed login attempts can be set. When this number is reached by a user, he or she is locked out of the switch. Only the system administrator cannot be locked out. Any other user can be unlocked by using the following command in Privileged EXEC mode:

```
G8272# remove aaa local user lockout username <username>
```

Configuring AAA Authentication

To configure authentication follow the steps below:

1. Configure authentication for console sessions:
 - a. Configure authentication to use an AAA group:

```
G8272(config)# aaa authentication login console group <group list>
```

- b. Configure authentication to use the local database:

```
G8272(config)# aaa authentication login console local
```

- c. Disable authentication:

```
G8272(config)# aaa authentication login console none
```


2. Configure authentication for telnet or SSH sessions:

a. Configure authentication to use an AAA group:

```
G8272(config)# aaa authentication login default group <group list>
```

b. Configure authentication to use the local database:

```
G8272(config)# aaa authentication login default local
```

c. Disable authentication:

```
G8272(config)# aaa authentication login default none
```

3. Configure the maximum number of failed login attempts:

```
G8272(config)# aaa local authentication attempts max-fail <1-25>
```

4. Optionally, you can configure the display of error messages for failed login attempts:

```
G8272(config)# aaa authentication login error-enable
```

5. Check the current AAA authentication settings:

```
G8272> display aaa authentication
```

In case no TACACS+ servers are reachable, you can configure local authentication by using the following commands:

- for console sessions:

```
G8272(config)# aaa authentication login console group <group list> local
```

- for telnet or SSH sessions:

```
G8272(config)# aaa authentication login default group <group list> local
```

Note: The switch will first try to contact the TACACS+ servers configured in the specified AAA group. If this fails, it will then turn to local authentication.

Authorization

Authorization makes a separation between commands available in Privileged EXEC mode and those available in Global Configuration mode. Each of the two categories can be configured separately using AAA groups. Every command is sent to the TACACS+ server and is executed only if the server authorizes it.

Configuring AAA Authorization

To configure authorization follow the steps bellow:

1. Configure authorization for Privileged EXEC commands:

- a. Configure authorization to use an AAA group:

```
G8272(config)# aaa authorization commands default group <group list>
```

- b. Configure authorization to use the local database:

```
G8272(config)# aaa authorization commands default local
```

2. Configure authorization for Global Configuration commands:

- a. Configure authorization to use an AAA group:

```
G8272(config)# aaa authorization config-commands default group  
<group list>
```

- b. Configure authorization to use the local database:

```
G8272(config)# aaa authorization config-commands default local
```

3. Check the current AAA authorization settings:

```
G8272> display aaa authorization
```

Accounting

Accounting is enabled by selecting the group of servers where accounting information should be stored. Local accounting can also be enabled. For TACACS+ only login/logout and failed login information will be logged.

Configuring AAA Accounting

To configure accounting follow the steps bellow:

1. Configure accounting to use one of the following:
 - a. an AAA group:

```
G8272(config)# aaa accounting default group <group list>
```

- b. the local database:

```
G8272(config)# aaa accounting default local
```

2. Check the current AAA accounting settings:

```
G8272> display aaa accounting
```

Chapter 6. Access Control Lists

An Access Control Lists (ACL) is a list of filters that permit or deny traffic for security purposes. They can also be used with Quality of Service (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 the actions that are performed when a match is made.

This section discusses the following topics:

- [“Supported ACL Types” on page 118](#)
- [“Summary of Packet Classifiers” on page 119](#)
- [“Summary of ACL Actions” on page 121](#)
- [“Configuring Port ACLs \(PACLs\)” on page 122](#)
- [“Configuring Router ACLs \(RACLs\)” on page 123](#)
- [“Configuring VLAN ACLs \(VACLs\)” on page 124](#)
- [“Configuring Management ACLs \(MACLs\)” on page 126](#)
- [“ACL Order of Precedence” on page 127](#)
- [“Creating and Modifying ACLs” on page 128](#)
- [“Viewing ACL Rule Statistics” on page 132](#)
- [“ACL Configuration Examples” on page 133](#)

Supported ACL Types

Cloud Network OS 10.1 supports the following types of ACLs:

- Internet Protocol version 4 (IPv4) ACLs:

IPv4 ACLs classify packets using Layer 3 and Layer 4 header fields of IPv4 packets. The following command creates an IPv4 ACL:

```
G8272(config)# ip access-list <ACL name>
```

Note: ACLs are not supported with IPv6.

- Media Access Control (MAC) ACLs:

MAC ACLs classify packets using Layer 2 fields. The following command creates a MAC ACL:

```
G8272(config)# mac access-list <MAC ACL name>
```

- Address Resolution Protocol (ARP) ACLs:

An ARP ACL filters ARP request and reply packets using various ARP fields. The following command creates an ARP ACL:

```
G8272(config)# arp access-list <MAC ACL name>
```

Based on where they are applied, ACLs can be classified as follows:

- Port-Based ACLs (PACLs)

A PACL can be applied to all packets received on a switched port. The PACLs can be MAC or IPv4 PACLs.

- Router ACLs (RACLs)

A RACL can be applied to all packets received on a routed port (ingress RACL) or all packets going out of a routed port (egress RACL). The RACLs can be MAC or IPv4 RACLs.

- VLAN ACLs (VACLs)

A VACL is an ACL that can be assigned to a VLAN rather than to a switch port as with IPv4 ACLs. Individual VACL filters are configured in a similar fashion to IPv4 ACLs.

- Management ACLs (MACLs)

MACLs filter inbound traffic heading towards the CPU through the management port.

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 and destination port number. Once classified, packet flows can be identified for more processing.

IPv4 ACLs, MAC ACLs, ARP ACLs, Port ACLs, Router ACLs, Management ACLs and VLAN ACLs allow you to classify packets based on the following:

- IPv4 ACLs:
 - AH Packet
 - any protocol packet
 - EIGRP protocol packet
 - ESP packet
 - GRE packet
 - IANA assigned protocol number (0-255)
 - ICMP Packet
 - IGMP packet
 - IPv4 Encapsulation packet
 - KA9Q NOS compatible IP over IP tunneling
 - OSPF packet
 - IPCOMP packet
 - PIM packet
 - TCP Packet
 - UDP Packet
 - Source IPv4 address and subnet mask
 - Destination IPv4 address and subnet mask
 - Type of Service value
 - TCP/UDP application source host and mask
 - TCP/UDP source port as shown in [Table 5](#)
 - TCP/UDP application destination host and mask
 - TCP/UDP destination port as shown in [Table 5](#)
 - TCP flag value as shown in [Table 6](#)
- MAC ACLs:
 - Source MAC address
 - Destination MAC address
 - Ethernet type (ARP, IPv4, MPLS, RARP, etc.)
 - Ethernet Priority (IEEE 802.1p Priority)
 - VLAN number and mask

- ARP ACLs:
 - Source IPv4 address and subnet mask
 - Destination IPv4 address and subnet mask
 - Source MAC address
 - Destination MAC address
- Port ACLs (PACLs):
 - IPv4 ACLs
 - MAC ACLs
- Router ACLs (RACLs):
 - IPv4 ACLs
 - MAC ACLs
- Management ACLs (MACLs):
 - IPv4 ACLs
- VLAN ACLs (VACLs):
 - IPv4 ACLs
 - MAC ACLs

Table 5. *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

Table 6. *Well Known TCP Flag Values*

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

Summary of ACL Actions

An ACL must have a set of rules to classify the packet. The first token of each rule indicates the action to be performed on those packets that match the rule.

The following actions are supported:

- All ACL types support Permit and Drop actions.
- VACLs also support Redirect actions.

Note: All ACLs have an implicit default deny action to drop those packets that do not match any of the ACL rules. This default action applies to all IPv4 packets in an IPv4 ACL, and to any packet for MAC and ARP ACLs on the respective port or VLAN.

Configuring Port ACLs (PACLs)

A PACL is configured if an ACL is assigned to a Layer 2 switched interface. The following is a sample PACL configuration:

```
G8272(config)# interface ethernet 1/5
G8272(config-if)# bridge-port
G8272(config-if)# ip port access-group <ACL name> in
G8272(config-if)# exit
```

```
G8272(config)# interface ethernet 1/10
G8272(config-if)# bridge-port
G8272(config-if)# mac port access-group <ACL name>
G8272(config-if)# exit
```

Configuring Router ACLs (RACLs)

A RACL is configured if an ACL is assigned to a Layer 3 routed interface. The following is a sample PACL configuration:

```
G8272(config)# interface ethernet 1/5
G8272(config-if)# no bridge-port
G8272(config-if)# ip access-group <ACL name> {in|out}
G8272(config-if)# exit
```

While a PACL, a RACL or both can be configured to an interface, only one of them will be active on that interface, depending upon whether that interface is a Layer 2 or a Layer 3 interface. However, two ACLs of the same type cannot be configured on the same interface.

A packet can match different ACL types at the same time. When that happens, actions are taken according to the order of precedence.

Configuring VLAN ACLs (VACLs)

A VACL is an ACL that can be assigned to a VLAN 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.

VACLs can be configured for all the 4K VLAN IDs.

Individual VACL filters are configured in a similar fashion to IPv4 ACLs.

Follow the steps below to create or change a VACL:

1. From Configuration Mode, enter VLAN Access Map Configuration Mode:

```
G8272(config)# vlan access-map <ACL name>
G8272(config-access-map)#
```

2. Enter an action for the VACL to take.

To drop or forward matching packets, enter:

```
G8272(config-access-map)# action {drop|forward}
```

To redirect matching packets, enter:

```
G8272(config-access-map)# action redirect {ethernet <chassis number/port
number>|port-aggregation <LAG number>}
```

Note: Because the redirected packet will be subjected to the usual VLAN membership check, make sure the interface to where the packet is redirected has the same VLAN membership as the VACL.

3. Enter a match for the VACL to use.

- To configure a rule to match an IP ACL, enter:

```
G8272(config-access-map)# match ip address <ACL name>
```

- To configure a rule to match a MAC ACL, enter:

```
G8272(config-access-map)# match mac address <MAC ACL name>
```

4. To preserve your configuration in case of a reboot, from Privileged EXEC mode, enter:

```
G8272# cp running-config startup-config
```

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

- To assign a VACL filter to a VLAN, enter:

```
G8272(config)# vlan filter <VACL name> vlan-list <VLAN number>
```

- To remove a VACL filter from a VLAN, enter:

```
G8272(config)# no vlan filter <VACL name> vlan-list <VLAN number>
```

VACL Configuration Example

In this example, traffic from VLAN 3 ports is forwarded.

```
G8272(config)# vlan access-map myVACL  
G8272(config-access-map)# action forward  
G8272(config-access-map)# exit  
  
G8272(config)# vlan filter myVACL vlan-list 3
```

Configuring Management ACLs (MACLs)

MACLs filter ingress or egress traffic through the management port. A MACL is configured the same way as an IPv4 or a MAC ACL and needs to be assigned on the management interface (mgmt 0).

In the following example, traffic to the switch management interface is permitted only for source IP address 12.12.12.12:

```
G8272(config)# ip access-list myMACL
G8272(config-acl)# 10 permit ip host 12.12.12.12 any
G8272(config-acl)# exit

G8272(config)# interface mgmt 0
G8272(config-if)# ip access-group myMACL in
G8272(config-if)# exit
```

Use the following command to view the MACL configuration:

```
G8272(config)# display access-list myMACL
IP access list myMACL
    10 permit ip host 12.12.12.12 any
```

ACL Order of Precedence

An ACL consists of a set of rules that are used to classify or filter incoming or outgoing traffic. The order in which each rule is applied to the port traffic will depend upon its sequence number. A rule with a lower sequence number will take precedence over a rule with a higher sequence number. If none of the rules are matched, the switch will apply an implicit 'deny all' rule, which drops all the traffic.

Consider the following example:

```
G8272(config)# ip access-list acl-1
G8272(config-acl)# 10 permit icmp any any
G8272(config-acl)# 20 deny ip any any
```

In the above configuration, the rule with sequence number 10 takes higher precedence over the rule with sequence number 20.

The packet will be compared to the rules starting from the lowest-numbered rule through the highest-numbered rule. If any rule matches the packet, that packet will not be compared to any more rules in that ACL.

A packet can match different types of ACLs that are configured on the same interface. If the packet matches rules in multiple ACLs, the resulting action will be in the following order:

- 1.Drop
- 2.Redirect
- 3.Permit or Forward

All statistics are counted for each matching rule.

Creating and Modifying ACLs

Each ACL is defined by a group of rules. You can create the following types of ACLs:

- **Individual IPv4 ACLs**

Each command defines one filter rule for matching traffic criteria. Each filter rule must also include an action (permit or deny the packet). For example:

```
G8272(config)# ip access-list myACL
G8272(config-acl)# permit udp any lt 301 10.2.2.1 0.0.0.255
```

This example configures ACL 'myACL' to permit UDP traffic from any host, that has a UDP source port smaller than 301, and that goes to hosts that belong to the network 10.2.2.0/24. All other traffic will be dropped by the implicit 'deny' rule.

- **MAC ACLs**

Each command defines one filter rule for matching traffic criteria. Each filter rule must also include an action (permit or deny the packet). For example:

```
G8272(config)# mac access-list myMAC-ACL
G8272(config-mac-acl)# permit 001b.6384.45e6 0000.0000.ffff
8c89.a502.93af 0000.0000.ffff cos 2
```

This example configures MAC ACL 'myMAC-ACL' to permit traffic from MAC host 001B.6384.45E6 with source wildcard 0000.0000.FFFF that goes to MAC host 8C89.A502.93AF with destination wildcard 0000.0000.FFFF using Class of Service (CoS) 2. All other traffic will be dropped by the implicit 'deny' rule.

- **ARP ACLs**

Each command defines one filter rule for matching traffic criteria. Each filter rule must also include an action (permit or deny the packet). For example:

```
G8272(config)# arp access-list myARP-ACL
G8272(config-arp-acl)# permit ip any mac 0007.e94d.4a85 0000.0000.ffff
```

This example configures ARP ACL 'myARP-ACL' to permit traffic from any IP source address with source MAC address 0007.E94D.4A85 and source wildcard 0000.0000.FFFF. All other traffic will be dropped by the implicit 'deny' rule.

Creating an IPv4 ACL

To create an IPv4 ACL, follow these steps:

1. From Configuration Mode, enter ACL Configuration Mode for the ACL:

```
G8272(config)# ip access-list <ACL name>
G8272(config-acl)#
```

Note: The ACL name is case-sensitive.

2. Assign rules to the ACL using the following syntax:

```
G8272(config-acl)# [<sequence number>] {deny|permit} <protocol> <source IP address>
<source wildcard IP mask> [<operand> <port number>] <destination IP address> <destination wildcard
IP mask> [<operand> <port number>] [<bit to match on>]
```

Note: You can enter more than one rule for an ACL.

3. Exit ACL Configuration Mode:

```
G8272(config-acl)# exit
```

To see the access list you have created, enter:

```
G8272(config)# display ip access-lists <ACL name>
```

4. To save your configuration, from Privileged EXEC mode, enter:

```
G8272# cp running-config startup-config
```

Note: To change a filtering rule, you must first remove it using the **no** command in front of the command that created the group:

```
G8272(config-acl)# no [<sequence number>] {deny|permit} <protocol> <source IP
address> <source wildcard IP mask> [<operand> <port number>] <destination IP address> <destination
wildcard IP mask> [<operand> <port number>] [<bit to match on>]
```

Removing an IPv4 ACL

To remove an ACL, use the following command:

```
G8272(config)# no ip access-list <ACL name>
```

Resequencing an IPv4 ACL

Sometimes you may want to fit a rule between two existing numbers. Use the **resequence** command in Configuration Mode to change the numbering scheme but keep the existing rules in order:

```
G8272(config)# resequence ip access-list <ACL name> <starting number> <increment>
```

Creating a MAC ACL

To create a MAC ACL, follow these steps:

1. From Configuration Mode, enter MAC ACL Configuration Mode for the ACL:

```
G8272(config)# mac access-list <ACL name>
G8272(config-mac-acl)#
```

Note: The ACL name is case-sensitive.

2. Assign rules to the ACL using the following syntax:

```
G8272(config-mac-acl)# [<sequence number>] {deny|permit} {{<source MAC address>
 [<source wildcard MAC mask>]}|any|host <source MAC address>} {{<destination MAC address>
 [<destination wildcard MAC mask>]}|any|host <destination MAC address>} {<protocol>|cos
 <cos>|vlan <VLAN number>}[cos <cos>] [vlan <VLAN number>]
```

Note: You can enter more than one rule for an ACL.

3. Exit MAC ACL Configuration Mode:

```
G8272(config-mac-acl)# exit
```

To see the access list you have created, enter:

```
G8272(config)# display mac access-lists <ACL name>
```

4. To save your configuration, from Privileged EXEC mode, enter:

```
G8272# cp running-config startup-config
```

Note: To change a filtering rule, you must first remove it using the **no** command in front of the command that created the group:

```
G8272(config-mac-acl)# no [<sequence number>] {deny|permit} {{<source MAC
 address> [<source wildcard MAC mask>]}|any|host <source MAC address>} {{<destination MAC
 address> [<destination wildcard MAC mask>]}|any|host <destination MAC address>} {<protocol>|
 cos <cos>|vlan <VLAN number>}[cos <cos>] [vlan <VLAN number>]
```

Removing a MAC ACL

To remove a MAC ACL, use the following command:

```
G8272(config)# no mac access-list <ACL name>
```

Resequencing a MAC ACL

Sometimes you may want to fit a rule between two existing numbers. Use the **resequence** command in Configuration Mode to change the numbering scheme but keep the existing rules in order:

```
G8272(config)# resequence mac access-list <ACL name> <starting number>
<increment>
```

Creating an ARP ACL

To create an ARP ACL, follow these steps:

1. From Configuration Mode, enter ARP ACL Configuration Mode for the ACL:

```
G8272(config)# arp access-list <ACL name>
G8272(config-arp-acl)#
```

Note: The ACL name is case-sensitive.

2. Assign rules to the ACL using the following syntax:

```
G8272(config-arp-acl)# [<sequence number>] {deny|permit} [request|response]
ip {<source IPv4 address>|any|host <single source address> {mac {<source MAC address>
<source wildcard MAC mask>|any|host <source MAC address>}}}
```

Note: You can enter more than one rule for an access list.

3. Exit ARP ACL Configuration Mode:

```
G8272(config-arp-acl)# exit
```

To see the access list you have created, enter:

```
G8272(config)# display arp access-lists <ACL name>
```

4. To save your configuration, from Privileged EXEC mode, enter:

```
G8272# cp running-config startup-config
```

Note: To change a filtering rule, you must first remove it using the no command in front of the command that created the group:

```
G8272(config-arp-acl)# no [<sequence number>] {deny|permit} [request|
response] ip {<source IPv4 address>|any|host <single source address> {mac {<source MAC
address> <source wildcard MAC mask>|any|host <source MAC address>}}}
```

Removing an ARP ACL

To remove an ARP ACL, use the following command:

```
G8272(config)# no arp access-list <ACL name>
```

Resequencing an ARP ACL

Sometimes you may want to fit a rule between two existing numbers. Use the **resequence** command in Configuration Mode to change the numbering scheme but keep the existing rules in order:

```
G8272(config)# resequence arp access-list <ACL name> <starting number>
<increment>
```

Viewing ACL Rule Statistics

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

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

```
G8272(config)# ip access-list myacl
G8272(config-acl)# statistics per-entry
G8272(config-acl)# exit

G8272(config)# display ip access-lists
```

ACL Configuration Examples

The following are examples of configuring ACLs:

ACL Example 1

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

1. Create an Access Control List.

```
G8272(config)# ip access-list myACL1
G8272(config-acl)#
```

2. Configure the Access Control List.

```
G8272(config-acl)# deny any any 100.10.1.1
G8272(config-acl)# permit any any any
G8272(config-acl)# exit
```

3. Add ACL 'myACL1' to ethernet interface 1/1 incoming traffic.

```
G8272(config)# interface ethernet 1/1
G8272(config-if)# ip port access-group myACL1 in
G8272(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 on ethernet port 1/2 with source IP from network 100.10.1.0/24 and destination IP 200.20.2.2 is denied.

1. Create an Access Control List.

```
G8272(config)# ip access-list myACL2
G8272(config-acl)#
```

2. Configure the Access Control List.

```
G8272(config-acl)# deny any 100.10.1.0/24 host 200.20.2.2
G8272(config-acl)# permit any any any
G8272(config-acl)# exit
```

3. Add ACL 'myACL2' to ethernet interface 1/2.

```
G8272(config)# interface ethernet 1/2
G8272(config-if)# ip port access-group myACL2 in
G8272(config-if)# exit
```

ACL Example 3

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

1. Create an Access Control List.

```
G8272(config)# ip access-list myACL3
G8272(config-acl)#
```

2. Configure the Access Control List.

```
G8272(config-acl)# deny igmp any any
G8272(config-acl)# permit any any any
G8272(config-acl)# exit
```

3. Add ACL 'myACL3' to ethernet interface 1/3.

```
G8272(config)# interface ethernet 1/3
G8272(config-if)# ip port access-group myACL3 in
G8272(config-if)# exit
```

ACL Example 4

Use the following configuration to permit access from any host to hosts with destination MAC address 11:05:00:10:00:00 and MAC wildcard FF:F5:FF:FF:FF:FF, and deny access to all other hosts.

1. Create a MAC Access Control List.

```
G8272(config)# mac access-list myMAC-ACL4
G8272(config-mac-acl)#
```

2. Configure the MAC Access Control List.

```
G8272(config-mac-acl)# permit any 11:05:00:10:00:00 FF:F5:FF:FF:FF:FF
G8272(config-mac-acl)# deny any any
G8272(config-mac-acl)# exit
```

3. Add MAC ACL 'myMAC-ACL4' to ethernet interface 1/4.

```
G8272(config)# interface ethernet 1/4
G8272(config-if)# mac port access-group myMAC-ACL4
G8272(config-if)# exit
```

ACL Example 5

This configuration blocks egress traffic from a specific network to any destination. All traffic from the network 100.10.1.0/24 to any destination egressing from routed port 3 is denied.

1. Create an Access Control List.

```
G8272(config)# ip access-list myACL5
G8272(config-acl)#
```

2. Configure the Access Control List.

```
G8272(config-acl)# deny any 100.10.1.0/24 any
G8272(config-acl)# permit any any any
G8272(config-acl)# exit
```

3. Add ACL 'myACL5' to ethernet interface 5.

```
G8272(config)# interface ethernet 1/5
G8272(config-if)# ip port access-group myACL5 out
G8272(config-if)# exit
```

ACL Example 6

This configuration creates an ARP ACL named 'myARP-ACL6':

1. Create the ARP ACL:

```
G8272(config)# arp access-list myARP-ACL6
G8272(config-arp-acl)#
```

2. Add rules to the ACL:

```
G8272(config-arp-acl)# statistics per entry
G8272(config-arp-acl)# 10 deny ip any mac host 0033.3333.3333
G8272(config-arp-acl)# 20 permit ip any mac 0044.4444.4444
0000.0000.FFFF
G8272(config-arp-acl)# 30 permit ip host 20.20.20.2 mac host
0011.1111.1111
G8272(config-arp-acl)# 50 permit ip host 40.40.40.2 mac 0055.5555.5555
0000.0000.FFFF
G8272(config-arp-acl)# 60 deny ip host 50.50.50.2 mac any
G8272(config-arp-acl)# 70 deny ip 30.30.30.2 0.0.0.255 mac
0022.2222.2222 0000.0000.FFFF
G8272(config-arp-acl)# 80 deny ip 40.40.40.2 0.0.0.255 mac any
G8272(config-arp-acl)# 90 permit ip 50.50.50.2 0.0.0.255 mac host
0055.5555.5555
G8272(config-arp-acl)# exit
```

3. Configure the inspection filter:

```
G8272(config)# ip arp inspection filter myARP-ACL vlan 20,30,40,50
```


Part 3: Switch Basics

This section discusses basic switching functions:

- [“Interface Management” on page 139](#)
- [“Forwarding Database” on page 153](#)
- [“VLANs” on page 157](#)
- [“Ports and Link Aggregation” on page 179](#)
- [“Multiple Spanning Tree Protocol” on page 197](#)
- [“Virtual Link Aggregation Groups” on page 211](#)
- [“Quality of Service” on page 241](#)

Chapter 7. Interface Management

The term interface can refer to a physical switch port or a logical port, a VLAN, a Link Aggregation Group (LAG) etc.

This section discusses the following topics:

- [“Interface Management Overview” on page 140](#)
- [“Management Interface” on page 141](#)
- [“Physical Ports” on page 142](#)
- [“Port Aggregation” on page 144](#)
- [“Loopback Interfaces” on page 145](#)
- [“Switch Virtual Interfaces” on page 146](#)
- [“Basic Interface Configuration” on page 147](#)

Interface Management Overview

Interface Management (IFM) is part of the system infrastructure that acts as a logical interface abstraction for software components (applications) running on the switch. It is provided by the Network Service Module (NSM). NSM follows a server-client model and, as a server, it handles interface interactions with other switch applications, which are clients of the NSM server.

NSM clients include applications like: Neighbor Discovery Protocol (NDP), Routing Information Base (RIB), Spanning Tree Protocol (STP), Link Aggregation Control Protocol (LACP), Internet Group Management Protocol (IGMP), Border Gateway Protocol (BGP) and more. NSM clients subscribe to the services provided by the NSM server to deal with interface events, like a link going up or down. When an interface event occurs, NSM handles the event accordingly and then notifies all of its clients.

All interfaces are assigned a unique internal identifier called an interface index (*ifindex*). The *ifindex* is a 32-bit integer attributed to the interface by NSM. Once an *ifindex* has been assigned to an interface, it is forever bound to it.

During CLI interactions with the switch, interfaces are identified by their names, and not their *ifindex*.

To view the *ifindex* of all interfaces, use the following command:

```
G8272> display interface snmp-ifindex
```

Port	IFMIB Ifindex (hex)
Ethernet1/1	410010 (0x0006419a)
Ethernet1/2	410020 (0x000641a4)
Ethernet1/3	410030 (0x000641ae)
Ethernet1/4	410040 (0x000641b8)
Ethernet1/5	410050 (0x000641c2)
Ethernet1/6	410060 (0x000641cc)
Ethernet1/7	410070 (0x000641d6)
Ethernet1/8	410080 (0x000641e0)
Ethernet1/9	410090 (0x000641ea)
Ethernet1/10	410100 (0x000641f4)
...	

Note: An interface can't be configured until it is created. Also, if the interface is removed, its configuration is lost. For example, the configuration for a 40 Gigabit Ethernet (GbE) port will be lost once the port is split into four 10 GbE breakout ports. There are some interfaces created by default on the switch that can't be removed.

The G8272 supports both IPv4 and IPv6 addressing on Layer 3 interfaces, such as routed ports, loopback interfaces or Switch Virtual Interfaces. For more details, see [Chapter 14, "Basic IP Routing"](#) and [Chapter 17, "Internet Protocol Version 6"](#).

Management Interface

The Management Interface is a special physical port on the switch that allows you to perform management tasks.

It is a Layer 3 interface and can't be configured as a Layer 2 port.

The management interface can't forward traffic. To keep it separate from other interfaces, the management port is part of the management Virtual Routing and Forwarding (VRF) instance, while other interfaces are part of the default VRF instance.

To configure the management interface, use the following command:

```
G8272(config)# interface mgmt 0
G8272(config-if)#
```

Note: After running the above command, you will enter Interface Configuration mode for the management port.

Note: By default, it is created by the switch and it can't be deleted.

To view detailed information about the management interface, use the following command:

```
G8272(config)# display interface mgmt 0

Interface mgmt0
  Hardware is Management Ethernet   Current HW addr: a897.dcde.2500
  Physical:a897.dcde.2500   Logical:(not set)
  index 3 metric 1 MTU 1500 Bandwidth 1000000 Kbit
  no bridge-port
  arp ageing timeout 1500
  <UP,BROADCAST,RUNNING,ALLMULTI,MULTICAST>
  VRF Binding: Associated with management
  Speed auto Duplex full
  IPV6 DHCP IA-NA client is enabled.
  inet 10.241.41.27/25 broadcast 10.241.41.127
  inet6 fe80::aa97:dcff:fede:2500/64
  RX
    131249 input packets 5 unicast packets 128053 multicast packets
    3191 broadcast packets 23903151 bytes
  TX
    6749 output packets 0 unicast packets 6749 multicast packets
    0 broadcast packets 939061 bytes
```

Physical Ports

The G8272 is a 1U rack-mountable aggregation switch. It uses a wire-speed, non-blocking switching fabric that provides simultaneous wire-speed transport of multiple packets at low latency on all ports.

The G8272 has the following port capabilities:

- Forty-eight 10 Gigabit Ethernet (GbE) Small Form Pluggable Plus (SFP+) ports which also support legacy 1 GbE connections;
- Six 40 GbE Quad Small Form Pluggable Plus (QSFP+) ports, each of which can optionally be used as four 10 GbE sports;

SFP+ ports, when used in 10 GbE mode, can be populated with optical transceiver modules or active or passive DACs. When used in legacy 1 GbE mode, the ports can be populated with optical or copper transceiver modules.

QSFP+ ports can be populated with optical QSFP+ modules or passive Direct-Attach Cables (DACs), including those that allow the port to be used as four 10 GbE ports.

By default, the G8272 has the following interface (port) configuration:

- ethernet interfaces 1-48 are configured as 10Gbit/s ports;
- ethernet interfaces 49-54 are configured as 40Gbit/s ports;

During CLI interactions, physical ports are referred to as ethernet interfaces and have the following port numbering scheme:

- 40 or 10 GbE ports: *<chassis number>/<port number>*

For example, physical port 10:

```
interface ethernet 1/10
```

- 10 GbE breakout ports: *<chassis number>/<port number>/<subport number>*

For example, breakout port 3 of physical port 52:

```
interface ethernet 1/52/3
```

In CLI commands, ethernet interfaces can be used in the following ways:

- as a single interface:

```
G8272(config)# interface ethernet 1/10
```

- as a continuous range:

```
G8272(config)# interface ethernet 1/10-24
```

- as a discrete range:

```
G8272(config)# interface ethernet 1/10, ethernet 1/20-22
```

To configure an ethernet interface, use the following command:

```
G8272(config)# interface ethernet <chassis number>/<port number>
G8272(config-if)#
```

For example:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)#
```

Note: After running the above command, you will enter Interface Configuration mode for the specified ethernet port.

Note: Ethernet interfaces can't be removed using the CLI.

To view detailed information about an ethernet interface, use the following command:

```
G8272(config)# display interface ethernet <chassis number>/<port number>

Interface Ethernet1/1
  Hardware is Ethernet   Current HW addr: a897.dcde.2503
  Physical:a897.dcde.2503   Logical:(not set)
  index 410010 metric 1 MTU 1500 Bandwidth 10000000 Kbit
  Port Mode is trunk
  <UP,BROADCAST,MULTICAST>
  VRF Binding: Not bound
  Speed 10000 Mb/s Duplex full
  Last link flapped never
  Last clearing of "display interface" counters never
  30 seconds input rate 0 bits/sec, 0 bytes/sec, 0 packets/sec
  30 seconds output rate 0 bits/sec, 0 bytes/sec, 0 packets/sec
  Load-Interval #2: 5 minute (300 seconds)
    input rate 0 bps, 0 pps; output rate 0 bps, 0 pps
  RX
    0 unicast packets  0 multicast packets  0 broadcast packets
    0 input packets  0 bytes
    0 jumbo packets  0 storm suppression packets
    0 giants  0 input error  0 short frame  0 overrun  0 underrun
    0 watchdog  0 if down drop
    0 input with dribble  0 input discard(includes ACL drops)
    0 Rx pause
  TX
    0 unicast packets  0 multicast packets  0 broadcast packets
    0 output packets  0 bytes
    0 jumbo packets
    0 output errors  0 collision  0 deferred  0 late collision
    0 lost carrier  0 no carrier  0 babble
    0 Tx pause
  0 interface resets
```

Port Aggregation

Multiple physical ports can be aggregated to act as a single logical port called a Link Aggregation Group (LAG). Port aggregation provides link redundancy, increased bandwidth and traffic balancing across the ethernet ports.

By default, no physical ports are aggregated. Thus, no LAGs are configured on the switch.

To create or modify an existing LAG, use the following command:

```
G8272(config)# interface port-aggregation <1-4096>
G8272(config-if)#
```

Note: After running the above command, you will enter Interface Configuration mode for the specified LAG.

To delete a LAG, use the following command:

```
G8272(config)# no interface port-aggregation <1-4096>
```

To view detailed information about a LAG, use the following command:

```
G8272(config)# display interface port-aggregation <1-4096>

Interface po1
  Hardware is AGGREGATE   Current HW addr: a897.dcde.0203
  Physical:(not set)   Logical:(not set)
  index 100001 metric 1 MTU 1500 Bandwidth 0 Kbit
  Port Mode is access
  <UP,BROADCAST,MULTICAST>
  VRF Binding: Not bound
  Members in this port-aggregation:
    Ethernet1/1
  Last clearing of "display interface" counters never
  30 seconds input rate 0 bits/sec, 0 bytes/sec, 0 packets/sec
  30 seconds output rate 0 bits/sec, 0 bytes/sec, 0 packets/sec
  Load-Interval #2: 5 minute (300 seconds)
    input rate 0 bps, 0 pps; output rate 0 bps, 0 pps
  RX
    0 unicast packets  0 multicast packets  0 broadcast packets
    0 input packets  0 bytes
    0 jumbo packets  0 storm suppression packets
    0 giants  0 input error  0 short frame  0 overrun  0 underrun
    0 watchdog  0 if down drop
    0 input with dribble  0 input discard(includes ACL drops)
    0 Rx pause
  TX
    0 unicast packets  0 multicast packets  0 broadcast packets
    0 output packets  0 bytes
    0 jumbo packets
    0 output errors  0 collision  0 deferred  0 late collision
    0 lost carrier  0 no carrier  0 babble
    0 Tx pause
  0 interface resets
```

For more details about port aggregation, see [Chapter 10, "Ports and Link Aggregation"](#).

Loopback Interfaces

A loopback interface is a virtual Layer 3 interface. It is usually used by different applications to communicate when they run on the same switch. Loopback interfaces are used exclusively on the switch and they don't pass packets to other interfaces. Any traffic that is sent to a loopback interface is immediately passed back to the switch as if it was received from another device. If a packet is routed to a loopback interface, but it's not destined for that interface, it will be discarded.

It is used to allow Border Gateway Protocol (BGP) sessions to be always up, even when the outgoing switch interface is not available. BGP sessions can use a loopback interface as their termination address.

Open Shortest Path First (OSPF) and other Layer 3 protocols can use the IP address of a loopback interface as their router ID.

To create or configure a loopback interface, use the following command:

```
G8272(config)# interface loopback <0-7>
G8272(config-if)#
```

Note: After running the above command, you will enter Interface Configuration mode for the specified loopback interface.

To delete a loopback interface, use the following command:

```
G8272(config)# no interface loopback <0-7>
```

Note: By default, loopback interface 0 is created on the switch and it can't be deleted.

Note: Loopback interfaces can't be configured as Layer 2 interfaces. They are automatically created as Layer 3 interfaces and this behavior can't be changed.

To view detailed information about a loopback interface, use the following command:

```
G8272(config)# display interface loopback <0-7>

Interface loopback0
  Hardware is Loopback
  index 8 metric 1 MTU 1500 Bandwidth 0 Kbit
  no bridge-port
  arp ageing timeout 1500
  <UP, LOOPBACK, RUNNING>
  VRF Binding: Not bound
  DHCP client is disabled.
  Encapsulation LOOPBACK
    0 packets input 0 bytes
    0 multicast frames 0 compressed
    0 input errors 0 frame 0 overrun 0 fifo
    0 packets output 0 bytes 0 underruns
    0 output errors 0 collisions 0 fifo
```

Switch Virtual Interfaces

A Switch Virtual Interface (SVI) is a VLAN that is used as a virtual Layer 3 interface. When a VLAN is created, its associated SVI is also created. A SVI is deleted when its related VLAN is also deleted.

By default, VLAN 1 is created on the switch, and thus, VLAN 1 SVI is also created.

A SVI as an ordinary Layer 3 interface:

- IP addresses can be configured;
- Layer 2 commands are not allowed;
- supports all features that can run on Layer 3 interfaces.

A VLAN can be associated with only one SVI. For more information about VLANs, see [Chapter 9, “VLANs”](#).

Note: You can configure up to a maximum of 256 SVIs.

Note: Before configuring a SVI interface, you must first create the appropriate Layer 2 VLAN. For more details, see [“Creating a VLAN” on page 160](#).

To configure a SVI, use the following command:

```
G8272(config)# interface vlan <1-4094>
G8272(config-if)#
```

Note: After running the above command, you will enter Interface Configuration mode for the specified SVI.

To remove the configuration of a SVI interface, you must delete its associated VLAN. For more details, see [“Deleting a VLAN” on page 161](#).

Note: The SVI associated with VLAN 1 can't be deleted, because VLAN 1 also cannot be removed from the switch.

To view detailed information about a SVI, use the following command:

```
G8272(config)# display interface vlan <1-4094>

Interface Vlan1
  Hardware is VLAN   Current HW addr: a897.dcde.2501
  Physical:(not set) Logical:(not set)
  index 9 metric 1 MTU 1500 Bandwidth 0 Kbit
  no bridge-port
  arp ageing timeout 1500
  <UP,BROADCAST,RUNNING,MULTICAST>
  VRF Binding: Not bound
  DHCP client is disabled.
  Last clearing of "display interface" counters never
```

Basic Interface Configuration

This section covers some of the basic options available when configuring an interface, like its port speed, MAC address or description.

To configure an interface, you must enter its Configuration mode.

To view a brief overview of all interfaces, use the following command:

```
G8272(config)# display interface brief
```

Ethernet Interface	VLAN	Type	Mode	Status	Reason	Speed	Port Agg#
Ethernet1/1	1	eth	access	down	Link not connected	auto	1
Ethernet1/2	1	eth	access	down	Link not connected	auto	--
Ethernet1/3	19	eth	access	up	none	10000	--
Ethernet1/4	20	eth	access	down	Administratively down	auto	--
Ethernet1/5	--	eth	routed	down	Administratively down	auto	--
Ethernet1/6	1	eth	access	down	Link not connected	auto	--
Ethernet1/7	1	eth	access	down	Link not connected	auto	--
Ethernet1/8	1	eth	access	down	Link not connected	auto	--
Ethernet1/9	1	eth	access	down	Link not connected	auto	--
Ethernet1/10	--	eth	routed	down	Administratively down	auto	--
...							

To view the basic interface capabilities, use the following command:

```
G8272(config)# display interface capabilities
```

```
Ethernet1/1
Model: G8272-X10G-SUP
Type (SFP capable): SFP+ 10G BASE
MDIX: no
Speed: 1000/10000
Duplex: full
Trunk encap. type: 802.1Q
Port-aggregation: yes
Flowcontrol: rx-(off/on),tx-(off/on)
QoS scheduling: rx-(8q1t),tx-(8q1t)
CoS rewrite: no
ToS rewrite: no
SPAN: yes
Port mode: Routed,Bridged
...
```

By default, the switch ethernet interfaces are configured as Layer 2 ports. To configure an ethernet interface as a Layer 3 port (routed port), see [Chapter 15, “Routed Ports”](#).

Interface Description

To add a short description to an interface, use the following command:

```
G8272(config-if)# description <text string>
```

Note: Interface descriptions can have up to a maximum of 80 characters.

To view the description of an interface, use the following command:

```
G8272(config)# display interface description
```

To reset the interface description to its default value, use the following command:

```
G8272(config-if)# no description
```

Interface Duplex

A duplex system is a point-to-point system consisting of two network devices that can communicate between them in both directions (transmission and reception).

There are three configurable duplex modes:

- full duplex - both devices can communicate with each other simultaneously;
- half duplex - both device can communicate with each other, but not simultaneously;
- auto-negotiation - the duplex mode is automatically decided by the device.

To configure the duplex mode of an interface, use the following command:

```
G8272(config-if)# duplex {auto|full|half}
```

To view the current duplex for an interface, use the following command:

```
G8272(config)# display interface status
```

Port	Name	Status	Vlan	Duplex	Speed	Type
Ethernet1/1	--	up	1	full	10000	eth
Ethernet1/2	--	up	trunk	full	10000	eth
Ethernet1/3	--	up	1	full	10000	eth
Ethernet1/4	--	up	1	full	10000	eth
Ethernet1/5	--	up	1	full	10000	eth
Ethernet1/6	--	up	1	full	10000	eth
...						

By default, the duplex mode is auto-negotiation.

To reset the duplex mode configuration to its default, use the following command:

```
G8272(config-if)# no duplex
```

Interface MAC Address

For all switch interfaces, you can configure a different MAC address than the one provided by the manufacturer (burn-in MAC address).

To change the MAC address, use the following command:

```
G8272(config-if)# mac-address <MAC address>
```

To view the current MAC address of an interface, use the following command:

```
G8272(config)# display interface mac-address
```

```
-----  
Interface                Mac-Address      Burn-in Mac-Address  
-----  
Ethernet1/1              a897.dcde.2503  a897.dcde.2503  
Ethernet1/2              a897.dcde.2504  a897.dcde.2504  
Ethernet1/3              a897.dcde.2505  a897.dcde.2505  
Ethernet1/4              a897.dcde.2506  a897.dcde.2506  
Ethernet1/5              a897.dcde.2507  a897.dcde.2507  
...  
Ethernet1/54             a897.dcde.2538  a897.dcde.2538  
mgmt0                    a897.dcde.2500  a897.dcde.2500  
po1                      0e00.0000.0001  (not set)  
Vlan1                    a897.dcde.2501  (not set)
```

Note: This command is available for every type of switch interface.

To reset the MAC address of a switch interface to its default value, use the following command:

```
G8272(config-if)# no mac-address
```

Interface Maximum Transmission Unit

The Maximum Transmission Unit (MTU) is the size (in bytes) of the largest data unit that a protocol can send across the interface.

For each network layer, there is a different size limitation:

- Layer 2 packets: 64 - 9216 bytes;
- Layer 3 IPv4 packets: 576 - 9216 bytes;
- Layer 3 IPv6 packets: 1280 - 9216 bytes.

To configure the MTU of an interface, use the following command:

```
G8272(config-if)# mtu <64-9216 (bytes)>
```

By default, the MTU is 1500 bytes.

To reset the MTU of an interface to its default value, use the following command:

```
G8272(config-if)# no mtu
```

Interface Shutdown

Each individual interface can be administratively brought up or down.

To put an interface in the administratively up state, use the following command:

```
G8272(config-if)# no shutdown
```

To put an interface in the administratively down state, use the following command:

```
G8272(config-if)# shutdown
```

To view the link state of an interface, use the following command:

```
G8272(config)# display interface status
```

Port	Name	Status	Vlan	Duplex	Speed	Type
Ethernet1/1	--	up	1	full	10000	eth
Ethernet1/2	--	up	trunk	full	10000	eth
Ethernet1/3	--	up	1	full	10000	eth
Ethernet1/4	--	up	1	full	10000	eth
Ethernet1/5	--	up	1	full	10000	eth
Ethernet1/6	--	up	1	full	10000	eth
...						

Interface Speed

For physical ethernet interfaces and the management interface, the port speed can be configured to different values than the default setting. Port speed refers to the maximum amount of data (in bits) that can pass across an interface at any given second.

The G8272 offers four possibilities:

- 1000 bits/s (1 Gb/s);
- 10000 bits/s (10 Gb/s);
- 40000 bits/s (40 Gb/s);
- auto-negotiation - the speed is automatically decided by the device.

To configure the port speed of a physical ethernet interface or the management interface, use the following command:

```
G8272(config-if)# speed {1000|10000|40000|auto}
```

By default, the port speed of an interface is auto-negotiation.

To reset the port speed to its default value, use the following command:

```
G8272(config-if)# no speed
```

Flow Control

Flow Control manages the rate of data transmission between two devices, when the transmitter send traffic faster then the receiver can process. It allows the receiver to control the transmission speed and prevents traffic loss.

Flow Control can be enabled for both egress and ingress traffic.

To enable or disable flow control on an interface, use the following command:

```
G8272(config-if)# flowcontrol {send|receive} {on|off}
```

To view the current flow control configuration, use the following command:

```
G8272(config)# display interface flowcontrol
```

Port	Send FlowControl admin	oper	Receive FlowControl admin	oper	RxPause	TxPause
Ethernet1/1	on	on	off	off	0	0
Ethernet1/2	off	off	off	off	0	0
Ethernet1/3	off	off	off	off	0	0
Ethernet1/4	off	off	off	off	0	0
Ethernet1/5	off	off	off	off	0	0
Ethernet1/6	off	off	off	off	0	0
Ethernet1/7	off	off	off	off	0	0
...						

By default, flow control is disabled on all interfaces.

To reset flow control to its default setting, use the following command:

```
G8272(config-if)# no flowcontrol
```

Storm Control

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

Storm Control allow the monitoring of incoming traffic levels for a set time interval. During that interval, traffic level (a percentage of the total interface bandwidth) is compared to the configured traffic storm control level. If ingress traffic reaches the storm control level, the switch drops all packets until the time interval ends.

The switch can be configured to limit the available bandwidth for broadcast, multicast and unknown unicast packets. Unicast packets whose destination MAC address is not in the Forwarding Database (FDB) are considered unknown unicast packets. When an unknown unicast packet 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.

To configure the bandwidth available on an interface for different types of packets, use the following command:

```
G8272(config-if)# storm-control {broadcast|multicast|unicast} level  
<bandwidth percentage>
```

Note: The *bandwidth percentage* value can be defined up to two decimals.

By default, storm control is disabled on all switch interfaces.

To disable storm control on an interface, use the following command:

```
G8272(config-if)# no storm-control {broadcast|multicast|unicast} level
```

Chapter 8. Forwarding Database

The forwarding database (FDB) table is used by a Layer 2 device (switch/bridge) to store the MAC addresses that have been learned and the corresponding ports. The MAC addresses are either dynamically learned from the source addresses of incoming traffic, or manually configured by the system administrator.

When an Ethernet frame is received, the switch dynamically builds the address table by using the source MAC address of the received frames. If the FDB table doesn't contain any information on the destination MAC address of the received frames, it will flood the Ethernet frame out to all ports in the broadcast domain. When the recipient replies, the switch adds its relevant source MAC address and port ID to the FDB table.

The following topics are discussed in this chapter:

- [“MAC Learning” on page 154](#)
- [“Static MAC addresses” on page 155](#)
- [“Aging Time” on page 156](#)

MAC Learning

MAC addresses can be learned only when both global and port based learning control are enabled. The feature is enabled by default.

Interface based MAC learning can only be enabled or disabled on Ethernet ports or Link Aggregation Groups (LAGs). All the Ethernet interfaces belonging to a LAG (static and dynamic) need to have the same MAC learning configuration.

To configure MAC learning control, use the following command:

```
G8272(config)# [no] mac-learn disable  
or  
G8272(config-if)# [no] mac-learn disable
```

To display current MAC learning status, use the following command:

```
G8272(config)# display mac address-table learning [interface ethernet  
<chassis number/port number>]
```

where *chassis number* is 1 and *port number* a value from 1 to 128.

Static MAC addresses

To avoid flooding all LAN ports, entries can be manually configured in the MAC address table. These static MAC entries are retained across a reboot of the switch.

A static MAC address cannot be configured with a non-existing VLAN.

To set a static MAC address, use the following command:

```
G8272(config)# mac address-table static [address <MAC address>] [interface {ethernet <chassis number/port number>|port-aggregation <LAG number>}] [vlan <VLAN number>]
```

where *chassis number* is 1, the *port number* is from 1 to 128, *LAG number* is from 1 to 4096 and *VLAN number* is a value from 1 to 4094.

In addition, you can enter a multicast address as a statically configured MAC address. A multicast address can accept more than one interface as its destination.

To delete a static entry from the MAC address table, use the following command:

```
G8272(config)# no mac address-table static address <MAC address> vlan <VLAN number>
```

To clear the static MAC addresses, use the following command:

```
G8272# remove mac address-table static [address <MAC address>] [interface {ethernet <chassis number/port number>|port-aggregation <LAG number>}] [vlan <VLAN number>]
```

Note: Static MAC address will override any matching dynamic entry in the database.

To display static and multicast MAC address configuration, use one of the following commands:

```
G8272# display mac address-table multicast [count] [vlan <VLAN number>] [user]
```

```
G8272# display mac address-table [dynamic|static] [address <MAC address>] [interface {ethernet <chassis number/port number>|port-aggregation <LAG number>}] [vlan <VLAN number>]
```

Aging Time

You can configure the amount of time that an entry remains in the MAC table.

To create or delete an aging time, use the following command:

```
G8272(config)# mac address-table aging-time <time value>  
or  
G8272(config)# no mac address-table aging-time
```

where *time value* is a number from 0 to 1000000 and 0 means disabled. The default value is 1800 seconds.

To display the global MAC address aging time, use the following command:

```
G8272(config)# display mac address-table aging-time
```

Note: For more information on the FDB available commands, see the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

Chapter 9. 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:

- [“VLAN Overview” on page 158](#)
- [“VLAN Configuration” on page 159](#)
- [“Native VLAN Tagging” on page 170](#)
- [“IPMC Flooding” on page 173](#)
- [“VLAN Topologies and Design Considerations” on page 174](#)

VLAN Overview

A Local Area Network (LAN) is a group of inter-connected computers and other end devices that share the same limited geographical area, such as a school, residence or office building. End devices in a LAN are physically separated from other end devices in different LANs.

A Virtual Local Area Network (VLAN) is a group of end devices that is logically separated by function or application and not by their physical locations. A VLAN has the same attributes as a physical LAN, but you can group together end devices that are part of different LANs.

Setting up VLANs is a way of segmenting 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 unicast, multicast and broadcast frames are forwarded and flooded only to ports belonging to the same VLAN.

VLAN Configuration

The G8272 supports up to 4094 VLANs per switch. Each can be identified with any number between 1 and 4094.

VLANs are divided into four ranges as described below:

Table 7. *VLAN range distribution*

VLAN number	Usage
1	This is the default VLAN. It is created automatically and all switch ports are part of it by default. You cannot modify or delete this VLAN.
2 - 3999	You can create, modify and delete any VLAN included in this range.
4000 - 4094	These 95 VLANs are allocated for internal switch use. You cannot create, modify or delete any VLAN in this range.
4095	This is reserved and unused in accordance with the IEEE 802.1Q standard.

Use the following command to view VLAN information:

```
G8272> display vlan

Flags:
u - untagged egress traffic for this VLAN
t - tagged egress traffic for this VLAN

VLAN    Name                               Status  IPMC FLOOD Ports
=====
1       default                            ACTIVE  Disabled Ethernet1/1(u)
2       VLAN0002                           ACTIVE  Disabled Ethernet1/1(t)
                                     Ethernet1/2(u)
3       VLAN0003                           ACTIVE  Disabled Ethernet1/1(t)
                                     Ethernet1/3(u)
4       VLAN0004                           ACTIVE  Disabled Ethernet1/1(t)
                                     Ethernet1/4(u)
5       VLAN0005                           ACTIVE  Disabled Ethernet1/1(t)
                                     Ethernet1/5(u)
6       VLAN0006                           ACTIVE  Disabled Ethernet1/1(t)
                                     Ethernet1/6(u)
...

```

Creating a VLAN

You can create VLANs falling in the normal VLAN range (VLANs 2-3999).

You can create a single VLAN or a range of VLANs using a single command. The VLAN range can be either continuous (e.g. 20-25) or discrete (e.g. 20, 22, 25, 30).

The default VLAN (VLAN 1) is created automatically when the switch boots up for the first time and persists across later reloads.

You cannot create an internally reserved VLAN (VLANs 4000-4094).

When you create a new VLAN, it will remain unused until you assign a port to it. By default, all switch ports belong to the default VLAN (VLAN 1).

To create a VLAN, use the following command:

```
G8272(config)# vlan <VLAN number>
G8272(config-vlan)#
```

Note: After running the above command, you will enter VLAN Configuration mode for the specified VLAN or range of VLANs.

For example:

```
G8272(config)# vlan 100-105
G8272(config-vlan)#
```

To view the current list of VLAN, use the following command:

```
G8272> display vlan

Flags:
u - untagged egress traffic for this VLAN
t - tagged egress traffic for this VLAN

VLAN    Name                               Status  IPMC FLOOD  Ports
=====
1       default                            ACTIVE  Disabled
                                             Ethernet1/1(u)
                                             ...
100     VLAN0100                           ACTIVE  Disabled
                                             Ethernet1/1(t)
                                             Ethernet1/2(u)
101     VLAN0101                           ACTIVE  Disabled
                                             Ethernet1/3(t)
                                             Ethernet1/4(u)
102     VLAN0102                           ACTIVE  Disabled
                                             Ethernet1/5(t)
                                             Ethernet1/6(u)
...

```


Deleting a VLAN

You can delete VLANs falling in the normal VLAN range (VLANs 2-3999).

You can delete a single VLAN or a range of VLANs using a single command. The VLAN range can be either continuous (e.g. 20-25) or discrete (e.g. 20, 22, 25, 30).

You cannot delete the default VLAN or an internally reserved VLAN (VLANs 4000-4094).

To delete a VLAN, use the following command:

```
G8272(config)# no vlan <VLAN number>
```

For example:

```
G8272(config)# no vlan 100-105
```

You can delete a VLAN which is already associated to a port. If the VLAN is the access VLAN of a switch access port, the VLAN remains configured as the access VLAN, but the switch access port won't be able to forward any traffic until the VLAN is created again.

If the VLAN is the native VLAN of a switch trunk port, the VLAN remains configured as the native VLAN, but the switch trunk port won't be able to forward any untagged traffic until the VLAN is created again.

If the VLAN belongs to the allowed VLAN list of a switch trunk port, the VLAN remains in the allowed VLAN list, but the switch trunk port won't be able to forward any traffic carried in the VLAN until the VLAN is created again.

Configuring the State of a VLAN

You can change the operational state of a VLAN belonging to the normal VLAN range (VLANs 2-3999) to be active or suspended. You can change the state of a single VLAN or a range of VLANs using a single command.

If a VLAN is in active state, it will forward traffic through the ports that are part of that VLAN. If the VLAN is in suspended state, it will not pass any traffic.

By default, when a VLAN is created, it will be in the active operational state.

You cannot change the state of the default VLAN. It will always remain in the active operational state.

You also cannot change the operational state of an internally reserved VLAN (VLANs 4000-4094).

To change the operational state of a VLAN, use the following commands:

- to change the state to active:

```
G8272(config)# vlan <VLAN number>
G8272(config-vlan)# state active
```

- to change the state to suspended:

```
G8272(config)# vlan <VLAN number>
G8272(config-vlan)# state suspend
```

You can reset the VLAN operational state to its default value by using the following command:

```
G8272(config)# vlan <VLAN number>
G8272(config-vlan)# no state
```

To view the current state of each VLAN, use the following command:

```
G8272> display vlan

Flags:
u - untagged egress traffic for this VLAN
t - tagged egress traffic for this VLAN

VLAN   Name                               Status  IPMC FLOOD Ports
=====
1      default                            ACTIVE  Disabled
                                             Ethernet1/1(u)
                                             ...
100    VLAN0100                           ACTIVE  Disabled
                                             Ethernet1/1(t)
                                             Ethernet1/2(u)
...
```

You can change the operation state of a VLAN to suspended even if that VLAN is already associated to a switch port. If the VLAN is the access VLAN of a switch access port, the VLAN remains configured as the access VLAN, but the switch access port won't be able to forward any traffic until the VLAN's state is changed to active.

If the VLAN is the native VLAN of a switch trunk port, the VLAN remains configured as the native VLAN, but the switch trunk port won't be able to forward any traffic carried in the VLAN until the VLAN's state is changed to active.

If the VLAN belongs to the allowed VLAN list of a switch trunk port, the VLAN remains in the allowed VLAN list, but the switch trunk port won't be able to forward any traffic carried in the VLAN until the VLAN's state is changed to active.

Configuring the Name of a VLAN

You can change the name of a VLAN belonging to the normal VLAN range (VLANs 2-3999) to any alphanumeric string with a maximum length of 32 characters.

The name of a VLAN must be unique to that VLAN. No two VLANs can have the same identical name. The names are case-sensitive. For example, the name "VLAN-test" is considered different from "VLAN-TEST".

The name of the default VLAN (VLAN 1) is "default". For all other VLANs the default name used when creating that VLAN is in the format "VLANxxxx", where xxxx represent the four digits of the VLAN number (including leading zeros). For example, the default name of VLAN 2 is VLAN0002.

You cannot change the name of the default VLAN or of an internally reserved VLAN (VLANs 4000-4094).

To change the name of a VLAN, use the following command:

```
G8272(config)# vlan <VLAN number>
G8272(config-vlan)# name <VLAN name>
```

To reset the name of a VLAN to its default value, use the following command:

```
G8272(config)# vlan <VLAN number>
G8272(config-vlan)# no name
```

To view the current names of each VLAN, use the following command:

```
G8272> display vlan

Flags:
u - untagged egress traffic for this VLAN
t - tagged egress traffic for this VLAN

VLAN    Name                               Status  IPMC FLOOD Ports
=====
1       default                            ACTIVE  Disabled
                                             Ethernet1/1(u)
                                             ...
100     VLAN0100                           ACTIVE  Disabled
                                             Ethernet1/1(t)
                                             Ethernet1/2(u)
...
```

Configuring a Switch Access Port

You can configure an ethernet physical interface or Link Aggregation Group (LAG) interface as a switch access port. By default, all the switch ports are in access mode.

A port in access mode can be part only of a single VLAN and it will carry traffic just for that one VLAN. That VLAN is called an Access VLAN.

To configure a port as an access port, use the following command (for this example ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# bridge-port mode access
```

To reset a port to its default operational mode, use the following command:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# no bridge-port mode
```

By default, the access VLAN for an interface is the default VLAN (VLAN 1). You can configure the access VLAN to be the default VLAN or a VLAN belonging to the normal VLAN range (VLANs 2-3999).

Configuring the Access VLAN

You can configure the access VLAN to be an uncreated VLAN. Until that VLAN is created, the switch access port won't be able to forward any traffic. As soon as it is created, the associated port will start forwarding traffic.

To change the access VLAN of a port, use the following command (for this example ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# bridge-port access vlan <VLAN number (1-3999)>
```

To reset a port's access VLAN to the default access VLAN (VLAN 1), use the following command:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# no bridge-port access vlan
```

Note: Any access mode configuration settings are reset when the port's operational mode is changed from access to another mode.

A switch port in access operational mode accepts untagged, priority-tagged and tagged traffic carried in the access VLAN and discards all tagged ingress traffic carried in other VLANs. Egress traffic out of the switch access port should be sent untagged. For more details, see [“Native VLAN Tagging” on page 170](#).

To view the current operational mode of a port, use the following command (for this example ethernet interface 1/12 is used):

```
G8272> display interface ethernet 1/12 bridge-port
```

```
Interface Ethernet1/12  
  Bridge-port           : enabled  
  Bridge-port mode     : access  
  Ingress filter        : enable  
  Acceptable frame types : all  
  Default Vlan          : 110  
  Configured Vlans      : 110  
  Enabled Vlans         : 110
```

Configuring a Switch Trunk Port

You can configure an ethernet physical interface or Link Aggregation Group (LAG) interface as a switch trunk port. By default, all the switch ports are in access mode.

A port in trunk mode can be part of any number of VLANs and it will carry traffic for all configured VLANs.

To configure a port as a trunk port, use the following command (for this example ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# bridge-port mode trunk
```

To reset a port to its default operational mode, use the following command:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# no bridge-port mode
```

Configuring the Allowed VLAN List

The allowed VLAN list is the group of VLANs assigned to a trunk port for which traffic will be forwarded across the trunk link. By default, the allowed VLAN list consists of all VLANs (1-4094). The trunk port will forward traffic only for VLANs that are created on the switch. If a VLAN is in the allowed VLAN list, but it isn't created, the trunk port will not forward traffic for that VLAN. If a VLAN that is in the allowed VLAN list is created after the port is configured as a trunk port, traffic for that VLAN will start to be forwarded as soon the VLAN is created.

For example, the currently created VLANs are the default VLAN and VLANs 100-105. When configuring ethernet interface 1/12 as a trunk port, by default its allowed VLAN list will include all VLANs, but will only forward traffic for VLANs 1 and 100-105. If VLAN 106 is created on the switch, ethernet interface 1/12 will start forwarding traffic for VLAN 106 as well.

You can configure an uncreated VLAN to a switch trunk port as a member of the allowed VLAN list. The switch trunk port won't be able to forward traffic carried in that VLAN until it is created.

To change the allowed VLAN list of a port, use the following command (for this example ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# bridge-port trunk allowed vlan <VLAN number (1-3999)>
```

For example:

```
G8272(config-if)# bridge-port trunk allowed vlan 100-105,300,320
```

You can also add new VLANs to the current allowed VLAN list, remove VLANs from the list or specify the list to contain all VLANs or none.

Considering the example above, the allowed VLAN list currently contains VLANs 100-105, 300 and 320. The following command will add VLAN 110 to that list:

```
G8272(config-if)# bridge-port trunk allowed vlan add 110
```

If you want to remove VLAN 110 from the list, use the following command:

```
G8272(config-if)# bridge-port trunk allowed vlan remove 110
```

If you want to add all VLANs to the list, use the following command:

```
G8272(config-if)# bridge-port trunk allowed vlan all
```

If you want to add all VLANs to the list except VLAN 110, use the following command:

```
G8272(config-if)# bridge-port trunk allowed vlan except 110
```

If you don't want any allowed VLANs on the current port, use the following command:

```
G8272(config-if)# bridge-port trunk allowed vlan none
```

To reset a port's allowed VLAN list to its default settings, use the following command:

```
G8272(config)# interface ethernet 1/12  
G8272(config-if)# no bridge-port trunk allowed vlan
```

Configuring the Native VLAN

When a port is configured as a switch trunk port, it will tag traffic packets with their appropriate VLAN number. All packets coming from the allowed VLAN list will be tagged when they're carried across the trunk, except for packets belonging to the native VLAN. Packets coming from the native VLAN are carried across the trunk untagged.

By default, the native VLAN for an interface is the default VLAN (VLAN 1). You can configure the native VLAN to be the default VLAN or a VLAN belonging to the normal VLAN range (VLANs 2-3999).

You can configure the native VLAN to be an uncreated VLAN. The switch trunk port won't be able to forward untagged traffic, priority tagged traffic or traffic carried in that VLAN until the native VLAN is created and added to its allowed VLAN list.

To change the native VLAN of a port, use the following command (for this example ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12  
G8272(config-if)# bridge-port trunk native vlan <VLAN number (1-3999)>
```


To reset a port's native VLAN to the default native VLAN (VLAN 1), use the following command:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# no bridge-port trunk native vlan
```

Note: Any trunk mode configuration settings are reset when the port's operational mode is changed from trunk to another mode.

A port configured in trunk operational mode accepts tagged traffic carried in allowed VLANs and discard traffic carried in VLANs excluded from the allowed VLAN list on ingress side. Egress traffic carried in any allowed VLAN except the native VLAN should be sent tagged on egress side.

The trunk port accepts untagged, priority-tagged and tagged traffic carried in the native VLAN on the ingress side if native VLAN tagging is disabled on the port. Egress traffic carried in the native VLAN out of the trunk port is sent untagged on egress side if native VLAN tagging is disabled on the trunk port. Otherwise, the traffic should be sent tagged with the native VLAN ID. For mode details, see [“Native VLAN Tagging” on page 170](#).

To view the current operational mode of a port, use the following command (for this example ethernet interface 1/12 is used):

```
G8272> display interface ethernet 1/12 bridge-port

Interface Ethernet1/12
  Bridge-port           : enabled
  Bridge-port mode     : trunk
  Ingress filter        : enable
  Acceptable frame types : all
  Default/Native Vlan  : 110
  Configured Vlans     : 100-105,300,320
  Enabled Vlans        : 100-105,300,320
```

Native VLAN Tagging

When a packet enters a switch trunk link, a tag is added to its frame header to represent the VLAN membership of the frame's port. Each frame must be distinguishable as being within exactly one VLAN. A frame entering the trunk link that does not contain a VLAN tag is assumed to be flowing on the native VLAN.

This can lead to a security vulnerability in the network. It is advised to enable the tagging of native VLAN traffic. This ensures that all packets going out of the trunk link will be tagged and it prevents the reception of untagged packets.

If native VLAN tagging is disabled, the switch is vulnerable to VLAN hopping attacks. The basic concept behind a VLAN hopping attack is for an attacking host on a VLAN to gain access to traffic on other VLANs that would normally not be accessible.

To address such potential attacks, native VLAN tagging performs the following:

- on the ingress side, all untagged data traffic is dropped when native VLAN tagging is not explicitly configured for egress traffic;
- on the egress side, all traffic is tagged. If traffic belongs to the native VLAN then it is tagged with the native VLAN ID.

Native VLAN tagging can be configured globally or on each interface. The switch port must be in trunk operational mode for native VLAN tagging to function.

By default, native VLAN tagging is globally disabled on the switch. Globally enabling or disabling native VLAN tagging affects all trunk ports that don't have native VLAN tagging individually configured, which is the default setting for 802.1Q trunk ports. Until a native VLAN tagging is configured on each trunk port, individual tagging behavior is dependent on the global configuration.

When native VLAN tagging is disabled on a switch trunk port, all untagged ingress packets are accepted and carried in the native VLAN. All tagged packets belonging to VLANs in the allowed VLAN list are accepted as long as that VLAN is active and tagged packets belonging to VLANs not in the allowed VLAN list are discarded. If the ingress packet is priority-tagged, it is regarded as an untagged packet and it is carried in the native VLAN.

If an egress packet belongs to the native VLAN then it is sent untagged. Other tagged packets belonging to VLANs in the allowed VLAN list are sent with the VLAN preserved. If the egress packet is priority-tagged and belongs to the native VLAN, it is sent untagged (the priority tag is removed). Otherwise, the priority-tagged packet is tagged with a new VLAN ID and its priority tag is removed in advance. In other words, the packet is not sent in Q-in-Q format.

When native VLAN tagging is enabled on a switch trunk port, all untagged ingress packets will be discarded unless native VLAN tagging is not explicitly configured for egress traffic. Otherwise, all untagged ingress packets are accepted in the trunk link and carried in the native VLAN. If the ingress packet is priority-tagged, it is accepted in the native VLAN only if tagging is explicitly configured for egress traffic. Otherwise, the packet is discarded.

All egress packets are tagged. If a packet belongs to the native VLAN then it is tagged with the native VLAN ID. If the egress packet is priority-tagged, it is tagged with a new VLAN ID and its priority tag is removed in advance. In other words, the packet is not sent in Q-in-Q format.

Note: Control protocol data units (PDU) are accepted as untagged on the native VLAN of a switch trunk port, even if native VLAN tagging is enabled.

Configuring Native VLAN Tagging

Native VLAN tagging is supported both on physical ethernet interfaces and on Link Aggregation Groups (LAG). A LAG inherits the native VLAN tagging configuration of its first trunk port member. When adding a switch trunk port to a LAG, the port must have the same native VLAN tagging configuration as the LAG. Otherwise, the trunk port will be rejected from the LAG.

To globally enable or disable native VLAN tagging, use the following command:

```
G8272(config)# [no] vlan dot1q tag native
```

If you want to enable native VLAN tagging just for egress traffic, use the following command:

```
G8272(config)# vlan dot1q tag native egress-only
```

Note: If native VLAN tagging is enabled only for egress traffic, all untagged ingress packets are accepted in the switch trunk port and are carried in the native VLAN. Otherwise, the ingress packets are discarded.

Native VLAN tagging can be configured independently on each ethernet interface or LAG. To enable or disable tagging on an interface, use the following command (for this example, ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# vlan dot1q tag native enable
```

To enable native VLAN tagging only for egress traffic, use the following command:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# vlan dot1q tag native enable egress-only
```

To disable native VLAN tagging on an interface, use the following command:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# vlan dot1q tag native disable
```

To delete native VLAN tagging configuration of an interface, use the following command:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# no vlan dot1q tag native
```

Note: This will make native VLAN tagging on that interface dependent on the global configuration.

To view the current native VLAN tagging settings for all interfaces, use the following command:

```
G8272> display vlan dot1q tag native
```

```
Tag native vlan global setting: enabled
```

```
-----  
Port Tag Native VLAN Status
```

```
-----  
Eth1 enabled
```

```
Eth5 disabled
```

```
Eth13 enabled (egress-only)
```

```
...
```

IPMC Flooding

IP Multicast (IPMC) is a method of sending IP traffic to a group of interested devices in a single transmission. When sending a multicast packet, a device uses an IP multicast address from the range:

- for IPv4: 224.0.0.0/4 (224.0.0.0 - 239.255.255.255)
- for IPv6: ff00::/8

If the switch does not understand multicast addresses then it will flood the unknown traffic to all the members of a VLAN. In this scenario the switch has to filter the packets sent to multicast groups they are not subscribed to.

By default, IPMC flooding is enabled on all VLANs, meaning that unknown IP multicast traffic is forwarded across all VLAN member interfaces. If IPMC flooding is disabled, the switch will discard all unknown multicast IP traffic that has a destination IP address outside the reserved multicast range. For IPv4 addressing that range is 224.0.0.X (224.0.0.0 - 224.0.0.255).

To enable or disable the flooding of unknown IPMC traffic in a VLAN, use the following command:

```
G8272(config-vlan)# [no] flood
```

To enable or disable IPMC flooding for a specific family of IP addresses, use the following command:

- for IPv4:

```
G8272(config-vlan)# [no] flood ipv4
```

- for IPv6:

```
G8272(config-vlan)# [no] flood ipv6
```

To check the IPMC flood configuration of a VLAN, use the following command:

```
G8272> display vlan id <1-4096>

Flags:
u - untagged egress traffic for this VLAN
t - tagged egress traffic for this VLAN

VLAN    Name                Status  IPMC FLOOD Ports
=====
1       default             ACTIVE  Disabled
                                             Ethernet1/3(u)
                                             Ethernet1/4(u)
                                             Ethernet1/5(u)
...

```

VLAN Topologies and Design Considerations

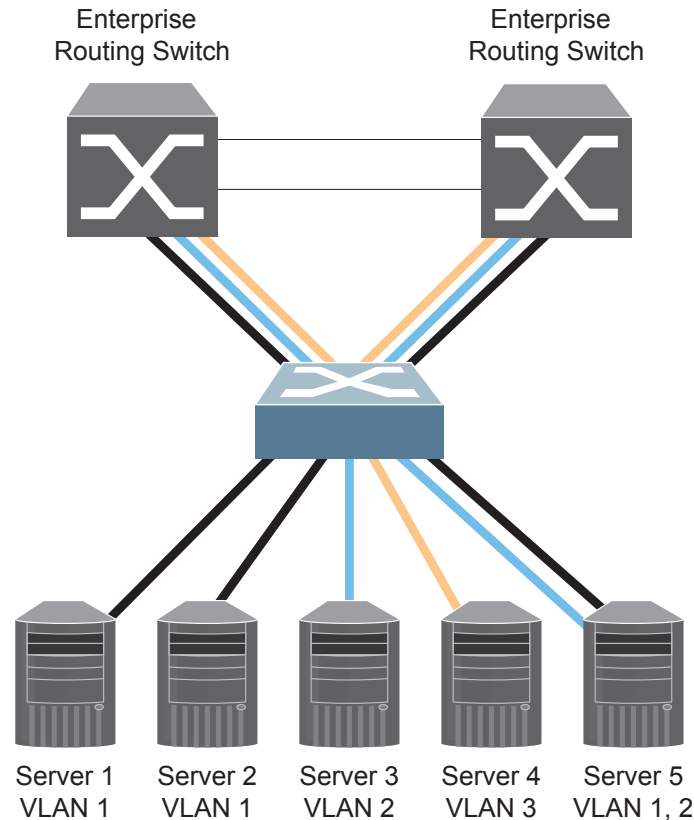
Note the following when working with VLAN topologies:

- By default, the G8272 software is configured so that trunk operational mode and native VLAN tagging are disabled on all switch ports.
- By default, the G8272 software is configured so that all switch ports are members of the default VLAN (VLAN 1).
- When using the Multiple Spanning Tree Protocol (MSTP), Spanning Tree Groups (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 10, “Ports and Link Aggregation”](#) and [Chapter 26, “Port Mirroring”](#).

Multiple VLANs with Trunk Mode Adapters

Figure 1 illustrates a network topology described in [Note: on page 176](#) and the configuration example on [page 177](#).

Figure 1. Multiple VLANs with VLAN-Tagged Gigabit Adapters



The features of this VLAN are described in the following table.

Table 8. Multiple VLANs Example

Component	Description
G8272 switch	This switch is configured with three VLANs that represent three different IP subnets. Five ports (ethernet ports 1/11-1/15) are connected downstream to servers. Two ports (ethernet ports 1/19 and 1/20) are connected upstream to routing switches. Uplink ports are members of all three VLANs and are in trunk mode.
Server 1	This server is a member of VLAN 1 and has presence in only one IP subnet. The associated switch port (ethernet port 1/11) is only a member of VLAN 1, so the port is in access mode.
Server 2	This server is a member of VLAN 1 and has presence in only one IP subnet. The associated switch port (ethernet port 1/12) is only a member of VLAN 1, so the port is in access mode.

Table 8. Multiple VLANs Example (continued)

Component	Description
Server 3	This server belongs to VLAN 2 and it is logically in the same IP subnet as Server 5. The associated switch port (ethernet port 1/13) is in access mode.
Server 4	A member of VLAN 3, this server can communicate only with other servers via a router. The associated switch port (ethernet port 1/14) is in access mode.
Server 5	A member of VLAN 1 and VLAN 2, this server can communicate with Server 1, Server 2, Server 3 and the upstream switches. It cannot communicate with Server 4. The associated switch port (ethernet port 1/15) is in trunk mode.
Enterprise Routing Switches	These switches must have all three VLANs (VLAN 1, 2 and 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. The switch ports are in trunk mode.

Note: Switch trunk port 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 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 port'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.

VLAN Configuration Example

Use the following procedure to configure the example network shown in [Figure 1 on page 175](#).

1. Enable trunk mode on server ports that support multiple VLANs and add VLANs to the trunk link.

```
G8272(config)# interface ethernet 1/15
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlans 1,2
G8272(config-if)# exit
```

2. Enable trunk mode on uplink ports that support multiple VLANs.

```
G8272(config)# interface ethernet 1/19
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# exit

G8272(config)# interface ethernet 1/20
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# exit
```

3. Configure server ports that belong to a single VLAN.

```
G8272(config)# interface ethernet 1/13
G8272(config-if)# bridge-port access vlan 2
G8272(config-if)# exit

G8272(config)# interface ethernet 1/14
G8272(config-if)# bridge-port access vlan 3
G8272(config-if)# exit
```

By default, all ports are members of the default VLAN (VLAN 1), so configure only those ports that belong to other VLANs. In this example, ethernet ports 1/11 to 1/12 are by default switch access ports and are members of VLAN 1, and don't require any further configuration.

Chapter 10. Ports and Link Aggregation

Link Aggregation (LAG) groups can provide super-bandwidth, multi-link connections between the G8272 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:

- [“G8272 Port Configuration Profiles” on page 180](#)
- [“Aggregation Overview” on page 183](#)
- [“Static LAGs” on page 184](#)
- [“Link Aggregation Control Protocol” on page 188](#)
- [“LAG Hashing” on page 193](#)

G8272 Port Configuration Profiles

The G8272 is a 1U rack-mountable aggregation switch. It uses a wire-speed, non-blocking switching fabric that provides simultaneous wire-speed transport of multiple packets at low latency on all ports.

The G8272 has the following port capabilities:

- Forty-eight 10 Gigabit Ethernet (GbE) Small Form Pluggable Plus (SFP+) ports which also support legacy 1 GbE connections;
- Six 40 GbE Quad Small Form Pluggable Plus (QSFP+) ports, each of which can optionally be used as four 10 GbE SFP+ ports;

SFP+ ports, when used in 10 GbE mode, can be populated with optical transceiver modules or active or passive DACs. When used in legacy 1 GbE mode, the ports can be populated with optical or copper transceiver modules.

QSFP+ ports can be populated with optical QSFP+ modules or passive Direct-Attach Cables (DACs), including those that allow the port to be used as four 10 GbE SFP+ ports.

By default, the G8272 has the following interface (port) configuration:

- ethernet interfaces 1-48 are configured in 10Gbit/s port mode;
- ethernet interfaces 49-54 are configured in 40Gbit/s port mode;

The G8272 supports three different port profiles:

- default - 48 10Gbit/s SFP+ ports and 6 40Gbit/s QSFP+ ports;
- 72x10G - 72 10Gbit/s SFP+ ports;
- custom port mode;

In 72x10G port mode, each of the six 40Gbit/s ports is split into four 10Gbit/s breakout ports, adding up to a total of 72 10Gbit/s ports.

To change the port mode to the 72x10G profile, use the following command:

```
G8272(config)# hardware profile portmode 72x10G
```

You can configure a custom port profile on the switch, changing an ethernet interface port mode from 40Gbit/s to 10Gbit/s or vice versa, and have different numbers of 10Gbit/s and 40Gbit/s ports than the default configuration. For example, you can configure 64 10Gbit/s ports and 2 40Gbit/s ports.

Note: The maximum number of 40Gbit/s ports configured on a switch is 6.

To create a custom port profile, use the following command:

```
G8272(config)# hardware profile portmode custom ethernet <chassis number/port number>
```

Note: This command will configure the specified port to run in 10Gbit/s mode.

You can specify a single ethernet port or a range of ports. The range can be either continuous (e.g. 20-25) or discrete (e.g. 20, 22, 25, 30).

For example:

1. Configure a single ethernet interface to run in 10Gbit/s mode:

```
G8272(config)# hardware profile portmode custom ethernet 1/50
```

2. Configure a range of ethernet interfaces to run in 10Gbit/s mode:

- a. a continuous range:

```
G8272(config)# hardware profile portmode custom ethernet 1/50-54
```

- b. a discrete range:

```
G8272(config)# hardware profile portmode custom ethernet 1/49, ethernet 1/53, ethernet 1/54
```

Note: When a 40Gbit/s port is split into four 10Gbit/s breakout ports, the ethernet interface associated with that port will include four subports.

For example, before changing ethernet port 54 to 10Gbit/s port mode, the port will be displayed as below:

```
G8272> display interface ethernet 1/54 brief
```

Ethernet Interface	VLAN	Type	Mode	Status	Reason	Speed	Port Agg#
Ethernet1/54	1	eth	access	down	Link not connected	auto	--

After splitting the port into four 10Gbit/s breakout ports, the port will be displayed as below:

```
G8272> display interface ethernet 1/54/1-4 brief
```

Ethernet Interface	VLAN	Type	Mode	Status	Reason	Speed	Port Agg#
Ethernet1/54/1	1	eth	access	down	Link not connected	auto	--
Ethernet1/54/2	1	eth	access	down	Link not connected	auto	--
Ethernet1/54/3	1	eth	access	down	Link not connected	auto	--
Ethernet1/54/4	1	eth	access	down	Link not connected	auto	--

To configure 10Gbit/s breakout ports, use the following command:

```
G8272(config)# interface ethernet <chassis number/port number/subport number>
```

For example:

```
G8272(config)# interface ethernet 1/54/2
```

When configuring a custom port profile on an ethernet interface, you can reset the interface's port mode to its default settings by using the following command:

```
G8272(config)# no hardware profile portmode custom ethernet <chassis  
number/port number/subport number>
```

To reset the switch port configuration to the default profile, use the following command:

```
G8272(config)# hardware profile portmode default
```

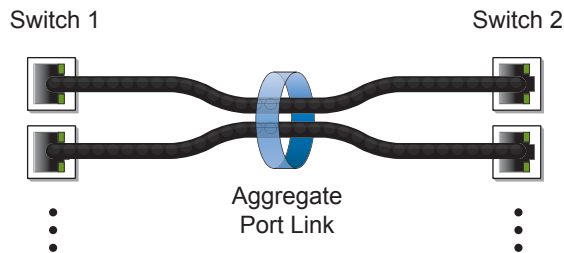
To view the current configured port profile, use the following command:

```
G8272# display boot portmode  
  
Current port mode:  
  Port Ethernet1/49 is set in 10G mode  
  Port Ethernet1/50 is set in 10G mode  
  Port Ethernet1/51 is set in 10G mode  
  Port Ethernet1/52 is set in 10G mode  
  Port Ethernet1/53 is set in 10G mode  
  Port Ethernet1/54 is set in 10G mode
```

Aggregation Overview

When using LAGs between two switches, as shown in [Figure 2](#), you can create a virtual link between the switches, operating with combined throughput levels that depends on how many physical ports are included. The physical links in a LAG operate as one single logical connection.

Figure 2. Port LAG



LAGs are also useful for connecting a G8272 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.

A LAG can be statically configured or dynamically aggregated through the use of a network protocol, like the Link Aggregation Control Protocol (LACP). For more details about LACP, see [“Link Aggregation Control Protocol” on page 188](#).

The switch supports up to a maximum of 72 LAGs, either statically configured or aggregated dynamically using LACP. When creating a LAG, you must assign it a number from 1 to 4096.

Creating a LAG

To create a LAG, use the following command:

```
G8272(config)# interface port-aggregation <1-4096>
G8272(config-if)#
```

Note: Upon creating a LAG, you will enter Interface Configuration command mode, where you'll be able to configure the settings of the LAG.

To remove a LAG, use the following command:

```
G8272(config)# no interface port-aggregation <1-4096>
```

Static LAGs

You can create a static LAG that contains multiple physical links. Static LAGs don't require protocols like LACP to run on the switch to keep the configured aggregation.

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.

Before configuring a static LAG, consider the following:

- Consider how existing VLANs will be affected by the addition of a LAG.
- Determine which switch ports (up to 32) are to become LAG members (the specific ports making up the LAG).
- Ensure that the chosen switch ports are in the up link state.
- Member ports must be all configured either as switch access ports or as switch trunk ports and have the same VLAN membership. For more details, see [“Configuring a Switch Access Port” on page 165](#) and [“Configuring a Switch Trunk Port” on page 167](#).
- LAG member ports must have the same Spanning Tree Protocol (STP) configuration.
- LAG member ports must have MAC address learning and flow control enabled.
- The same maximum transmission unit (MTU) size must be configured on all LAG member ports.
- The same bandwidth, speed, duplex and auto-negotiation settings must be configured on all LAG member ports.
- Consider how the existing Spanning Tree will react to the new LAG configuration.

Static LAG Configuration Rules

When configuring a static LAG, you should consider the following configuration rules that determine how a LAG will react in any network topology:

- All links must originate from one logical source 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, such as when using VLAGs (see [Chapter 12, “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.
- A physical switch port can't be member of multiple LAGs. It can belong to only one single LAG.
- Aggregation from third-party devices must comply with Cisco EtherChannel technology.
- Each LAG inherits its port configuration (speed, flow control, native VLAN tagging) from the first member port. As additional ports are added to the LAG, their settings must be changed to match the LAG configuration.

- The maximum number of physical links in a LAG is 32. After reaching this limit, no more links can be added to the LAG.
- When a port becomes a member of a LAG, its MAC address and STP configuration are replaced with the ones of the LAG.
- When a port joins or leaves a LAG, the following parameters are unaffected: interface description, LLDP settings, LACP port priority and link state.
- 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

To add a physical port to a static LAG, use the following steps:

1. Configure the ethernet interface's port mode and VLAN membership.
 - a. Configure the interface as a switch access port, or:

```
G8272(config)# interface ethernet <chassis number/port number>
G8272(config-if)# bridge-port mode access
G8272(config-if)# bridge-port access vlan <1-3999>
G8272(config-if)# exit
G8272(config)#
```

- b. As a switch trunk port:

```
G8272(config)# interface ethernet <chassis number/port number>
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan <1-3999>
G8272(config-if)# exit
G8272(config)#
```

2. Create a LAG and configure the same port mode and VLAN membership as the previously configured ethernet interface.
 - a. Access port mode, or:

```
G8272(config)# interface port-aggregation <1-4096>
G8272(config-if)# bridge-port mode access
G8272(config-if)# bridge-port access vlan <1-3999>
G8272(config-if)# exit
G8272(config)#
```

- b. Trunk port mode:

```
G8272(config)# interface port-aggregation <1-4096>
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan <1-3999>
G8272(config-if)# exit
G8272(config)#
```

3. Add the interface to a LAG without an aggregation protocol:

```
G8272(config)# interface ethernet <chassis number/port number>
G8272(config-if)# aggregation-group <1-4096> mode on
G8272(config-if)# exit
G8272(config)#
```

4. Check the current LAG configuration:

```
G8272(config)# display interface port-aggregation <1-4096>

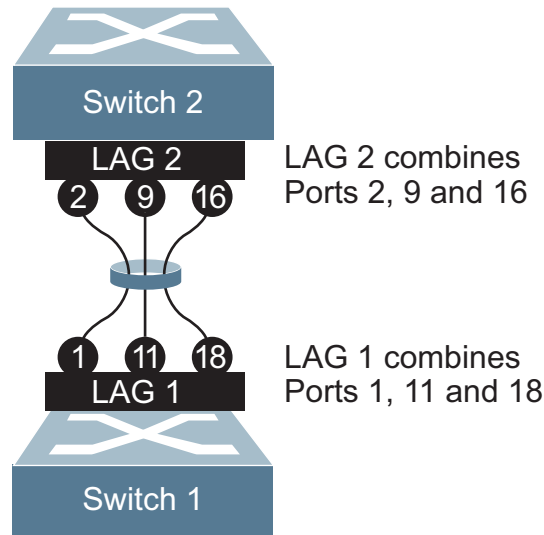
Interface po1
  Hardware is AGGREGATE Current HW addr: a897.dcde.2503
  Physical:(not set) Logical:(not set)
  index 100001 metric 1 MTU 1500 Bandwidth 0 Kbit
  Port Mode is access
  <UP,BROADCAST,MULTICAST>
  VRF Binding: Not bound
  Members in this port-aggregation:
    Ethernet1/1, Ethernet1/2, Ethernet1/3, Ethernet1/4, Ethernet1/5
  ...
```

To remove an interface from a LAG, use the following command:

```
G8272(config)# interface ethernet <chassis number/port number>
G8272(config-if)# no aggregation-group
```

In the following example, three ports are aggregated between two switches.

Figure 3. 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 *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

1. On Switch 1:
 - a. Add ethernet interfaces 1, 11 and 18 to LAG 1:

```
G8272(config)# interface ethernet 1/1, ethernet 1/11, ethernet 1/18
G8272(config-if-range)# aggregation-group 1 mode on
G8272(config-if-range)# exit
```

Note: LAG 1 is automatically created when you assign ethernet interfaces 1, 11 and 18 to it.

- b. Check the configured LAG settings:

```
G8272(config)# display interface port-aggregation 1
```

2. On Switch 2:
 - a. Add ethernet interfaces 2, 9 and 16 to LAG 2:

```
G8272(config)# interface ethernet 1/2, ethernet 1/9, ethernet 1/16
G8272(config-if-range)# aggregation-group 2 mode on
G8272(config-if-range)# exit
```

Note: LAG 2 is automatically created when you assign ethernet interfaces 2, 9 and 16 to it.

- b. Check the configured LAG settings:

```
G8272(config)# display interface port-aggregation 2
```

3. Connect the switch ports that will be members in the LAG.

LAG 1 on Switch 1 is now connected to LAG 2 on Switch 2.

Note: In this example, two G8272 switches are used. If a third-party device supporting link aggregation is used (such as Cisco routers and switches with EtherChannel 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.

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 - LAG) 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 a LACP LAG fails, traffic is reassigned dynamically to the remaining link(s) of the dynamic LAG.

LACP ensures that a local LAG doesn't attempt to send traffic to a remote single interface. With LAG enabled on the switch, a local LAG can only transmit packets to a remote LAG also configured with LACP.

LACP allows the exchange of aggregation information between the two LAG participants in the form of Link Aggregation Control Protocol Data Units (LACPDU). If the information isn't agreed upon, the LAG will not form.

In comparison with a static LAG, LACP has the following advantages:

- If a physical link fails, LACP will detect the error and remove the link from the aggregate, ensuring packets are not lost due to the failure.
- Both of the participating devices can mutually confirm the LAG configuration. In the case of static link aggregation, configuration errors are often not detected as quickly.

Up to a maximum of 32 physical ports can be configured on a single LACP LAG.

By default, LACP is globally enabled on the switch and it can't be disabled. However, it can be enabled or disabled for individual interfaces.

Configuring LACP

Before configuring ethernet ports as members of a LACP LAG, you must ensure they meet the same configuration requirements (switch port mode, VLAN membership etc.) as specified for static LAGs. For more details, see [page 184](#).

To include a physical port in a LACP LAG, you must add the port to the desired LAG and specify its channel mode. You can configure a port with the following channel modes:

- on - configures the port as part of a static LAG without an aggregation protocol;
- active - enables LACP on the interface and configures the port as an active member of the LACP negotiation that starts the negotiations by sending LACP packets to other ports;
- passive - configures the port as a passive member of the LACP negotiation that only responds to received LACP packets from ports in the active state.

Note: LACP is enabled on ports configured in the passive state only when an active LACP device is detected at the other end of the physical link.

A port in LACP active state will be able to form a LAG with another port that is either in the active or passive state. A port in LACP passive state will form a LAG with another port, only if that port is in the active state.

Note: Two ports in LACP passive state will not form a LAG, because neither will start the negotiation process.

An interface in the on state doesn't participate in LACP traffic. Thus, when a LACP active port tries to negotiate with that interface, it will not receive any reply and will become an individual link with the interface.

To configure an interface as a LACP LAG member, use the following command to add that interface to the desired LAG and specify its LACP state:

```
G8272(config)# interface ethernet <chassis number/port number>
G8272(config-if)# aggregation-group <1-4096> mode {active|passive}
```

To remove an interface from a LACP LAG, use the following command:

```
G8272(config)# interface ethernet <chassis number/port number>
G8272(config-if)# no aggregation-group
```

System Priority

A switch running LACP has an assigned system priority. LACP uses the configured system priority together with the switch's MAC address to generate a LACP system ID.

The system priority is also used during negotiations with other devices. The switch with the lowest configured system priority determines which physical ports participate in the LACP aggregation. If both switches have the same system priority, then the LACP system ID is used to determine the controlling device.

To configure the switch LACP system priority, use the following command:

```
G8272(config)# lacp system-priority <1-65535>
```

By default, the system priority has a value of 32768. To reset the switch's system priority to the default value, use the following command:

```
G8272(config)# no lacp system-priority
```

Port Priority

Each physical interface that is member of a LACP LAG has an assigned port priority. LACP uses the configured port priority together with the port number to form the LAC port ID.

When there is a restriction that doesn't allow all compatible interfaces to aggregate, LACP uses port priority to decide what interfaces should be put in hot-standby mode and what interfaces should be put in active mode. A port with a lower configured priority will be more likely to be chosen as an active link. If two interfaces have the same port priority, the port ID is used to determine their roles.

To configure the port priority of an interface, use the following command:

```
G8272(config)# interface ethernet <chassis number/port number>  
G8272(config-if)# lACP port-priority <1-65535>
```

By default, the port priority of an interface has a value of 32768. To reset the port priority to the default value, use the following command:

```
G8272(config-if)# no lACP port-priority
```

LACP Timeout

An interface that is a member of a LACP LAG will send LACPDUs at a regular interval. After not receiving any response to three consecutive LACPDUs, the link is timed out and the state of the interface transitions to individual. Interfaces in the individual or suspended state will function as normal physical ports.

You can configure a short or a long timeout interval. A short timeout interval is equal to three seconds, meaning that LACPDUs are sent every second. A long timeout interval has a value of 90 seconds, with LACPDUs being sent every 30 seconds.

To configure the timeout interval of an interface, use the following command:

```
G8272(config)# interface ethernet <chassis number/port number>  
G8272(config-if)# lACP timeout {short|long}
```

By default, LACP uses a long timeout interval. To reset an interface LACP timeout interval to this value, use the following command:

```
G8272(config-if)# no lACP timeout
```

LACP Individual

A port with LACP enabled (active or passive) that can't join a LACP LAG has its state changed to individual and is able to form individual links. This can be useful when a LACP enabled port is connected to port with LACP disabled or when two LACP enabled ports are both in the passive state. In such scenarios, LACP individual helps avoid the loss of traffic.

By default, LACP individual is globally enabled on the switch and can't be disabled. However, it can be enabled or disabled individually on each interface.

Note: LACP individual can only be enabled or disabled for LACP enabled ports. Static LAGs member are unaffected.

Although, this behavior is helpful in preventing traffic loss, it can cause problems by permitting loops to be created by mismatching port configurations.

Ports with LACP enabled can be configured to transition to suspended state instead of individual. Ports in the suspended state are not able to process traffic.

To configure a LACP enabled interface to transition to suspended state after it times out, use the following command:

```
G8272(config)# interface port-aggregation <1-4096>
G8272(config-if)# shutdown
G8272(config-if)# lacp suspend-individual
G8272(config-if)# no shutdown
```

Note: The LAG must be disabled (shut down) to configure LACP individual. Once the desired changes are made, the LAG can be re-enabled.

To configure a LACP enabled interface to transition to individual state after it times out, use the following command:

```
G8272(config-if)# shutdown
G8272(config-if)# no lacp suspend-individual
G8272(config-if)# no shutdown
```

Note: LACP individual can only be configured for LAGs. When the LACP individual configuration is changed, it will apply the modification to all member ports.

LACP Configuration Example

Use the following procedure to configure LACP for ethernet ports 1, 2 and 3 to participate in link aggregation.

1. Ensure that ethernet ports 1, 2 and 3 have the same port mode settings and are members of the same VLAN (in this example, VLAN 10 is used).
 - a. Configure them as switch access ports, or:

```
G8272(config)# interface ethernet 1/1-3
G8272(config-if-range)# bridge-port mode access
G8272(config-if-range)# bridge-port access vlan 10
G8272(config-if-range)# exit
G8272(config)#
```

- b. As switch trunk ports:

```
G8272(config)# interface ethernet 1/1-3
G8272(config-if-range)# bridge-port mode trunk
G8272(config-if-range)# bridge-port trunk native vlan 10
G8272(config-if-range)# bridge-port trunk allowed vlan 10
G8272(config-if-range)# exit
G8272(config)#
```

2. Create a LAG (in this example, LAG 1 is used) and ensure that its port mode settings and VLAN membership is the same as the ones configured on ethernet ports 1, 2 and 3.

a. Configure the LAG with access port mode, or:

```
G8272(config)# interface port-aggregation 1
G8272(config-if)# bridge-port mode access
G8272(config-if)# bridge-port access vlan 10
G8272(config-if)# exit
G8272(config)#
```

b. With trunk port mode:

```
G8272(config)# interface port-aggregation 1
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk native vlan 10
G8272(config-if)# bridge-port trunk allowed vlan 10
G8272(config-if)# exit
G8272(config)#
```

3. (Optional) Allow member ports to individually participate in normal data traffic if no LACPDU's are received:

```
G8272(config)# interface port-aggregation 1
G8272(config-if)# shutdown
G8272(config-if)# no lACP suspend-individual
G8272(config-if)# no shutdown
G8272(config-if)# exit
G8272(config)#
```

4. (Optional) Configure the port priority and timeout of the LAG members:

```
G8272(config)# interface ethernet 1/1-3
G8272(config-if)# lACP port-priority <1-65535>
G8272(config-if)# lACP timeout {short|long}
G8272(config-if)# exit
G8272(config)#
```

5. (Optional) Configure the LACP system priority:

```
G8272(config)# lACP system-priority <1-65535>
```

6. Configure the ports as members of LAG 1 and set them in LACP active mode:

```
G8272(config)# interface ethernet 1/1-3
G8272(config-if-range)# aggregation-group 1 mode active
```

LAG Hashing

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 physical port the frame will use.

LAG hashing can be only configured globally on the switch, thus affecting all existing LAGs.

Note: LAG hashing is enabled by default on the switch and can't be disabled.

The G8272 uses RTAG7 as its hashing method. It offers different information options from which to choose the parameters used to calculate the hash value. To achieve the most even traffic distribution, configure 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 switch can use the following packet information options to calculate the hash value for traffic balancing:

- Packet ingress physical interface;
- Layer 2 information:
 - destination MAC address;
 - source MAC address;
 - source and destination MAC addresses;
- Layer 3 information:
 - destination IP address;
 - source IP address;
 - source and destination IP addresses;
- Layer 4 information:
 - destination port number;
 - source port number;
 - source and destination port numbers.

Note: All packet information options are mutually exclusive when calculating the hash value. Only one of these options can be selected.

Note: The packet ingress physical interface can also be used in combination with of the above options to calculate the hash value.

When the hash value is calculated using Layer 3 or Layer 4 packet information, the associated Layer 2 is also applied to non-IP protocol traffic. [Table 9](#) presents the applied hashing criteria associated with every hashing configuration:

Table 9. *Hashing Criteria*

Hash configuration	Layer 2 Criteria	Layer 3 Criteria	Layer 3 Criteria
destination mac	destination mac	destination mac	destination mac
source mac	source mac	source mac	source mac
source and destination mac	source and destination mac	source and destination mac	source and destination mac
destination IP	destination mac	destination IP	destination IP
source IP	source mac	source IP	source IP
source and destination IP	source and destination mac	source and destination IP	source and destination IP
destination port	destination mac	destination IP	destination IP, destination port
source port	source mac	source IP	source IP, source port
source and destination port	source and destination mac	source and destination IP	source and destination IP, source and destination port

By default, the hash value is calculated using the source and destination IP addresses. This means, that for Layer 3 or Layer 4 packets, the hash value is calculated using the source and destination IP addresses, and for Layer 2 packets, the hash value is calculated using the source and destination MAC addresses.

To check the current LAG hash configuration, use the following command:

```
G8272> display port-aggregation load-balance

Port-aggregation Load-Balancing Configuration:
System: source-dest-ip

Port-aggregation Load-Balancing Addresses Used Per-Protocol:
Non-IP: source-dest-mac
IP: source-dest-ip
```

LAG Hashing Configuration

To configure choose what packet information options is used to calculate the hash value, use the following command:

```
G8272(config)# port-aggregation load-balance ethernet <hash method>

G8272(config)# port-aggregation load-balance ethernet ?
destination-ip  Load distribution on the destination IP address
destination-mac Load distribution on the destination MAC address
destination-port Load distribution on the destination TCP/UDP port
source-dest-ip  Load distribution on the source and destination IP
                address
source-dest-mac Load distribution on the source and destination MAC
                address
source-dest-port Load distribution on the source and destination
                TCP/UDP port
source-interface Load distribution on the source ethernet interface
source-ip        Load distribution on the source IP address
source-mac       Load distribution on the source MAC address
source-port      Load distribution on the source TCP/UDP port
```

For more details about this command, see the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

To reset the hashing method to its default settings, use the following command:

```
G8272(config)# no port-aggregation load-balance ethernet
```

To view the LAG hash value, use the following command:

```
G8272> display port-aggregation load-balance forwarding-path interface
port-aggregation <1-4096> <hash parameters>

G8272> display port-aggregation load-balance forwarding-path interface
port-aggregation 1 ?
src-interface  Optional source interface (physical switch port only)
dst-mac        Destination MAC Address
src-mac        Source MAC Address
dst-ip         Destination IP Address
src-ip         Source IP Address
dst-ipv6       Destination IPV6 Address
src-ipv6       Source IPV6 Address
l4-dst-port    Destination Port
l4-src-port    Source Port
```

For more details about this command, see the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

The *hash parameters* include the following:

- Packet ingress physical interface;
- source and destination MAC addresses;
- source and destination IPv4 or IPv6 addresses;
- source and destination port numbers;

The *hash parameters* used to calculate the hashing value are decided by the LAG hashing configuration and the traffic type. If there is IP/port information input, then the traffic type is Layer 3/Layer 4. If there is no such information input, then the traffic type is Layer 2. Any unspecified hash parameters are substituted with 0.

For example, if the LAG hashing configuration is source and destination IP addresses, and IP information is contained in the input, then the Layer 2 and Layer 4 information is ignored when calculating the hash value.

```
G8272> display port-aggregation load-balance forwarding-path interface
port-aggregation <1-4096> dst-ip 10.158.70.96 src-ip 10.169.2.36 src-mac
78AA.F560.EF0F 14-src-port 14255 14-dst-port 16874
```

```
Missing params will be substituted by 0's.
Load-balance Algorithm on switch: source-dest-ip
Outgoing port id: Ethernet 1/1
```

```
Param(s) used to calculate load balance:
destination-ip: 10.158.70.96
source-ip: 10.169.2.36
```

If the LAG hashing configuration is source and destination IP addresses, and no IP information is contained in the input, then the default IPv4 address (0.0.0.0) is used when calculating the hash value:

```
G8272> display port-aggregation load-balance forwarding-path interface
port-aggregation <1-4096> src-mac 78AA.F560.EF0F 14-src-port 14255
14-dst-port 16874
```

```
Missing params will be substituted by 0's.
Load-balance Algorithm on switch: source-dest-ip
Outgoing port id: Ethernet 1/1
```

```
Param(s) used to calculate load balance:
destination-ip: 0.0.0.0
source-ip: 0.0.0.0
```

Chapter 11. Multiple Spanning Tree Protocol

When multiple paths exist between two points on a network, Spanning Tree Protocol (STP) or one of its enhanced variants ensure that the RackSwitch G8272 uses only the most efficient network path.

Cloud Network OS 10.1 supports Multiple Spanning Tree Protocol (MSTP).

IEEE 802.1Q (2005) MSTP provides both rapid convergence and load balancing in a VLAN environment. MSTP allows multiple STGs, with multiple VLANs in each.

This chapter covers the following topics:

- [“MSTP Overview” on page 198](#)
- [“Global MSTP Control” on page 199](#)
- [“Bridge Protocol Data Units” on page 200](#)
- [“MST Region” on page 204](#)
- [“Global MST Instance Parameters” on page 205](#)
- [“Port Type and Link Type” on page 207](#)
- [“MSTP Configuration Guidelines” on page 208](#)

MSTP Overview

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 (2005).

In MSTP mode, the G8272 supports up to 64 instances of Spanning Tree, corresponding to STGs 1-64, 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* ports (["Port Type and Link Type" on page 207](#)) bypass the Discarding and Learning states, and enter directly into the Forwarding state.

Global MSTP Control

By default, the Multiple Spanning Tree feature is globally enabled on the switch (MST mode). Multiple Spanning Tree can be globally disabled using the following command:

```
G8272(config)# spanning-tree mode disable
```

MSTP can be re-enabled by specifying the MSTP mode:

```
G8272(config)# spanning-tree mode mst
```

MSTP can also be individually enabled or disabled on each interface. To enable MSTP on an interface, use the following command (for this example, ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# spanning-tree enable
```

By default, MSTP is also enabled on each interface. To disable it, use the following command (for this example, ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# spanning-tree disable
```

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.

Port States

The port state controls the forwarding and learning processes of Spanning Tree. In MSTP, 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 207](#)) 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.

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 G8272 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.

BPDU Guard

The BPDU Guard will put a switch interface into the shut down (error disabled) state when a BPDU is received on that interface. The administrator can manually put the interface in the up state or, if Error Disable Recovery is enabled, it will automatically bring the interface up (for more information, see “[Error Disable Recovery](#)” on page 203).

Note: BPDU Guard will function even when MSTP is disabled. However, it will not work when BPDU Filter is also enabled.

By default, BPDU Guard is disabled. You can enable it on each of the switch interfaces by using the following command:

```
G8272(config)# interface ethernet <chassis number/port number>  
G8272(config-if)# spanning-tree bpduguard enable
```

To disable BPDU Guard on an interface, use the following command:

```
G8272(config-if)# spanning-tree bpduguard disable
```

To reset the BPDU Guard to its default configuration, use the following command:

```
G8272(config-if)# no spanning-tree bpduguard
```


BPDU Filter

The BPDU Filter prevents the transmission or reception of BPDUs on a switch interface.

Note: When the BPDU Filter is enabled on an interface, the BPDU Guard will not function, because all BPDUs will be dropped by the switch.

Note: BPDU Filter will function even when MSTP is disabled.

By default, BPDU Filter is disabled. You can enable it on each of the switch interfaces by using the following command:

```
G8272(config)# interface ethernet <chassis number/port number>  
G8272(config-if)# spanning-tree bpdufilter enable
```

To disable BPDU Filter on an interface, use the following command:

```
G8272(config-if)# spanning-tree bpdufilter disable
```

To reset the BPDU Filter to its default configuration, use the following command:

```
G8272(config-if)# no spanning-tree bpdufilter
```

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 `auto` (the default) indicates that the default cost will be computed for an auto-negotiated link or LAG speed.

Use the following command to globally configure the port path cost:

```
G8272(config)# spanning-tree mst 0 cost {<1-200,000,000>|auto}
```

You can also configure the port path cost differently on each interface by using the following command:

```
G8272(config)# interface ethernet <chassis number/port number>  
G8272(config-if)# spanning-tree cost {<1-200,000,000>|auto}
```

You can choose the calculation method for the default path cost by using the following command:

```
G8272(config)# spanning-tree pathcost method <method type>
```

where *method type* is one of the following:

- *long*: 32 bit based values;
- *short*: 16 bit based values.

However, when using MST spanning tree mode, the switch uses only the long method for calculating path cost.

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:

```
G8272(config)# spanning-tree mst <0-64> priority <0-61440>
```

Note: Priority must be configured in increments of 4096 (such as 0, 4096, 8192, 12288 and so on).

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:

```
G8272(config)# interface ethernet <chassis number/port number>  
G8272(config-if)# spanning-tree port-priority <0-224>
```

Note: Port priority must be configured in increments of 32 (such as 0, 64, 96, 128 and so on). If the specified port priority value is not evenly divisible by 32, the value will be automatically rejected.

The default port priority is 128. To reset it to its default value, use the following command:

```
G8272(config-if)# no spanning-tree port-priority
```

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 and then automatically recovered.

Note: This differs from BPDU Guard, where the port is blocked and recovered only if Error Disabled Recovery is also enabled, or the port is manually brought back up by an administrator.

You can configure the root guard at the port level using the following commands:

```
G8272(config)# interface ethernet <chassis number/port number>  
G8272(config-if)# spanning-tree guard root
```

The default state is disabled.

Error Disable Recovery

When BPDU Guard is enabled on a switch interface, it will put that interface in the error disabled (shutdown) state when a BPDU is received.

To bring the interface back up, the administrator can manually put it in the up state or he can configure Error Disable Recovery to automatically transition interface from the error disabled state to the up state after a certain time interval.

To enable or disable Error Disable Recovery, use the following command:

```
G8272(config)# [no] errdisable recovery cause bpduguard
```

By default, Error Disable Recovery will bring the port back up after 30 seconds since it was detected as being in the error disabled state.

To configure the recovery time interval (in seconds), use the following command:

```
G8272(config)# errdisable recovery interval <30-65535>
```

To reset the recovery time interval to its default value, use the following command:

```
G8272(config)# no errdisable recovery interval
```

MST 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:

- Region name (up to a maximum of 32 alphanumeric characters)
- 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.

To enter the MST Region Configuration mode, use the following command:

```
G8272(config)# spanning-tree mst configuration
G8272(config-mst)#
```

To assign a name to the MST region, use the following command:

```
G8272(config-mst)# name <region name>
```

To assign a revision number to the MST region, use the following command:

```
G8272(config-mst)# revision <0-65535>
```

By default, the MST region revision number is 0.

The administrator may manually assign VLANs to specific MSTIs (Multiple Spanning Tree instances).

1. If no VLANs exist (other than default VLAN 1), see [“Creating a VLAN” on page 160](#) for information about creating VLANs and assigning ports to them.
2. Assign the VLAN to a MSTI using the following command:

```
G8272(config)# spanning-tree mst configuration
G8272(config-mst)# instance <0-64> vlan <VLAN number>
```

Global MST Instance Parameters

You can configure the following global parameters for MST instances:

- hop count
- forward delay
- hello timer
- maximum age interval

Hop Count

To compute the STP topology inside the MST region, the protocol uses, along with the path cost to the root, a hop-count system similar to the IP time-to-live (TTL) mechanism.

Use the following command to configure the maximum number of hops inside the region:

```
G8272(config)# spanning-tree mst max-hops <1-255>
```

By default, the maximum number of hops is 20. To reset it to its default value, use the following command:

```
G8272(config)# no spanning-tree mst max-hops
```

Forward Delay

Forward delay specifies the amount of time that a bridge port has to wait before it changes from the discarding and learning states to the forwarding state.

Use the following command to configure the forward delay (in seconds):

```
G8272(config)# spanning-tree mst forward-time <4-30>
```

By default, the forward delay is 15 seconds. To reset it to its default value, use the following command:

```
G8272(config)# no spanning-tree mst forward-time
```

Hello Timer

The Hello timer specifies how often the bridge transmits a configuration BPDU. Any bridge that is not the root bridge uses the root bridge hello timer value.

Use the following command to configure the hello timer (in seconds):

```
G8272(config)# spanning-tree mst hello-time <1-10>
```

By default, the forward delay is 2 seconds. To reset it to its default value, use the following command:

```
G8272(config)# no spanning-tree mst hello-time
```

Maximum Age Interval

The maximum age interval specifies the maximum time the bridge waits without receiving a configuration BPDU before it reconfigures the MSTP network.

Use the following command to configure the maximum age interval (in seconds):

```
G8272(config)# spanning-tree mst max-age <6-40>
```

By default, the forward delay is 20 seconds. To reset it to its default value, use the following command:

```
G8272(config)# no spanning-tree mst max-age
```

Port Type and Link Type

Edge 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 connected 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 configure a port as an edge port:

```
G8272(config)# interface ethernet <chassis number/port number>
G8272(config-if)# [no] spanning-tree port type edge
```

Link Type

The link type determines how the port behaves in regard to Multiple Spanning Tree. Use the following commands to define the link type for the port:

```
G8272(config)# interface ethernet <chassis number/port number>
G8272(config-if)# [no] spanning-tree link-type <type>
```

where *type* corresponds to the duplex mode of the port, as follows:

- `point-to-point` 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.

MSTP Configuration Guidelines

This section provides important information about configuring Multiple Spanning Tree Groups:

- You may configure up to 64 MSTP instances. Up to 4k VLANs are supported in MSTP mode.
- Switches must have the same MST configuration identification elements (name, revision level number and VLAN to MSTI mapping) to be in the same MST region.
- MSTP is fully compatible with RSTP. A MST region appears as a virtual bridge connecting to single spanning tree entities.
- When you enable MSTP, a default revision number of 0 and a blank region name are automatically configured.
- The switch supports a single instance of the MSTP Algorithm consisting of the CIST and up to 64 MSTIs.
- A VLAN can only be mapped to one MSTI or to the CIST. One VLAN mapped to multiple spanning trees is not allowed. All the VLANs are mapped to the CIST by default. Once a VLAN is mapped to a specified MSTI, it is removed from the CIST.

MSTP Configuration Example

This section provides steps to configure MSTP on the switch.

1. Configure port and VLAN membership on the switch.
2. Configure Multiple Spanning Tree region parameters.

```
G8272(config)# spanning-tree mst configuration
G8272(config-mst)# name Lenovo
G8272(config-mst)# revision 5
G8272(config-mst)# instance 1 vlan 10
G8272(config-mst)# exit
```

3. Configure bridge priority for instance number 1.

```
G8272(config)# spanning-tree mst 1 priority 4096
```

4. Configure timers.

```
G8272(config)# spanning-tree mst hello-time 5
G8272(config)# spanning-tree mst max-age 10
G8272(config)# spanning-tree mst forward-time 25
```

5. Configure edge ports.

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# spanning-tree port type edge
```


6. Configure BPDU guard.

```
G8272(config-if)# spanning-tree bpduguard enable
G8272(config-if)# exit
```

7. Enable the timeout mechanism for a port to be recovered automatically from being operationally shut down by BPDU.

```
G8272(config)# errdisable recovery cause bpduguard
G8272(config)# errdisable recovery interval 30
```

8. Configure Root Guard.

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# spanning-tree guard root
```

9. Enable BPDU filtering.

```
G8272(config-if)# spanning-tree bpdufilter enable
```

10. Configure link type.

```
G8272(config-if)# spanning-tree link-type point-to-point
```

11. Configure path cost value and port priority.

```
G8272(config-if)# spanning-tree cost 10000
G8272(config-if)# spanning-tree port-priority 64
```

Note: To configure these parameters per instance, use the following commands:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# spanning-tree mst 1-10 port-priority 64
G8272(config-if)# spanning-tree mst 1-10 cost 10000
```

12. Configure max hops value.

```
G8272(config)# spanning-tree mst max-hops 20
```

13. Configure path cost method.

```
G8272(config)# spanning-tree pathcost method long
```

14. Verify MSTP configuration.

```
G8272# display spanning-tree mst
```

Chapter 12. Virtual Link Aggregation Groups

Virtual Link Aggregation Groups (vLAGs) is a mechanism which allows the redundant uplinks to remain active by utilizing all available bandwidth.

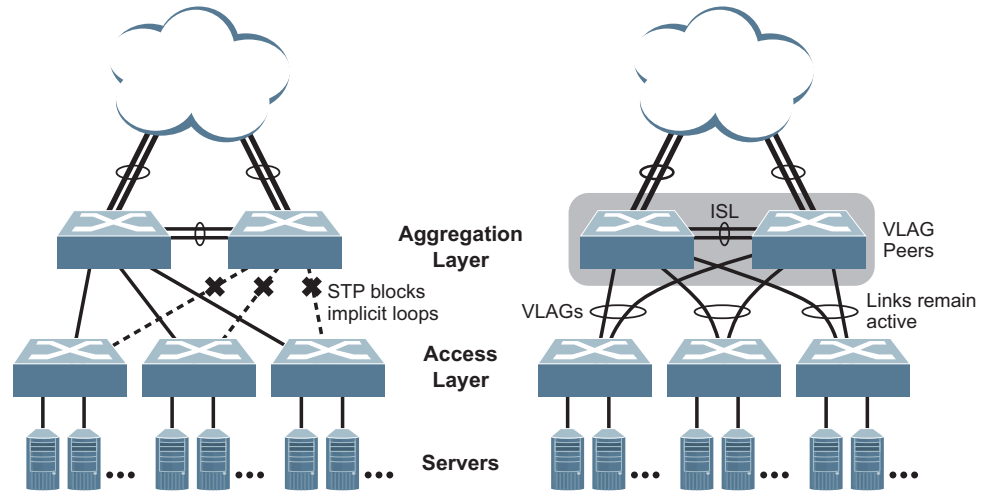
This section discusses the following topics:

- [“vLAG Overview” on page 212](#)
- [“vLAG Capacities” on page 214](#)
- [“vLAGs versus regular LAGs” on page 219](#)
- [“Configuring vLAGs” on page 220](#)
- [“Health Check” on page 225](#)
- [“Basic vLAG Configuration Example” on page 227](#)
- [“vLAG Configuration - VLANs Mapped to a MST Instance” on page 230](#)
- [“Configuring vLAGs in Multiple Layers” on page 232](#)

vLAG Overview

In many data center environments, downstream servers or switches connect to upstream devices which consolidate traffic. For example, see [Figure 4](#).

Figure 4. 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 (MSTP—see [Chapter 11, “Multiple Spanning Tree Protocol”](#)) 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.

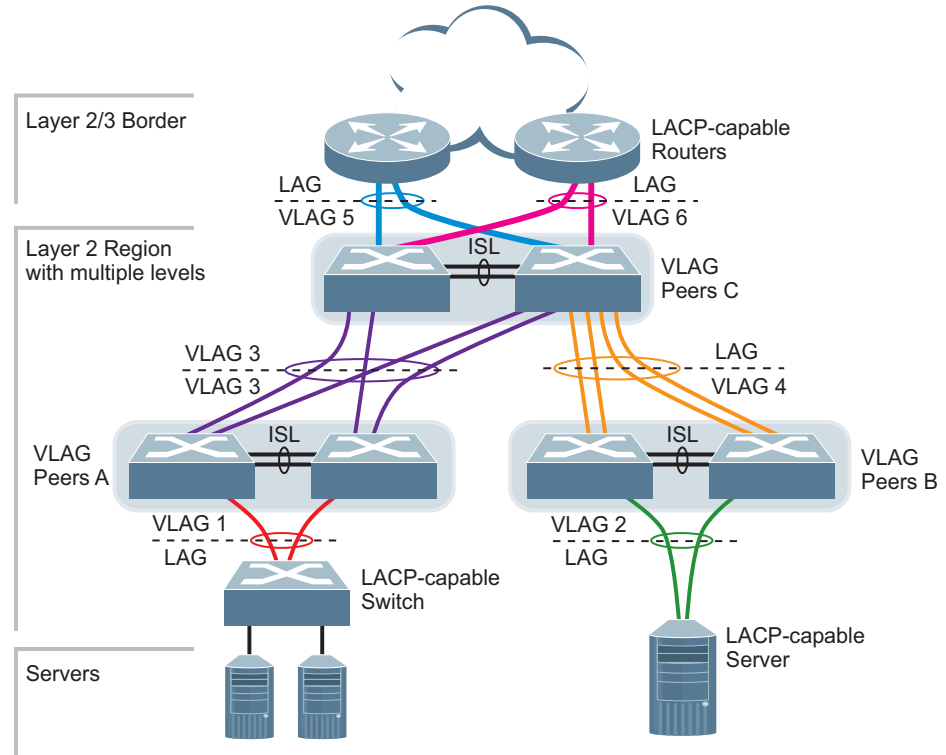
Two switches are paired into vLAG peers and act as a single virtual entity for the purpose of establishing a multi-port aggregation. Ports from both peers can be grouped into a vLAG and connected to the same LAG-capable target device. From the perspective of the target device, the ports connected to 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 also support Layer 2 MultiPathing, which offers redundancy and traffic load balancing across multiple paths between network nodes.

A vLAG can be set up as either a static LAG or a LACP LAG.

vLAGs are useful in multi-layer environments for both uplink and downlink redundancy to any regular LAG-capable device. For example:

Figure 5. 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 LACP 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.

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 LAGs on the vLAG client end.

On the vLAG peers, the vLAGs are configured similarly to regular LAGs, using many of the same features and rules. See [Chapter 10, “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 or a LACP LAG, and consumes one slot from the overall 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 72 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 64.

vLAG Benefits

Using vLAGs offers the following benefits:

- enables a single device to use a LAG across two upstream switches;
- STP blocked interfaces are eliminated;
- topology loops are also eliminated;
- enables the use of all available uplink bandwidth;
- allows fast convergence times in case of link or device failure;
- allows link-level resilience;
- enables high availability.

vLAG Synchronization Mechanism

All vLAG status and protocol synchronization information is transmitted between vLAG peers using the inter-switch link (ISL) LAG. vLAG uses a Layer 2 transport protocol (Edge Control Protocol or ECP) for communicating across the ISL LAG.

If the transport channel used by ECP for communication stops working, the ISL is marked as down. If vLAG health check detects that the primary switch is still functioning, the ports of the secondary switch that are members of the vLAG instance are shut down (error disabled) to avoid creating loops.

To view information about the ECP transport channel, use the following command:

```
G8272# display vlag ecp channels
```

The state of the transport channel used by ECP is detected periodically by vLAG hello packets. Both vLAG peers send hello packets every 5 seconds. If no hello packets are received for 30 seconds, the channel is marked as down.

ECP is a transport protocol that operates between two peers over an IEEE 802 LAN. ECP provides reliable, in-order delivery of ULP (Upper Layer Protocol) PDUs (Protocol Data Units).

To view information about the upper layer protocols associated with ECP, use the following command:

```
G8272# display vlag ecp upper-layer-protocols
```

ECP has the following service characteristics:

- Reliable delivery of ULPDUs, that is resilient against frame loss;
- Delivery of ULPDUs to the recipient ULP in the order that they were transmitted by the sending ULP;
- Delivery of a single copy of each ULPDU to the recipient;
- Flow control that provides protection against buffer overrun on the receive side.

vLAG System MAC

vLAG supports both static LAGs and LACP LAGs. For connections between an access switch and an end host, vLAGs that are configured using LACP LAGs will use a special reserved MAC address called vLAG system MAC.

The vLAG system MAC address is created using the vLAG tier ID. Together with the LACP system priority, the MAC address is used to generate a LACP system ID, which is used during LACP negotiation. For more details about LACP, see [“Link Aggregation Control Protocol” on page 188](#).

vLAG and LACP Individual

By default, physical interfaces that are members of a LACP LAG and don't receive any Link Aggregation Control Protocol Data Units (LACPDU) after a set amount of time will be expelled from the LAG and enter individual state. Interfaces in the individual state will function as normal physical ports. For more details, see [“LACP Individual” on page 190](#).

If the interface is part of a vLAG, it will be marked as a non-vLAG port when it enters individual state. The port will forward and block traffic according to the upper layer protocol, for e.g. STP.

This behavior can lead to the creation of loops when the vLAG doesn't connect the switch to an end host. To avoid such scenarios, only LACP LAGs with LACP individual enabled interfaces that are connected to an end device should be added to the vLAG. When you assign a LACP LAG to a vLAG, the switch will display a warning if it detects that the LAG member ports have LACP individual enabled.

```
warning: port-aggregation 1 with lacp individual enabled is added for
vLAG instance 1
```

Note: The LAG used for the vLAG ISL can't have LACP individual enabled.

Note: When the upper layer protocol (for e.g. STP) is disabled, you must manually ensure that there are no loops.

vLAG and LACP System Priority

To reduce configuration dependency, LACP system priority information is synchronized between the vLAG peers. The secondary switch will use the system priority of the primary switch to reaggregate with the access switch, regardless of its own priority. This ensures that LACP LAG member ports that participate in a vLAG are aggregated in the same LAG on the access switch, even if the LACP system priorities of the primary and secondary switches don't match.

If the secondary switch becomes the primary, it will replace its current LACP system priority with its original value and then reaggregate with the access switch.

vLAG LACP Misconfigurations or Cabling Errors

On a single switch, LACP can ensure that only one LAG will form within an aggregation group, despite incorrect configurations or cabling. In a vLAG topology, LACP misconfiguration or erroneous cabling can cause the formation of two LAGs within an aggregation group, thus creating a network loop.

The switch detects such LACP misconfigurations or cabling errors. After detection, vLAG will suspend the ports that are part of the vLAG on the secondary switch and a syslog message will be generated. Once the misconfiguration or cabling error is corrected, vLAG will include the suspended ports to be part of the LAG.

Any misconfiguration will be checked only on the secondary vLAG switch. The vLAG primary switch will synchronize partner information with the secondary switch, followed by a misconfiguration check on the secondary vLAG switch. When a misconfiguration occurs, the following syslog message will be generated:

```
LACP port <port name>'s received partner info is mismatch with vLAG instance <1-128>'s expected one, it will be suspended.
```

Note: On the primary vLAG switch, no misconfiguration syslog messages will be displayed.

FDB Synchronization

To prevent the flooding of unknown unicast packets and offer local preference, the Forwarding Databases (FDB) are synchronized between the vLAG peers.

MAC addresses learned by a switch on its vLAG ports are synchronized across the corresponding vLAG ports of its peer. MAC addresses not learned on vLAG ports are synchronized over the ISL of the vLAG peers.

A peer MAC address is added as a static unicast FDB entry across the ISL to avoid unnecessary traffic flooding of the vLAG peer.

FDB synchronization will start as soon as the first vLAG instance is formed and will stop when the last vLAG instance will be disbanded.

The member ports of the vLAG ISL will disable MAC learning once FDB synchronization starts and enable it once the FDB synchronization stops. This happens regardless of the configured MAC learning settings of the ports prior to the starting of FDB synchronization.

Synchronized MAC entries will not be aged or cleaned by the local switch. When this happens, the MAC entries will be re-installed in the FDB. To avoid certain scenarios where an aged MAC entry isn't re-installed fast enough, thus triggering a traffic flood, vLAG uses a timer to periodically check the status of all FDB entries and re-install any MAC addresses that are scheduled for cleaning. For more details, see [“FDB Refresh” on page 223](#).

To get FDB learning and aging callbacks, vLAG will use the FDB interrupt mechanism.

vLAG and MSTP

The G8272 supports Multiple Spanning Tree Protocol (MSTP) running over individual vLAG instances. For more details about MSTP, see [Chapter 11, “Multiple Spanning Tree Protocol”](#).

Bridge protocol data units (BPDU) are received by both vLAG peers and are sent only through the link that received the BPDUs from the upstream switches. A BPDU is a data message transmitted across a local area network to detect loops in network topologies. A BPDU contains information regarding ports, switches, port priority and addresses, which is necessary to configure and maintain the spanning tree topology.

Both vLAG peers use the vLAG system MAC address as the designated bridge ID on their vLAG member ports. The root bridge ID remains unchanged.

The ISL LAG forwards traffic even if the root bridge is configured on a switch not participating in the vLAG. In this scenario, it is possible that the switch will display two root ports. MSTP will consider the ports that are members of the ISL LAG as such only if the ISL has formed. If the ISL fails, the ports are considered regular ports and might be blocked.

MSTP will calculate the path cost of non-vLAG ports as normal. For vLAG ports, the protocol assigns a fixed value of 1, regardless of the path cost method used.

For regular LAGs (static or LACP), MSTP uses a dynamic calculation to determine the path cost of the LAG. This method is based on the path cost of the LAG member ports. When a LAG member's state changes to down, the path cost of the LAG is recalculated.

For LAGs participating in a vLAG, MSTP uses a fixed path cost value of 1, instead of the dynamic calculation described above. This means that a vLAG has better path preference than a regular single port. Also, when a vLAG member's state changes to down, the path cost of the LAG is not recalculated.

The path cost of the LAGs involved in the vLAG can be changed to another value than 1, but this must be manually configured on both vLAG peers. The new path cost should be identical on both primary and secondary switches. When this condition is not met, the results of the MSTP path cost calculation might not be correct.

To ensure that vLAG works properly with MSTP, the vLAG peers should be part of the same STP region. When this condition is not met, the vLAG ports of the secondary will be considered as being disabled (not error disabled) in the MSTP configuration. This will negatively affect traffic flow through the vLAG. When the vLAG peers are correctly configured as being part of the same STP region, the vLAG ports of the secondary will transition to the up state.

BPDU guard and BPDU filter will consider the vLAG as a single link. For example, when BPDU guard will receive a BPDU packet, the vLAG ports from both vLAG peers will be put in the error disabled state.

As long as one of the vLAG instances running on the vLAG peers has not failed, a vLAG port is visible in the MSTP configuration, even if the interface is in the administratively down or error disabled state. This means that the status of the vLAG port is not determined of its physical state, but instead by the vLAG. MSTP considers the vLAG port as being a single logical link and is oblivious if one of the vLAG members fails. MSTP will mark the vLAG port as being in the down state only if the vLAG link fails on both vLAG peers.

vLAGs versus regular LAGs

Though similar to regular LAGs in many regards, vLAGs differ from regular 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 regular LAG member ports.
- 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 be configured for all vLAG VLANs;
 - ISL ports must be placed into a regular LAG (static or LACP);
 - A minimum of two ports on each switch are recommended for ISL use.
- Dynamic routing protocols, such as OSPF, cannot terminate on vLAGs.
- 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

It is recommended that any vLAG configuration is done without traffic flowing through the vLAG peers. This can cause FDB flush events and resulting in a delayed stabilization of the vLAG configuration.

Note: The vLAG peers can be reloaded while traffic is flowing.

When configuring vLAG or making changes to your vLAG configuration, consider the following vLAG behavior:

- 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 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

When configuring a vLAG, the participating switches are checked for matching on the following parameters:

- Tier ID - It is used to divide vLAG domains. Different domains can't be aggregated together;
- System type - Both vLAG peers must be the same switch model to reduce the impact of different system performance and capability;
- CNOS version - It is used to avoid vLAG feature compatibility issues.

If the Tier ID or system type don't match, the vLAG won't form. If the peers use different CNOS versions, the vLAG will form, but notification messages will be displayed every 10 seconds, informing you to upgrade to the same CNOS version.

By default, the vLAG feature is disabled.

To enable or disable vLAG, use the following command:

```
G8272(config)# [no] vlag enable
```

vLAG ISL

vLAG uses an Inter-Switch Link (ISL) connection to synchronize information between the two vLAG peers. The ISL is a dedicated regular LAG (static or LACP) that must be formed from at least two physical interfaces.

It is recommended that the total bandwidth of the ISL LAG be larger than the total bandwidth of the vLAG ports. When a vLAG port fails, its traffic will be forwarded through the ISL LAG.

To configure the LAG used by the ISL, use the following command:

```
G8272(config)# vlag isl port-aggregation <1-4096>
```

Note: This command also triggers the forming of the ISL between the vLAG peers.

To delete the vLAG ISL, use the following command:

```
G8272(config)# no vlag isl port-aggregation
```

vLAG Role Election

For a centralized vLAG operation, vLAG assigns a role to each of the two switches participating in the vLAG: primary and secondary. The primary switch controls the centralized operation.

Each vLAG role is determined by comparing the local vLAG priorities and MAC addresses of the participating switches. The switch with the lower priority will become the primary. If the switches are configured with identical priorities, the switch with the smaller MAC address will be elected as the primary.

The vLAG role election is not preemptive. If a switch is already elected as the primary, another switch that joins the vLAG will become the secondary and remain in that role even if its priority or MAC address are smaller than those of the primary switch.

To configure the vLAG priority of a switch, use the following command:

```
G8272(config)# vlag priority <0-65535>
```

By default, the switch has a vLAG priority of 0. To reset the priority to the default value, use the following command:

```
G8272(config)# no vlag priority
```

vLAG Instance

vLAGs are configured using an instance ID. The instance ID is used to identify which LAGs are connected to downstream devices as a vLAG. A single vLAG instance can group LAGs with different IDs on the vLAG peers. For example, vLAG instance 1 can bind together LAG 1 from the primary vLAG peer and LAG 6 from the secondary vLAG peer.

The instance ID must be the same for vLAG peers.

The maximum number of configurable vLAG instances is 64.

When creating a vLAG, you must also assign a LAG to it. To create a vLAG instance, use the following command:

```
G8272(config)# vlag instance <1-64> port-aggregation <1-4096>
```

Note: Only one LAG can be assign to a vLAG on a single switch.

By default, vLAG instances are disabled. To enable or disable the vLAG instance, use the following command:

```
G8272(config)# [no] vlag instance <1-64> enable
```

To remove the associated LAG from the vLAG instance, use the following command:

```
G8272(config)# no vlag instance <1-64> port-aggregation
```

Note: When removing the assigned LAG from an enabled vLAG instance, the instance will also be disabled.

To remove a vLAG instance, use the following command:

```
G8272(config)# no vlag instance <1-64>
```

vLAG formation depends on the state of the two participating LAGs (local and remote). The possible states of the vLAG are listed below:

- DOWN - both the local and remote LAGs are in the down state;
- LOCAL_UP - only the local LAG is in the up state;
- REMOTE_UP - only the remote LAG is in the up state;
- FORMED - both the local and remote LAGs are in the up state.

Note: If the ISL is in the down state, vLAG synchronization isn't possible and the remote vLAG will be treated as being in the down state regardless of its real state.

FDB Refresh

vLAG uses a timer (FDB refresh interval) to periodically check the status of synchronized FDB entries. If a MAC address is detected as waiting to be removed from the FDB, the address will be re-installed in the FDB.

The FDB refresh interval is shorter than MAC aging time, which can be configured by using the following command:

```
G8272(config)# mac address-table aging-time <0-1000000 seconds>
```

By default, FDB refresh is enabled. To enable or disable FDB refresh, use the following command:

```
G8272(config)# [no] vlag mac-address-table refresh
```

vLAG Startup Delay

A startup delay is used to prevent traffic loss while a vLAG peer reloads. During the reload process, a switch's vLAG ports will be in the error disabled state. Once the ISL is established, the vLAG ports will transition to the up state one at a time only after the configured startup delay has passed.

The startup delay is initiated only after the ISL is established. Until this timer has expired, all vLAG ports will be maintained in the error disabled state. This delay will allow the switch to properly initiate BGP or OSPF for routing traffic through the upstream links. This ensures that traffic will flow smoothly once the vLAG ports are up.

Any vLAG ports configured in the administrative up or down state will retain their states even after the startup delay has passed. During the delay, their status will be error disabled.

During the startup delay interval, administratively up ports that become members of the vLAG instance will be error disabled and administratively up ports that leave the vLAG instance will transition back to the up state.

Also, any vLAG ports that are configured in the administrative up state during the startup delay will be immediately put in the error disabled state until the delay interval expires.

Note: vLAG configuration changes are not recommended during the reload process of a vLAG peer on which the startup delay timer has not expired.

To configure the startup delay, in seconds, use the following command:

```
G8272(config)# vlag startup-delay <0-3600>
```

By default, the startup delay is 120 seconds (2 minutes). To reset the interval to its default value, use the following command:

```
G8272(config)# no vlag startup-delay
```

Note: The startup delay is cancelled if the two vLAG peers reload simultaneously or the ISL fails.

vLAG Auto-recovery

During network operations, it is common for the vLAG peers to be reloaded simultaneously. In this scenario, if the vLAG ports are in the error disabled state, traffic will be lost.

Usually, to avoid traffic loss, the ports should be manually brought up. When this is omitted, an auto-recovery mechanism is required to automatically transition ports from the error disabled state to the up state.

vLAG Auto-recovery allows the switch to automatically re-enable any error disabled vLAG ports after a specified time interval, if it detects that its vLAG peer is not functional.

During the reloading process, the auto-recovery timer is initiated. If the ISL is established between the vLAG peers before the auto-recovery timer expires, the timer is cancelled.

If the auto-recovery timer expires or the ISL fails before that, auto-recovery terminates the vLAG startup-delay process, and then determines the status of the vLAG peer through Health Check. If the peer is functional, all of the switch's vLAG ports are maintained in the error disabled state. If the vLAG peer is not detected as functional, the switch assumes the role of the primary, and then transitions all of its vLAG ports to the up state.

To configure the auto-recovery interval, in seconds, use the following command:

```
G8272(config)# vlag auto-recovery <240-3600>
```

Note: It is recommended that the auto-recovery interval be longer than the reload duration of the switch. Some vLAG ports might remain in the error disabled state, if this condition is not met.

Note: Any changes to the vLAG auto-recovery mechanism will not take effect until the switch is reloaded.

vLAG auto-recovery is always enabled and can't be disabled. By default, the auto-recovery interval is set to 300 seconds (5 minutes). To reset the interval to its default value, use the following command:

```
G8272(config)# no vlag auto-recovery
```

vLAG Tier ID

vLAG uses a Tier Identifier to separate different vLAG domains. This is helpful when configuring multiple vLAGs in a multi-tier network environment, because vLAG domains with different configured tier IDs can't be aggregated together.

To configure the tier ID of a vLAG, use the following command:

```
G8272(config)# vlag tier-id <1-512>
```

To remove a vLAG's tier ID, use the following command:

```
G8272(config)# no vlag tier-id
```

Health Check

The status of a vLAG peer is usually monitored through the ISL. You can configure vLAG Health Check to provide an alternate way of checking the status of the peer when the ISL fails.

Health Check detects and mitigates the failure of the vLAG ISL. When Health Check determines that the vLAG peer is still functioning, but the ISL has failed, the path from the secondary switch to the access switch will be shut down.

The Health Check mechanism can be configured either over the management interface or over an ethernet port.

If a switch is reloaded and its ISL is in the down state, the switch is not elected. When both the primary and secondary switches are reloaded with ISL in the down state, all the ports that are members of the vLAG from both switches will transition to the error disabled state.

For the Health Check mechanism to work, the vLAG peers should be in the same subnetwork (preferably directly connected).

The Health Check mechanism uses hello messages to determine the status of the vLAG peer. The hello messages are exchanged periodically between the vLAG peers over a TCP session.

To configure the interval, in seconds, between consecutive hello messages, use the following command:

```
G8272(config)# vlag hlthchk keepalive-interval <2-300>
```

By default, hello messages are exchanged every 5 seconds. To reset the hello interval to its default value, use the following command:

```
G8272(config)# no vlag hlthchk keepalive-interval
```

The hello messages are used to maintain the TCP session between the vLAG peers. After sending a certain number of consecutive hello messages and not receiving any reply, the local switch will declare its vLAG peer as being in the down state.

To configure the number of unanswered hello messages, use the following command:

```
G8272(config)# vlag hlthchk keepalive-attempts <1-24>
```

By default, the vLAG peer is marked as down after the local switch doesn't receive a reply to 3 consecutive hello messages. To reset the number of attempts to the default value, use the following command:

```
G8272(config)# no vlag hlthchk keepalive-attempts
```

When a TCP session between the two vLAG peers cannot be established, the Health Check mechanism will retry to reconnect to the peer at a certain interval. To configure the retry interval, in seconds, use the following command:

```
G8272(config)# vlag h1thchk retry-interval <1-300>
```

By default, the retry interval is set at 30 seconds. To reset it to this value, use the following command:

```
G8272(config)# no vlag h1thchk retry-interval
```

Basic Health Check Configuration Example

To configure the vLAG Health Check mechanism, follow the steps below:

1. Configure the management interface or an ethernet port of the switch.

```
G8272(config)# interface mgmt 0
G8272(config-if)# ip address 10.10.10.1 255.255.255.0
G8272(config-if)# exit
```

Note: Configure a similar interface on vLAG Peer 2. For example, use IP address 10.10.10.2.

2. Specify the IPv4 address of the vLAG peer:

```
G8272(config)# vlag h1thchk peer-ip 10.10.10.2 vrf management
```

Note: For vLAG Peer 2, the interface should be configured as 10.10.10.2 and the health check peer IP should be configured as 10.10.10.1, pointing back to vLAG Peer 1.

The local health check peer IP address cannot be configured manually. It is automatically configured within the same subnet as the peer's vLAG health check IP address.

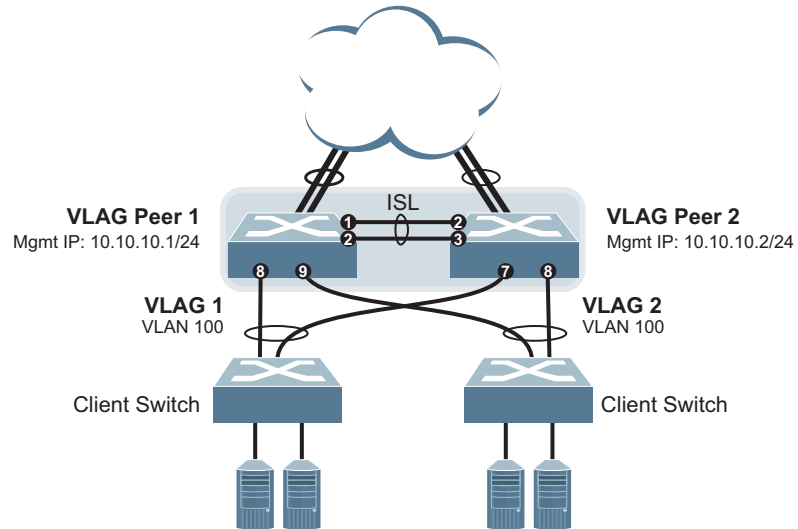
After configuring the peer IP address, the Health Check mechanism will attempt to establish a TCP session with the specified switch. To close the connection, use the following command:

```
G8272(config)# no vlag h1thchk peer-ip
```

Basic vLAG Configuration Example

Figure 6 shows an example configuration where two vLAG peers are used for aggregating traffic from downstream devices.

Figure 6. 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 the ISL ports and place them into a port LAG:

```
G8272(config)# interface ethernet 1/1-2
G8272(config-if-range)# bridge-port mode trunk
G8272(config-if-range)# bridge-port trunk allowed vlan 100
G8272(config-if-range)# aggregation-group 10 mode active
G8272(config-if-range)# exit
```

Notes:

- In this case, a LACP LAG is shown. A static LAG could be configured instead.
- ISL ports and vLAG ports must be members of the same VLANs.
- LAG 10 is automatically created when you assign the ethernet ports to it.

2. Associate the LAG to the ISL:

```
G8272(config)# vlag isl port-aggregation 10
```

3. Configure Health Check:

```
G8272(config)# vlag htlhchk peer-ip 10.10.10.2 vrf management
```

4. Configure VLAG Tier ID:

```
G8272(config)# vlag tier-id 10
```

5. Configure the ISL for the vLAG peer.

Make sure you configure the vLAG peer (vLAG Peer 2) using the same ISL aggregation type (LACP 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, follow the steps below.

1. Create VLAN 100:

```
G8272(config)# vlan 100
G8272(config-vlan)# exit
```

2. Configure the vLAG ports:

```
G8272(config)# interface ethernet 1/8
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 100
G8272(config-if)# aggregation-group 1 mode active
G8272(config-if)# exit

G8272(config)# interface ethernet 1/9
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 100
G8272(config-if)# aggregation-group 2 mode active
G8272(config-if)# exit
```

Note: In MSTP mode, VLANs are automatically mapped to CIST.

Note: LAGs 1 and 2 are automatically created when you assign ethernet interface 8 and 9 to them.

3. Associate the ports with their respective vLAGs:

```
G8272(config)# vlag instance 1 port-aggregation 1
G8272(config)# vlag instance 1 enable
G8272(config)# vlag instance 2 port-aggregation 2
G8272(config)# vlag instance 2 enable
```

4. Enable vLAG globally:

```
G8272(config)# vlag enable
```

5. Verify the completed configuration:

```
G8272# display vlag information
```

6. Repeat the configuration for vLAG Peer 2.

For each corresponding vLAG on the peer, the port LAG type (LACP or static), VLAN, STP mode and Tier ID must be the same as on vLAG Peer 1.

vLAG Configuration - VLANs Mapped to a MST Instance

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:

```
G8272(config)# spanning-tree mode mst
```

2. Configure the ISL ports and place them into a LAG (LACP or static):

```
G8272(config)# interface ethernet 1/1-2
G8272(config-if-range)# bridge-port mode trunk
G8272(config-if-range)# bridge-port trunk allowed vlan 100
G8272(config-if-range)# aggregation-group 1 mode active
G8272(config-if-range)# exit
```

Notes:

- In this case, a LACP LAG is shown. A static LAG could be configured instead.
- ISL ports and vLAG ports must be members of the same VLANs.

3. Associate the LAG to the ISL:

```
G8272(config)# vlag isl port-aggregation 1
```

4. Configure the vLAG Tier ID:

```
G8272(config)# vlag tier-id 10
```

5. Configure the ISL for the vLAG peer.

Make sure you configure the vLAG peer (vLAG Peer 2) using the same ISL aggregation type (LACP 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, follow the steps below.

1. Create the VLAN:

```
G8272(config)# vlan 100
G8272(config-vlan)# exit
```

2. Configure the vLAG ports:

```
G8272(config)# interface ethernet 1/3
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 100
G8272(config-if)# aggregation-group 3 mode active
G8272(config-if)# exit

G8272(config)# interface ethernet 1/4
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 100
G8272(config-if)# aggregation-group 4 mode active
G8272(config-if)# exit
```

Note: In MSTP mode, VLANs are automatically mapped to CIST.

Note: LAGs 3 and 4 are automatically created when you assign ethernet interface 3 and 4 to them.

3. Map the VLAN to an MST instance:

```
G8272(config)# spanning-tree mst configuration
G8272(config-mst)# instance 1 vlan 100
```

4. Associate the ports with their respective vLAGs:

```
G8272(config)# vlag instance 1 port-aggregation 3
G8272(config)# vlag instance 1 enable
G8272(config)# vlag instance 2 port-aggregation 4
G8272(config)# vlag instance 2 enable
```

5. Enable vLAG:

```
G8272(config)# vlag enable
```

6. Verify the completed configuration:

```
G8272# display vlag information
```

7. Configure the ISL for the vLAG peer.

Make sure you configure the vLAG peer (vLAG Peer 2) using the same ISL aggregation type (LACP 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 vLAGs in Multiple Layers

Figure 7. vLAG in Multiple Layers

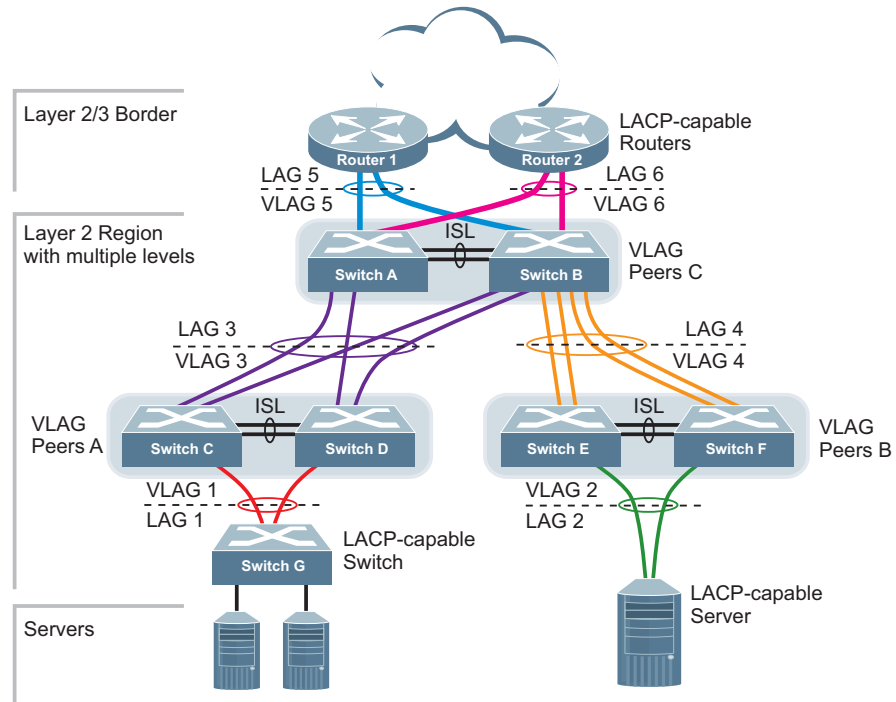


Figure 7 shows an example of vLAG being used in a multi-layer environment. Following are the configuration steps for the topology.

Task 1: Configure Layer 2/3 Border Region

Consider the following:

- Border routers 1 and 2 are connected to Switch A through port 1
- Border routers 1 and 2 are connected to Switch B through port 2

Configuring Border Router 1

1. Create a VLAN:

```
G8272(config)# vlan 50
G8272(config-vlan)# exit
```

2. Configure ports on border router 1 as follows:

```
G8272(config)# interface ethernet 1/1-2
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 50
G8272(config-if)# aggregation-group 5 mode active
G8272(config-if)# exit
```

Configuring Border Router 2

1. Create a VLAN:

```
G8272(config)# vlan 60
G8272(config-vlan)# exit
```

2. Configure ports on border router 2 as follows:

```
G8272(config)# interface ethernet 1/1-2
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 60
G8272(config-if)# aggregation-group 6 mode active
G8272(config-if)# exit
```

Task 2: Configure switches in the Layer 2 region

Consider the following:

- ISL ports on switches A - F: ports 1, 2
- Ports connecting switches A and B to the Layer 2/3 border routers: ports 5, 6
- Ports on switches A and B connecting to switches C and D: ports 10, 11
- Ports on switches C and D connecting to switches A and B: 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
- Ports on switch E connecting to switch B: ports 15, 16
- Ports on switch F connecting to switch B: ports 17, 18
- Port on switches C and D connecting to switch G: port 3
- Port on switches E and F connecting to server: port 3

Configuring Switch A

1. Configure vLAG tier ID and enable vLAG globally:

```
G8272(config)# vlag tier-id 10
G8272(config)# vlag enable
```

2. Configure ISL ports on Switch A:

```
G8272(config)# interface ethernet 1/1-2
G8272(config-if-range)# bridge-port mode trunk
G8272(config-if-range)# bridge-port trunk allowed vlan 30,50,60,100
G8272(config-if-range)# aggregation-group 10 mode active
G8272(config-if-range)# exit
```

3. Associate the LAG to the ISL:

```
G8272(config)# vlag isl port-aggregation 10
```

4. Create VLANs 30, 50, 60 and 100:

```
G8272(config)# vlan 30,50,60,100
G8272(config-vlan)# exit
```

5. Configure ports on Switch A connecting to the Layer 2/3 border routers:

```
G8272(config)# interface ethernet 1/5
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 50
G8272(config-if)# aggregation-group 5 mode active
G8272(config-if)# exit

G8272(config)# interface ethernet 1/6
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 60
G8272(config-if)# aggregation-group 6 mode active
G8272(config-if)# exit
```

6. Configure ports on switch A connecting to switches C and D:

```
G8272(config)# interface ethernet 1/10-11
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 30
G8272(config-if)# aggregation-group 3 mode active
G8272(config-if)# exit
```

7. Associate the ports with their respective vLAGs:

```
G8272(config)# vlag instance 3 port-aggregation 3
G8272(config)# vlag instance 3 enable

G8272(config)# vlag instance 5 port-aggregation 5
G8272(config)# vlag instance 5 enable

G8272(config)# vlag instance 6 port-aggregation 6
G8272(config)# vlag instance 6 enable
```

Configuring Switch B

1. Configure vLAG tier ID and enable vLAG globally:

```
G8272(config)# vlag tier-id 10
G8272(config)# vlag enable
```

2. Configure ISL ports on Switch B:

```
G8272(config)# interface ethernet 1/1-2
G8272(config-if-range)# bridge-port mode trunk
G8272(config-if-range)# bridge-port trunk allowed vlan 30,40,50,60,100
G8272(config-if-range)# aggregation-group 10 mode active
G8272(config-if-range)# exit
```

3. Associate the LAG to the ISL:

```
G8272(config)# vlag isl port-aggregation 10
```

4. Create VLANs 30, 40, 50, 60 and 100:

```
G8272(config)# vlan 30,40,50,60,100  
G8272(config-vlan)# exit
```

5. Configure ports on Switch B connecting to the Layer 2/3 border routers:

```
G8272(config)# interface ethernet 1/5  
G8272(config-if)# bridge-port mode trunk  
G8272(config-if)# bridge-port trunk allowed vlan 50  
G8272(config-if)# aggregation-group 5 mode active  
G8272(config-if)# exit  
  
G8272(config)# interface ethernet 1/6  
G8272(config-if)# bridge-port mode trunk  
G8272(config-if)# bridge-port trunk allowed vlan 60  
G8272(config-if)# aggregation-group 6 mode active  
G8272(config-if)# exit
```

6. Configure ports on switch B connecting to switches C and D:

```
G8272(config)# interface ethernet 1/10-11  
G8272(config-if)# bridge-port mode trunk  
G8272(config-if)# bridge-port trunk allowed vlan 30  
G8272(config-if)# aggregation-group 3 mode active  
G8272(config-if)# exit
```

7. Configure ports on switch B connecting to switches E and F:

```
G8272(config)# interface ethernet 1/15-18  
G8272(config-if)# bridge-port mode trunk  
G8272(config-if)# bridge-port trunk allowed vlan 40  
G8272(config-if)# aggregation-group 4 mode active  
G8272(config-if)# exit
```

8. Associate the ports with their respective vLAGs:

```
G8272(config)# vlag instance 3 port-aggregation 3  
G8272(config)# vlag instance 3 enable  
  
G8272(config)# vlag instance 4 port-aggregation 4  
G8272(config)# vlag instance 4 enable  
  
G8272(config)# vlag instance 5 port-aggregation 5  
G8272(config)# vlag instance 5 enable  
  
G8272(config)# vlag instance 6 port-aggregation 6  
G8272(config)# vlag instance 6 enable
```

Configuring Switches C and D

1. Configure vLAG tier ID and enable vLAG globally:

```
G8272(config)# vlag tier-id 20
G8272(config)# vlag enable
```

2. Configure ISL ports on Switch C:

```
G8272(config)# interface ethernet 1/1-2
G8272(config-if-range)# bridge-port mode trunk
G8272(config-if-range)# bridge-port trunk allowed vlan 10,30,100
G8272(config-if-range)# aggregation-group 10 mode active
G8272(config-if-range)# exit
```

3. Associate the LAG to the ISL:

```
G8272(config)# vlag isl port-aggregation 10
```

4. Create VLANs 10, 30 and 100:

```
G8272(config)# vlan 10,30,100
G8272(config-vlan)# exit
```

5. Configure ports on Switch C connecting to switches A and B:

```
G8272(config)# interface ethernet 1/10-11
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 30
G8272(config-if)# aggregation-group 3 mode active
G8272(config-if)# exit
```

6. Configure ports on switch C connecting to switch G:

```
G8272(config)# interface ethernet 1/3
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 10
G8272(config-if)# aggregation-group 1 mode active
G8272(config-if)# exit
```

7. Associate the ports with their respective vLAGs:

```
G8272(config)# vlag instance 1 port-aggregation 1
G8272(config)# vlag instance 1 enable

G8272(config)# vlag instance 3 port-aggregation 3
G8272(config)# vlag instance 3 enable
```

8. Repeat the Above steps for Switch D.

Configuring Switch E

1. Configure vLAG tier ID and enable vLAG globally:

```
G8272(config)# vlag tier-id 30
G8272(config)# vlag enable
```

2. Configure ISL ports on Switch E:

```
G8272(config)# interface ethernet 1/1-2
G8272(config-if-range)# bridge-port mode trunk
G8272(config-if-range)# bridge-port trunk allowed vlan 20,40,100
G8272(config-if-range)# aggregation-group 10 mode active
G8272(config-if-range)# exit
```

3. Associate the LAG to the ISL:

```
G8272(config)# vlag isl port-aggregation 10
```

4. Create VLANs 20, 40 and 100:

```
G8272(config)# vlan 20,40,100
G8272(config-vlan)# exit
```

5. Configure ports on Switch E connecting to switch B:

```
G8272(config)# interface ethernet 1/15-16
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 40
G8272(config-if)# aggregation-group 4 mode active
G8272(config-if)# exit
```

6. Configure ports on switch E connecting to the server:

```
G8272(config)# interface ethernet 1/3
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 20
G8272(config-if)# aggregation-group 2 mode active
G8272(config-if)# exit
```

7. Associate the ports with their respective vLAGs:

```
G8272(config)# vlag instance 2 port-aggregation 2
G8272(config)# vlag instance 2 enable

G8272(config)# vlag instance 4 port-aggregation 4
G8272(config)# vlag instance 4 enable
```

Configuring Switch F

1. Configure vLAG tier ID and enable vLAG globally:

```
G8272(config)# vlag tier-id 30
G8272(config)# vlag enable
```

2. Configure ISL ports on Switch F:

```
G8272(config)# interface ethernet 1/1-2
G8272(config-if-range)# bridge-port mode trunk
G8272(config-if-range)# bridge-port trunk allowed vlan 20,40,100
G8272(config-if-range)# aggregation-group 10 mode active
G8272(config-if-range)# exit
```

3. Associate the LAG to the ISL:

```
G8272(config)# vlag isl port-aggregation 10
```

4. Create VLANs 20, 40 and 100:

```
G8272(config)# vlan 20,40,100
G8272(config-vlan)# exit
```

5. Configure ports on Switch E connecting to switch B:

```
G8272(config)# interface ethernet 1/17-18
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 40
G8272(config-if)# aggregation-group 4 mode active
G8272(config-if)# exit
```

6. Configure ports on switch E connecting to the server:

```
G8272(config)# interface ethernet 1/3
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 20
G8272(config-if)# aggregation-group 2 mode active
G8272(config-if)# exit
```

7. Associate the ports with their respective vLAGs:

```
G8272(config)# vlag instance 2 port-aggregation 2
G8272(config)# vlag instance 2 enable

G8272(config)# vlag instance 4 port-aggregation 4
G8272(config)# vlag instance 4 enable
```

Chapter 13. 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” on page 242](#)
- [“Class Maps” on page 243](#)
- [“Policy Maps” on page 252](#)
- [“WRED” on page 260](#)
- [“Interface Service Policy” on page 262](#)
- [“Control Plane Protection” on page 257](#)
- [“Microburst Detection” on page 263](#)

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.

The basic QoS model works as follows:

- Classify traffic:
 - Match DSCP or IP Precedence value
 - Match IP RTP priority
 - Match Protocol (ARP, DHCP, IS-IS etc) value
 - Match 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
 - Mark 802.1p Priority, DSCP or IP Precedence values
 - 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.
 - Configure Traffic Shaping.
 - Define a guaranteed bandwidth.
 - Configure Strict Priority queuing.

Class Maps

A *class map* is a named object that represents a class of traffic. In the class map, you specify a set of match criteria for classifying the packets. You can then reference class maps in policy maps.

You can then reference class maps in policy maps.

You define the following class-map and policy-map object types when you create them:

- *control plane (CoPP)*: receives packets that are necessary to configure the switch hardware and for remote switch management
- *qos*: defines a QoS class map that is used to match packets to a specified class.
- *queuing*: defines a queuing class map that is used to match packets to a specified class.

To configure a class map, use one of the following commands:

```
G8272(config)# class-map type control-plane match-any <CoPP class map name>
```

```
G8272(config)# class-map type qos {match-all|match-any} <class map name>
```

```
G8272(config)# class-map type queuing match-any <queue name>
```

QoS Classification Types

You can classify traffic by matching packets based on predefined classification criteria, as follows:

- Access Control List (ACL)
- Class of Service (CoS)
- DiffServ Code Point (DSCP)
- UDP or TCP port number
- Precedence
- Protocol

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.

To set an QoS ACL filter, use the following command:

```
G8272(config-cmap-qos)# match [not] access-group name <ACL name>
```

The G8272 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/UDP: Source port, destination port, TCP flag.
- Packet format

For ACL details, see [Chapter 6, "Access Control Lists."](#)

Summary of QoS Actions

Actions determine how the traffic is treated. The G8272 QoS actions include the following:

- Re-mark a new DiffServ Code Point (DSCP) or IP Precedence
- Re-mark the 802.1p field
- Set the COS queue

Using Class of Service Filters

You can classify traffic based on the class of service (CoS). To do that, use the following command:

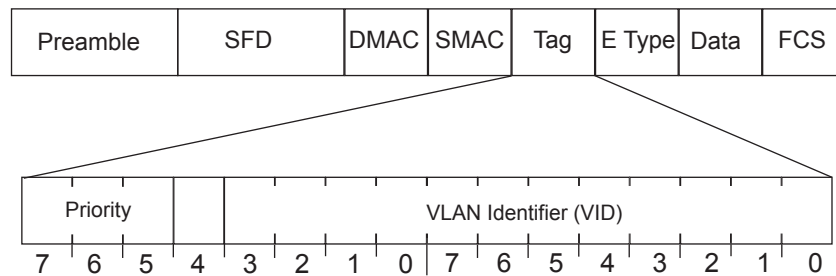
```
G8272(config-cmap-qos)# match [not] cos <CoS value (0-7)>
```

Using 802.1p Priority to Provide QoS

The G8272 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 8. Layer 2 802.1q/802.1p VLAN tagged packet



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 scheduling weight of the COS queue. COS queue number. Higher COS queue numbers provide forwarding precedence.

Using DiffServ Code Point (DSCP) Filters

The switch uses the Differentiated Services (DiffServ) architecture to provide QoS functions. DiffServ is described in IETF RFCs 2474 and 2475.

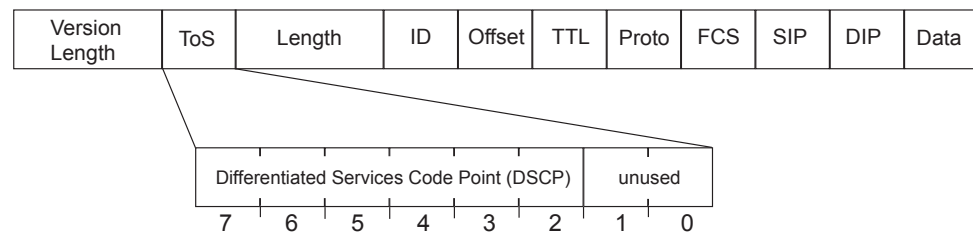
The six most significant bits in the TOS byte of the IP header are defined as DiffServ Code Points (DSCP). Packets are marked with a certain value depending on the type of treatment the packet must receive in the network device. DSCP is a measure of the Quality of Service (QoS) level of the packet.

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 9. Layer 3 IPv4 packet



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.

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 10 shows the default service levels provided by the switch, listed from highest to lowest importance:

Table 10. *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

Using TCP/UDP Port Filters

Defines the TCP/UDP port used by Real-time Transport Protocol (RTP) processes as the classification criterion.

To apply this filter, use the following command:

```
G8272(config-cmap-qos)# match [not] ip rtp <TCP port or UDP port>
```

Using Precedence Filters

Defines the precedence as the classification criterion. The precedence value is from 0 to 7 or it can be one of the following:

- **routine**: Routine precedence (0)
- **priority**: Priority precedence (1)
- **immediate**: Immediate precedence (2)
- **flash**: Flash precedence (3)
- **flash-override**: Flash override precedence (4)
- **critical**: Critical precedence (5)
- **internet**: Internetwork control precedence (6)
- **network**: Network control precedence (7)

To apply this filter, use the following command:

```
G8272(config-cmap-qos)# match [not] precedence <precedence value (0-7)>
```


Using Protocol Filters

To define a protocol as a QoS filter, use the following command:

```
G8272(config-cmap-qos)# match [not] protocol <protocol name>
```

where *protocol name* is one of the following:

- **arp** - Address Resolution Protocol
- **bridging** - Bridging
- **cdp** - CISCO Discovery Protocol
- **clns** - Connectionless Network Service
- **clns-es** - CLNS End Systems
- **clns-is** - CLNS Intermediate Systems
- **dhcp** - Dynamic Host Configuration
- **isis** - Intermediate System to Intermediate System
- **ldp** - Label Distribution Protocol
- **netbios** - NetBIOS extended user interface

Queuing Classification Types

The switch has the following eight predefined queues:

- 1p7q1t-out-q-default (queue 0 or default queue)
- 1p7q1t-out-pq1 (queue 1 or priority queue)
- 1p7q1t-out-q2 (queue 2)
- 1p7q1t-out-q3 (queue 3)
- 1p7q1t-out-q4 (queue 4)
- 1p7q1t-out-q5 (queue 5)
- 1p7q1t-out-q6 (queue 6)
- 1p7q1t-out-q7 (queue 7)

1p7q1t refers to how many priority queues, standard queues and thresholds are supported. In this case, there are:

- 1 priority queue
- 7 standard queues
- 1 threshold

To apply this filter, use the following command:

```
G8272(config)# class-map type queuing match-any <queue name>
```

You can filter packets to different queues based on CoS or QoS group.

Class Map Configuration Examples

Below are two basic configuration examples for QoS and queueing class maps.

QoS Class Map Configuration Example

To configure a QoS class map, follow the steps below:

1. Create a new QoS class map:

```
G8272(config)# class-map type qos cmap-qos-01
G8272(config-cmap-qos)#
```

2. Define the classification criteria (multiple ones can be configured):

```
G8272(config-cmap-qos)# match protocol arp
G8272(config-cmap-qos)# match precedence 3
```

3. Verify the QoS class map configuration:

```
G8272# display class-map type qos

Type qos class-maps
=====
class-map type qos match-any class-default

class-map type qos cmap-qos-01
  match precedence 3
  match protocol arp
```

Queueing Class Map Configuration Example

To configure a queueing class map, follow the steps below:

1. Create a new queueing class map:

```
G8272(config)# class-map type queueing match-any 1p7q1t-out-q2
G8272(config-cmap-que)#
```

2. Define the classification criteria (multiple ones can be configured):

```
G8272(config-cmap-que)# match cos 5
G8272(config-cmap-que)# match qos-group 2
```

3. Verify the QoS class map configuration:

```
G8272# display class-map type queueing

Type queueing class-maps
=====
class-map type queueing match-any 1p7q1t-out-q4

class-map type queueing match-any 1p7q1t-out-q2
  match qos-group 2
  match cos 5
...
```

Policy Maps

A *policy map* is a named object that represents a set of policies that are to be applied to a set of traffic classes, such as limiting the bandwidth or dropping packet.

The following predefined policy maps are used as default service policies:

- `control-plane`: Control Plane Protection (CoPP)
- `qos`: Quality of Service (QoS)
- `queuing`

To set a policy map, use the following command:

```
G8272(config)# policy-map type {control-plane <policy-map name>|qos  
<policy-map name>|queuing}
```

Ingress Policing

The switch can use ingress policing to monitor the data rates for a particular class of traffic. When the data rate exceeds user-configured values, the switch drops packets immediately.

Defining Single-Rate and Dual-Rate Policers

You can define single-rate and dual-rate policers. While single-rate policers monitor the committed information rate (CIR) of traffic, dual-rate policers monitor both CIR and peak information rate (PIR) of traffic. Also, the system monitors associated burst sizes.

One action can be configured for each condition.

To configure polices, use the following command:

```
G8272(config-pmap-c-qos)# police [cir] {<committed-rate> [<data-rate>]|percent  
<CIR-link-percent>} [bc <committed-burst-rate> [<link-speed>]] [pir] {<peak-rate>  
[<data-rate>]|percent <CIR-link-percent>} [be <peak-burst-rate> [<link-speed>]] [conform  
<conform actions>] [exceed <exceed actions>] [violate <violate actions>]
```

where:

- `cir` specifies to use the committed information rate, or desired bandwidth, as a bit rate or a percentage of the link rate.
- `committed-rate` can range from 8,000 bps to the maximum rate of the link, in steps that are multiples of 8,000 bps.
- `data-rate` can have one of the following units:
 - `bps`: bits per second
 - `kbps`: 1,000 bits per second
 - `mbps`: 1,000,000 bits per second
 - `Gbps`: 1,000,000,000 bits per second
- `percent` specifies the rate as a percentage of the interface bandwidth.

- **bc** indicates how much the cir can be exceeded, either as a bit rate or an amount of time at CIR.
- *committed-burst-rate* has the following options:
 - **bytes**: bytes
 - **kbytes**: 1,000 bytes
 - **mbytes**: 1,000,000 bytes
 - **ms**: milliseconds
 - **us**: microseconds
- **pir** sets a peak information rate value.
- **be** indicates how much the pir can be exceeded, either as a bit rate or an amount of time at pir.
- **conform** specifies to take action if the traffic data rate is within bounds. The following actions are allowed:

Table 11. *Conform Actions*

Conform Actions	Description
transmit	Sends over the packet. This action is applicable only when the packet conforms to the parameters.
set-prec-transmit	Sets the IP precedence field to a specified value and transmits the packet.
set-dscp-transmit	Sets the Differentiated Service Code Point (DSCP) field to a specified value and transmits the packet.
set-cos-transmit	Sets the class of service (CoS) field to a specified value and transmits the packet.

The default action is **transmit**.

- **exceed** specifies to take action if traffic data rate is exceeded.
- **violate** specifies to take action if traffic data rate violates the set rate values.

The following actions are available for both **exceed** and **violate** scenarios:

Table 12. *Exceed and Violate Actions*

Exceed and Violate Actions	Description
drop	Drops the packet. This is only available when the packet exceeds or violates the parameters.
set-dscp-transmit	Sets the Differentiated Service Code Point (DSCP) field to a specified value and transmits the packet.
set-cos-transmit	Sets the class of service (CoS) field to a specified value and transmits the packet.

Marking

Marking is a method that you use to modify the QoS fields of the incoming and outgoing packets.

To do so, use the following command:

```
G8272(config-pmap-c-qos)# apply {cos <0-7>|dscp <0-63>}precedence <0-7>|  
qos-group <0-7>}
```

Queuing Policing

Traffic queuing is a mechanism allowing the switch to order the packets and it applies to output data. Device modules can support multiple queues, which you can use to control the sequencing of packets in different traffic classes.

Bandwidth

Bandwidth on egress queues can be configured to allocate a minimum percentage of the interface bandwidth to a queue. To set the bandwidth, use the following command:

```
G8272(config-pmap-c-que)# bandwidth {<rate> {bps|kpbs|mbps|gbps}|percent  
<percent value>}
```

You can configure the bandwidth by specifying its bit rate, or as a percentage of the underlying link rate.

You can also configure the data rate as a percentage of the bandwidth that isn't allocated to other classes by using the following command:

```
G8272(config-pmap-c-que)# bandwidth remaining percent <percent value>
```

Shaping

You can configure shaping on an egress queue to impose a maximum rate on it. A data rate may be configured by either using the bit rate or as a percentage of the underlying interface link rate.

To se shaping, use the following command:

```
G8272(config-pmap-c-que)# shape [average] {<rate> {bps|kpbs|mbps|gbps}|  
percent <percent value>}
```

Priority

You can configure only one level of priority on an egress priority queue. To do so, use the following command:

```
G8272(config-pmap-c-que)# priority [level 1]
```

Policy Map Configuration Example

Below are two basic examples of configuring QoS and queueing policy maps.

QoS Policy Map Configuration Example

To configure a QoS policy map, follow the steps below:

1. Create a new QoS policy map:

```
G8272(config)# policy-map type qos pmap-qos-01
G8272(config-pmap-qos)#
```

2. Select an existing QoS class map:

```
G8272(config-pmap-qos)# class type qos cmap-qos-01
G8272(config-pmap-c-qos)#
```

3. Configure the ingress policing or marking:

```
G8272(config-pmap-c-qos)# police cir 2 mbps conform transmit
G8272(config-pmap-c-qos)# apply cos 3
```

4. Verify the QoS policy map configuration:

```
G8272# display policy-map type qos

Type qos policy-maps
=====

policy-map type qos pmap-qos-01
  class type qos cmap-qos-01
    apply cos 3
    police cir 2 mbps conform transmit
```

Queuing Policy Map Configuration Example

To configure a queuing policy map, follow the steps below:

1. Create a new queuing policy map:

```
G8272(config)# policy-map type queuing pmap-que-01
G8272(config-pmap-que)#
```

2. Select an existing queuing class map:

```
G8272(config-pmap-que)# class type queuing 1p7q1t-out-q3
G8272(config-pmap-c-que)#
```

3. Configure the queuing policing:

```
G8272(config-pmap-c-que)# bandwidth percent 25
G8272(config-pmap-c-que)# random-detect minimum-threshold percent 5
maximum-threshold percent 35
```

4. Verify the queuing policy map configuration:

```
G8272# display policy-map type queuing

Type queuing policy-maps
=====

policy-map type queuing default-out-policy
  class type queuing 1p7q1t-out-pq1
    priority level 1
  class type queuing 1p7q1t-out-q2
  class type queuing 1p7q1t-out-q3
  class type queuing 1p7q1t-out-q-default
    bandwidth remaining percent 25

policy-map type queuing pmap-que-01
  class type queuing 1p7q1t-out-q2
  class type queuing 1p7q1t-out-q3
    random-detect minimum-threshold percent 5 maximum-threshold percent 35
  bandwidth percent 25
```

Control Plane Protection

Control Plane Protection (CoPP) receives packets that are required for the internal protocol state machines. This type of traffic is usually received at low rate. However, in some situations such as DOS attacks, the switch may receive this traffic at a high rate. If the control plane protocols are unable to process the high rate of traffic, the switch may become unstable.

The control plane receives packets that are channeled through protocol-specific packet queues. Multiple protocols can be channeled through a common packet queue. However, one protocol cannot be channeled through multiple packet queues. These packet queues are applicable only to the packets received by the software and does not impact the regular switching or routing traffic. Packet queue with a higher number has higher priority.

You can configure the bandwidth for each packet queue. Protocols that share a packet queue will also share the bandwidth.

Given below are the commands to configure CoPP:

- Enter Control Plane Configuration mode by using the following command:

```
G8272(config)# control-plane
```

- To apply the control plane service policy, use the command below:

```
G8272(config-cp)# service-policy input copp-system-policy
```

- Modifies a control-plane policy-map:

```
G8272(config)# [no] policy-map type control-plane <policy-map name>
```

- Modifies a control-plane class map that is used to match packets to a specified class:

```
G8272(config)# [no] class-map type control-plane match-any <class map name>
```

To view CoPP switch configuration, use the following commands:

```
G8272(config)# display class-map type control-plane  
G8272(config)# display policy-map type control-plane  
G8272(config)# display policy-map interface control-plane
```

Control Plane Configuration Examples

To configure a CoPP class map, use the following steps:

1. Since new CoPPs cannot be created, modify one of the predefined CoPP class maps. To view the current CoPP class maps, use the following command:

```
G8272(config)# display class-map type control-plane

Type control plane class-maps
=====
class-map match-any copp-s-lacp

class-map match-any copp-s-default

class-map match-any copp-s-bfd

class-map match-any copp-s-arpreponse

class-map match-any copp-s-arprequest

class-map match-any copp-s-authentication
      QOS-ACCESS-LIST-NAME: copp-system-acl-authentication

...
```

2. Enter configuration mode for one of the above CoPP class maps:

```
G8272(config)# class-map type control-plane match-any copp-s-bfd
G8272(config-cmap-control-plane)#
```

3. Configure the classification criteria for the CoPP class map:

```
G8272(config-cmap-control-plane)# match access-group name myACL
```

Note: You cannot define multiple classification criteria for a single CoPP class map.

4. Verify the configuration:

```
G8272# display class-map type control-plane

Type control plane class-maps
=====
class-map match-any copp-s-lacp

class-map match-any copp-s-default

class-map match-any copp-s-bfd
      QOS-ACCESS-LIST-NAME: myACL

class-map match-any copp-s-arpreponse

class-map match-any copp-s-arprequest

class-map match-any copp-s-authentication
      QOS-ACCESS-LIST-NAME: copp-system-acl-authentication

...
```

To configure a CoPP policy map, use the following steps:

1. Since new CoPPs cannot be created, modify the predefined CoPP policy map (copp-system-policy). To view the current CoPP policy maps, use the following command:

```
G8272(config)# display policy-map type control-plane

Type control-plane policy-maps
=====

policy-map type control-plane copp-system-policy
  class type control-plane copp-s-default
    police pps 20000
  class type control-plane copp-s-ntp
    police pps 200
  class type control-plane copp-s-arprequest
    police pps 500
  class type control-plane copp-s-nd
    police pps 500
  ...
```

2. Enter configuration mode for one of the above CoPP policy maps:

```
G8272(config)# policy-map type control-plane match-any copp-system-policy
G8272(config-pmap-control-plane)#
```

3. Enter configuration mode for one of the above CoPP class maps:

```
G8272(config-pmap-control-plane)# class type control-plane copp-s-ntp
G8272(config-pmap-c-control-plane)#
```

4. Configure the number of packet per second processed by the switch for the current class map:

```
G8272(config-pmap-c-control-plane)# police pps 100
```

Note: The switch will not process more than the configured number of packets per second for each type of traffic. The excess packets will be dropped.

5. Verify the configuration:

```
G8272# display policy-map type control-plane

Type control-plane policy-maps
=====

policy-map type control-plane copp-system-policy
  class type control-plane copp-s-default
    police pps 20000
  class type control-plane copp-s-ntp
    police pps 100
  class type control-plane copp-s-arprequest
    police pps 500
  class type control-plane copp-s-nd
    police pps 500
  ...
```

WRED

Weighted Random Early Detection (WRED) is a congestion avoidance algorithm that helps prevent a TCP collapse, where a congested port indiscriminately drops packets from all sessions. The transmitting hosts wait to retransmit resulting in a dramatic drop in throughput. Often times, this TCP collapse repeats in a cycle, which results in a saw-tooth pattern of throughput. WRED selectively drops packets before the queue gets full, allowing majority of the traffic to flow smoothly.

WRED discards packets based on the CoS queues. Packets marked with lower priorities are discarded first.

For implementing WRED, you must define a profile with minimum threshold, maximum threshold, and a maximum drop probability. The profiles can be defined on a port or a CoS.

Configuring WRED

WRED can be configured only on physical ports and not on LAG. WRED is applicable only to unicast traffic.

To configure WRED, use the following command:

```
G8272(config)# policy-map type queuing <policy map name>
G8272(config-pmap-que)# class type queuing <class map name>
G8272(config-pmap-c-que)# random-detect {[minimum-threshold] {<min-threshold>
[packets|bytes|kbytes|mbytes|ms|us] |percent <min-percent>}
[maximum-threshold] {<max-threshold> [packets|bytes|kbytes|mbytes|ms|us]
|percent <max-percent>}
```

WRED Configuration Example

Follow these steps to enable WRED and configure a global and/or port-level profile. If you configure global and port-level profile, WRED uses the port-level profile to make transmit/drop decisions when experiencing traffic congestion.

1. Configure a policy-map of type queuing.

```
G8272(config)# policy-map type queuing policy-1
```

2. Assign a system defined class-map to the policy-map configured at step 1.

```
G8272(config-pmap-que)# class type queuing 1p7q4t-out-pq2
```

3. Enable WRED for the class-map by specifying (in bytes) the minimum and the maximum thresholds to be used.

```
G8272(config-pmap-c-que)# random-detect [minimum-threshold] <min-threshold>
bytes [maximum-threshold] <max-threshold> bytes
```

4. Use the following command to view global WRED information:

```
G8272(config)# display policy-map type queuing

Type queuing policy-maps
=====

policy-map type queuing default-out-policy
class type queuing 1p7q1t-out-pq1
  priority level 1
class type queuing 1p7q1t-out-q2
class type queuing 1p7q1t-out-q3
class type queuing 1p7q1t-out-q-default
  bandwidth remaining percent 25

policy-map type queuing policy_queue1
class type queuing 1p7q1t-out-q5
  shape average 4 bps
  random-detect minimum-threshold 12000 packets maximum-threshold 30000
  packets <<<<
class type queuing 1p7q1t-out-q3
...
```

- Use the following command to view port-level WRED information:

```
G8272(config)# display policy-map interface ethernet 1/4 output

Global statistics status : disabled

Ethernet1/4

Service-policy (queuing) output: policy-1

Class-map (queuing): 1p7q1t-out-q2 (match any)
  random-detect minimum-threshold 35000 bytes maximum-threshold 250000 bytes
```

Interface Service Policy

An *input* QoS policy is a service policy applied to incoming traffic on an Ethernet interface for classification. For type *queuing*, the output policy is applied to all outgoing traffic that matches the specified class.

To apply a policy to an interface, use one of the following commands:

```
G8272(config-if)# service-policy [type qos] input <policy-map-name>
```

```
G8272(config-if)# service-policy type queuing output <policy-map-name>
```

Limitations

- Service policy is applied only in the ingress for type `qos` and egress for type `queuing`.
- The device restricts QoS policies to one per interface per direction (ingress or egress) for each of the policy types `qos` and `queuing`.
- Queuing Policy-map is not support for Aggregator Port.

Microburst Detection

Microbursts are short peaks in data traffic that manifest as a sudden increase in the number of data packets transmitted over a specific millisecond level time frame, potentially overwhelming network buffers. Microburst detection allows users to analyze and mitigate microburst-related incidents, thus preventing network congestion.

Microburst Detection can be enabled or disabled on each switch interface. To enable it on an interface, use the following command (for this example, ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# microburst-detection enable threshold <threshold value>
```

To disable microburst detection on an interface, use the following command:

```
G8272(config-if)# no microburst-detection enable
```

To configure the polling interval (in milliseconds) used by microburst detection to evaluate traffic burst:

```
G8272(config)# microburst-detection interval <5-5000>
```

Note: By default, microburst detection is disabled.

To see the current microburst statistics, use the following command:

```
G8272# display statistics microburst
```

Below is a basic configuration example for Microburst Detection.

1. Enter Interface Configuration mode and enabled Microburst Detection, choosing a threshold value:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# microburst-detection enable threshold 20
G8272(config-if)# exit
```

2. Configure the polling interval:

```
G8272(config)# microburst-detection interval 100
```

3. See the current microburst statistics:

```
G8272# display statistics microburst
-----
Interface      # of uburst  avg size  max size  avg duration  max duration
-----
Ethernet1/12          0         0         0           0             0
```


Part 4: 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” on page 267](#)
- [“Routed Ports” on page 287](#)
- [“Address Resolution Protocol” on page 293](#)
- [“Internet Protocol Version 6” on page 301](#)
- [“Internet Group Management Protocol” on page 313](#)
- [“Border Gateway Protocol” on page 339](#)
- [“Open Shortest Path First” on page 389](#)
- [“Route Maps” on page 411](#)

Chapter 14. Basic IP Routing

This chapter provides configuration background and examples for using the G8272 to perform IP routing functions. The following topics are addressed in this chapter:

- [“IP Routing” on page 268](#)
- [“Routing Information Base” on page 271](#)
- [“Bidirectional Forwarding Detection” on page 272](#)
- [“Routing Between IP Subnets” on page 277](#)
- [“ECMP Static Routes” on page 282](#)
- [“Dynamic Host Configuration Protocol” on page 284](#)

IP Routing

IP Routing is the mechanism by which traffic travels across multiple networks, from its source to its destination. IP routing sends packets outside the local network by using Layer 3 protocols, like Address Resolution Protocol (ARP), Border Gateway Protocol (BGP) or Open Shortest Path First (OSPF).

The switch builds a forwarding table that correlates destination IP addresses with their next-hop addresses. If a packet is intended for a certain device, the switch will check its forwarding table for the IP address of the destination device. Once a match is found, the switch will send the packet through the interface on which the next-hop address is present. The next-hop device will perform the same process until the packet reaches its destination.

The switch uses a combination of configurable 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.

The G8272 supports both IPv4 and IPv6 routing. For more detailed information about IPv6 routing, see [Chapter 17, “Internet Protocol Version 6”](#).

IP routing runs only on Layer 3 interfaces, such as:

- Switch Virtual Interfaces (SVIs);
- The management interface;
- Ethernet interfaces configured as routed ports.

For more information about Layer 3 interfaces, see [Chapter 7, “Interface Management”](#) and [Chapter 15, “Routed Ports”](#).

By default, IP routing is enabled on the switch.

To enable or disable the forwarding of IPv4 traffic, use the following command:

```
G8272(config)# [no] ip forwarding
```

To enable or disable the forwarding of IPv6 traffic, use the following command:

```
G8272(config)# [no] ipv6 forwarding
```

Direct and Indirect Routing

Direct routing is used when both source and destination devices are present in the same network, the switch sends the packet directly from one device to the other. Forwarding table entries for direct routes are referred to as directly connected routes and are marked in the table with the letter 'C'.

When the source and destination devices aren't in the same network, the packet must be forwarded by a network node (router) that knows how to reach the destination device.

Static and Dynamic Routing

Static routing means that the packet is forwarded using a manually configured route. Configuring a static route teaches the switch how to reach the specified destination IP address. Forwarding table entries for static routes are marked with the letter 'S'.

Static routing is recommended for small networks, because every static route needs to be manually configured on each switch or router in the network.

To add or remove a static IPv4 route, use the following command:

```
G8272(config)# [no] ip route <IPv4 address>/<prefix length> <next-hop address>
```

For example:

```
G8272(config)# ip route 10.10.10.0/24 10.28.12.10
```

To add or remove a static IPv6 route, use the following command:

```
G8272(config)# [no] ipv6 route <IPv6 address>/<prefix length> <next-hop address>
```

For example:

```
G8272(config)# ipv6 route 2001:1::1/128 fe80::a8bb:ccff:fe00:300
```

For more details about the above commands, consult the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

Note: The switch supports a maximum of 256 static routes.

To view configured static routes, use the following command:

```
G8272# display {ip|ipv6} static-route
```

Dynamic routing means that the packet is forwarded using routes that the switch learns automatically via routing protocols, like BGP or OSPF. Routing protocols are used by switches or routers to exchange routing information. Forwarding table entries for dynamic routes are marked with the different letters depicting the protocol through which the route was learned. For example, 'B' is used for routes learned through BGP or 'O' for routes learned through OSPF.

Dynamic routing is fault tolerant. If a learned link goes down, the switch can learn another route to the destination IP address. Another advantage is the low administrative overhead compared to configuring static routing.

One of the disadvantages of dynamic routing is the increased memory and CPU usage due to the processing of routing information received from other routers.

For more information about routing protocols, see [Chapter 19, “Border Gateway Protocol”](#) and [Chapter 20, “Open Shortest Path First”](#).

Note: The switch supports a maximum of 15,872 IPv4 dynamic routes and 256 IPv4 static routes, or 6,144 IPv6 dynamic routes and 256 IPv6 static routes.

To remove a route from the forwarding table, use the following command:

```
G8272# remove {ip|ipv6} route <IP address>/<prefix length>
```

To delete all routes from the forwarding table, use the following command:

```
G8272# remove {ip|ipv6} route *
```

Default Gateway

The default gateway is the next-hop address that the switch uses when forwarding packets for which it doesn't have a dynamic or static route. The default gateway is a static route configured with 0.0.0.0/0 as its destination IP address.

To configure an IPv4 address as the default gateway for IPv4 traffic, use the following command:

```
G8272(config)# ip route 0.0.0.0/0 <next-hop address>
```

To configure an IPv6 address as the default gateway for IPv6 traffic, use the following command:

```
G8272(config)# ipv6 route 0::0/0 <next-hop address>
```

Note: Routes that have a specific destination take precedence over the default gateway.

Routing Information Base

The Routing Information Base (RIB) is a table in which the switch keeps information about known routes, like the destination IP address, the next-hop address or the associated metric (path cost). It consists of static entries (routes manually configured) and dynamic entries (routes learned through routing protocols like BGP or OSPF).

The main objective of routing protocols is the construction of the RIB. This implies the exchange of routing information between neighboring network devices.

A route redistribution can be requested by a routing protocol. When the switch receives the request, it will share its routing information based upon the protocol that made the request. For example, if BGP requested a route redistribution, the switch will share all the routes it has learned through BGP.

Any route changes on the switch will trigger route updates to be sent to the neighboring devices. A route update is shared only if the Hardware Specific Layer (HSL) gives a positive acknowledgement for that route.

Routing protocols, such as BGP, can use the Next-hop Lookup service to check the availability of a route.

RIB supports the maintaining of routes in a stale state for the duration of a graceful restart of the BGP service. This means that traffic forwarding isn't disturbed for the duration of the restart.

Routes with Indirect Next-hops

There are scenarios in which the switch isn't directly connect to the next-hop through an interface. Such occurrences are called indirect next-hops.

RIB supports both normal and ECMP routes with indirect next-hops. If there is a direct route through which the next-hop is reachable (the route leads directly to the next-hop), RIB will install the direct route, instead of an indirect one.

Bidirectional Forwarding Detection

Bidirectional Forwarding Detection (BFD) is a Layer 3 detection protocol used to detect failures in the forwarding path of adjacent network routers.

For BFD to work properly, a session must be established between two devices (or BFD peers). This means BFD must be configured on both BFD peers. Once BFD has been enabled on the connecting interfaces and on the peers, a BFD session is established, followed by the negotiation of BFD timers.

Note: The switch supports up to a maximum of 100 BFD sessions.

BFD uses few system resources and provides a fast failure detection, that is independent of all media types, encapsulations, topologies and routing protocols, like BGP. Once a failure is detected, BFD notifies the routing protocols of the local switch to trigger a routing table update, thus reducing network convergence time.

BFD control packets are encapsulated in User Datagram Protocol (UDP) packets, that have an assigned destination port of 3784 or 3785 and source port between 49152 and 65535. BFD control packets are always sent as unicast packets to the configured BFD peer.

BFD sessions are supported for both IPv4 and IPv6 addressing.

By default, BFD is enabled on the switch. It can be enabled or disabled for each individual interface.

To enable or disable BFD on an interface, use the following command:

```
G8272(config-if)# [no] bfd {ipv4|ipv6}
```

Note: BFD cannot be enabled or disabled globally on the switch. It can be enabled or disabled on individual interfaces.

To configure the exchange of BFD control packets during BFD sessions, use the following command:

```
G8272(config)# [no] bfd {ipv4|ipv6} interval <50-999> minrx <50-999> multiplier <3-50>
```

where:

- **interval** - the time interval, in milliseconds, between the transmission of two consecutive BFD control packets (the default value is 50 milliseconds);
- **minrx** - the time interval, in milliseconds, at which BFD control packets are expected to be received (the default value is 50 milliseconds);
- **multiplier** - the number of consecutive BFD control packets that must be not be received before the switch declares its BFD peer unavailable and sends a notification the peer (the default is 3 missed control packets);

You can also configure these parameters on an interface. They can be different than those configured globally on the switch.

```
G8272(config-if)# [no] bfd interval <50-999> minrx <50-999> multiplier <3-50>
```


Notes:

- The configured intervals at which the switch sends and expects BFD control packets are only the desired values. BFD will negotiate with its BFD peer the actual values used during the BFD session when exchanging packets.
- These parameters apply to BFD control packets while BFD echo mode is disabled. If BFD echo mode is enabled, the configured parameters are applied to BFD echo packets instead.

BFD Asynchronous Mode

BFD operates in asynchronous mode, meaning that control packets are periodically exchanged between BFD peers. If a BFD peer sends a number of consecutive control packets to its peer and doesn't receive any reply from the other device, it will declare the BFD session to be down.

BFD Echo Mode

If a switch is configured with BFD Echo mode, it sends its peer a series of echo packets and requests the peer to send back (echo) those hello packets the switch previously sent. If a number of hello packets from the echoed traffic are not received by the switch, it will declare the BFD session as being down.

BFD Echo mode can run independently on each BFD peer.

By default, BFD echo mode is disabled and it can be enabled on each switch interface. To enable or disable BFD echo mode on an interface, use the following command:

```
G8272(config-if)# [no] bfd echo
```

BFD Echo mode uses a timer called slow timer that determines how fast a new BFD session is established. The slow timer also specifies the interval that asynchronous BFD sessions use to exchange control packets, replacing the BFD interval. While control packets are transmitted using the slow timer, echo packets still use the configured BFD interval.

To globally configure the BFD slow timer, in milliseconds, use the following command:

```
G8272(config)# bfd slow-timer <1000-30000>
```

By default, the slow timer is set to 2000 milliseconds (2 seconds). To reset the slow timer to its default value, use the following command:

```
G8272(config)# no bfd slow-timer
```

BFD and BGP

When BFD detects a forwarding failure, it immediately notifies its client protocols.

BFD can work together with BGP to prevent the use of aggressive keep-alive timers. To keep this timer below 10 or 30 seconds, it is recommended to use BFD as well. BFD is compatible with both eBGP and iBGP single or multi hop peers.

BFD Peer Support

BFD provides failure detection in the forwarding path for destinations that are directly connected or are one hop away, or for destinations that are more than one hop away from the switch. The maximum number of hops connecting a switch and its BFD peer is 255.

To add or remove a static single hop BFD peer on an interface, use the following command:

```
G8272(config-if)# [no] bfd neighbor src-ip <source IP address> dest-ip  
<destination IP address>
```

Note: The *IP address* can be either an IPv4 address or an IPv6 address.

To add or remove a static multi hop BFD peer on an interface, use the following command:

```
G8272(config-if)# [no] bfd neighbor src-ip <source IP address> dest-ip  
<destination IP address> multihop
```

Different parameters (interval, minrx and multiplier) can be globally configured for multi hop peers:

```
G8272(config)# [no] bfd multihop-peer <peer IP address> interval <50-999> minrx  
<50-999> multiplier <3-50>
```

BFD Static Routes

BFD can also be enabled on static IP routes. Only one BFD session is created for multiple static routes using the same next-hop address that run through a particular switch interface.

Note: BFD is supported only on single hop static routes, and not on multi hop static routes.

To enable or disable BFD on a static IP route, use the following command:

```
G8272(config)# [no] {ip|ipv6} route static bfd <interface> <BFD peer address>/  
<prefix length>
```

Note: The *BFD peer address* can be either an IPv4 address or an IPv6 address, depending on the type of static route.

For example:

```
G8272(config)# ip route static bfd ethernet 1/12 10.10.10.34/24
```

BFD Authentication

BFD supports the enabling or disabling of authentication for BFD sessions between peers. If authentication doesn't match for both peers, received packets are discarded.

BFD authentication is enabled by specifying an authentication algorithm and key-chain:

- authentication algorithms:
 - simple - plain text password;
 - keyed-md5 - Keyed Message Digest 5 hash algorithm;
 - keyed-sha1 - Keyed Secure Hash Algorithm I;
 - meticulous-keyed-md5 - Meticulous Keyed Message Digest 5 hash algorithm;
 - meticulous-keyed-sha1 - Meticulous Keyed Secure Hash Algorithm I;
- authentication key-chains:
 - key-id - a single key;
 - key-chain - multiple keys;

A key contains the secret data and the time when it becomes valid. The authentication algorithm and key-chain must be configured on both BFD peers and they must be identical. If any mismatch occurs, the BFD session cannot be established.

By default, BFD sessions are established without authentication.

To configure BFD authentication on an interface, use one of the following commands:

- configure authentication to use a single key:

```
G8272(config-if)# [no] bfd [ipv4|ipv6] authentication <authentication algorithm>
key-id <0-255 (key ID)> key <key string>
```

- configure authentication to use a key-chain, containing multiple keys:

```
G8272(config-if)# [no] bfd [ipv4|ipv6] authentication <authentication algorithm>
key-chain <key-chain name>
```

For example:

```
G8272(config-if)# bfd authentication meticulous-keyed-sha1 key-chain kch7
```

Authentication can also be configured for multi hop BFD peers:

```
G8272(config)# [no] bfd multihop-peer <peer IP address> authentication
<authentication algorithm> key-id <0-255 (key ID)> key <key string>
```

```
G8272(config-if)# [no] bfd multihop-peer <peer IP address> authentication
<authentication algorithm> key-chain <key-chain name>
```

Generalized TTL Security Mechanism

The Generalized TTL Security Mechanism (GTSM) provides the switch's IP based control plane with protection from CPU utilization based attacks. It relies on a packet's TTL (for IPv4) or hop limit (for IPv6) to protect the switch from packets sent in a rapid succession.

By default, GTSM is disabled. To enable or disable GTSM for BFD, use the following command:

```
G8272(config)# bfd gtsm {enable|disable}
```

By default, GTSM uses a TTL or hop limit of 255. To change the TTL or hop limit, run the following command:

```
G8272(config)# bfd gtsm ttl <1-255>
```

To reset the TTL or hop limit to its default value of 255, use the following command:

```
G8272(config)# no bfd gtsm ttl <1-255>
```

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 G8272 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.

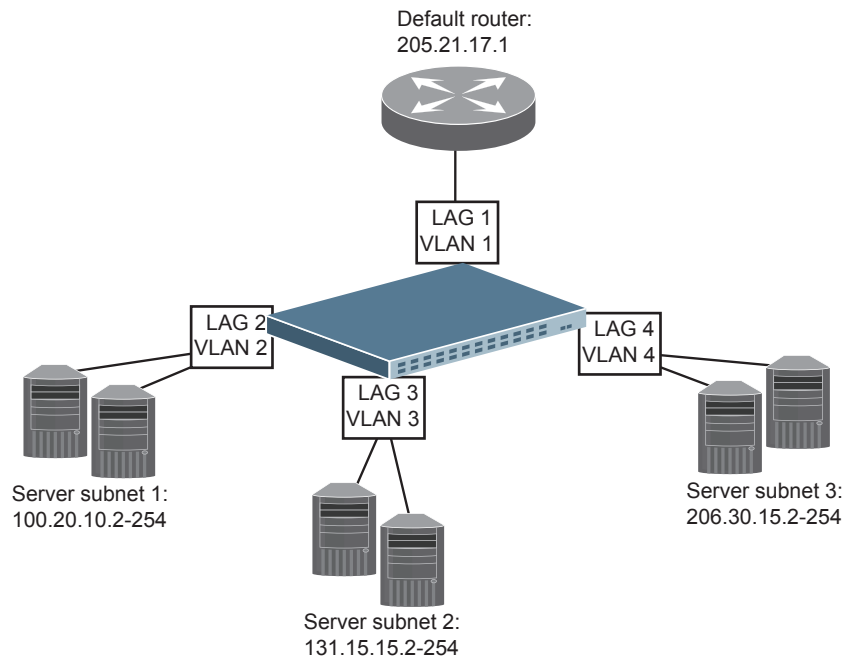
Even if every end-station could be moved to better logical subnets (a daunting task), competition for access to common server pools on different subnets still burdens the routers.

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 G8272 in the following configuration example:

Figure 10. 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 [Figure 10 on page 278](#).

1. Assign an IP address for each router and each server.

The following IP addresses are used:

Table 13. *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 interfaces are needed:

Table 14. *Subnet Routing Example: IP Interface Assignments*

Interface	Devices	IP Interface Address
Port 1	Default router	205.21.17.3
Port 2	Web servers	100.20.10.1
Port 3	Database servers	131.15.15.1
Port 4	Terminal Servers	206.30.15.1

3. Determine which switch ports belong to which VLANs.

The following table adds port and VLAN information:

Table 15. *Subnet Routing Example: Optional VLAN Ports*

Devices	LAG	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 Command Line Interface (CLI) as the administrator.

4. Add the switch ports to their respective VLANs:

```
G8272(config)# vlan 1-4
G8272(config-vlan)# exit

G8272(config)# interface ethernet 1/22
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 1
G8272(config-if)# aggregation-group 1 mode on
G8272(config-if)# exit

G8272(config)# interface ethernet 1/1-2
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 2
G8272(config-if)# aggregation-group 2 mode on
G8272(config-if)# exit

G8272(config)# interface ethernet 1/3-4
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 3
G8272(config-if)# aggregation-group 3 mode on
G8272(config-if)# exit

G8272(config)# interface ethernet 1/5-6
G8272(config-if)# bridge-port mode trunk
G8272(config-if)# bridge-port trunk allowed vlan 4
G8272(config-if)# aggregation-group 4 mode on
G8272(config-if)# exit
```

5. Now that the ports are separated into LAGs, the VLANs are assigned to the appropriate interfaces for each subnet. From [Table 15 on page 279](#), the settings are made as follows:

```
G8272(config)# interface vlan 1
G8272(config-if)# ip address 205.21.17.3/24
G8272(config-if)# exit

G8272(config)# interface vlan 2
G8272(config-if)# ip address 100.20.10.1/24
G8272(config-if)# exit

G8272(config)# interface vlan 3
G8272(config-if)# ip address 131.15.15.1/24
G8272(config-if)# exit

G8272(config)# interface vlan 4
G8272(config-if)# ip address 206.30.15.1/24
G8272(config-if)# exit
```

6. Configure the default gateway to the routers' addresses.

The default gateway allows the switch to send outbound traffic to the router:

```
G8272(config)# ip route 0.0.0.0/0 205.21.17.1
```


7. Verify the configuration.

```
G8272# display vlan
```

VLAN	Name	Status	IPMC FLOOD	Ports
1	default	ACTIVE	IPv4, IPv6	Ethernet1/1(u) Ethernet1/2(u) Ethernet1/6(u) Ethernet1/7(u) Ethernet1/8(u) Ethernet1/9(u) Ethernet1/11(u) Ethernet1/12(u) Ethernet1/13(u) Ethernet1/17(u) Ethernet1/18(u) Ethernet1/19(u) Ethernet1/21(u)

...

```
G8272# display interface brief
```

Ethernet Interface	VLAN	Type	Mode	Status	Reason	Speed	Port Agg#
Ethernet1/1	1	eth	access	down	Link not connected	auto	--
Ethernet1/2	1	eth	access	down	Link not connected	auto	--
Ethernet1/3	19	eth	access	up	none	10000	--
Ethernet1/4	20	eth	access	down	Administratively down	auto	--
Ethernet1/5	--	eth	routed	down	Administratively down	auto	--
Ethernet1/6	1	eth	access	down	Link not connected	auto	--
Ethernet1/7	1	eth	access	down	Link not connected	auto	--
Ethernet1/8	1	eth	access	down	Link not connected	auto	--
Ethernet1/9	1	eth	access	down	Link not connected	auto	--
Ethernet1/10	--	eth	routed	down	Administratively down	auto	--

...

```
G8272# display ip interface brief
```

Interface	IP-Address	Admin-Status	Link-Status
Ethernet1/5	21.1.1.1	administratively down	down
Ethernet1/10	191.1.1.1	administratively down	down
Ethernet1/14	16.1.1.2	up	up
Ethernet1/20	116.0.0.1	up	down
Ethernet1/22	22.1.1.2	administratively down	down
mgmt0	10.241.40.13	up	up
Vlan17	17.1.1.2	up	down
Vlan18	18.1.1.2	up	down
Vlan19	19.1.1.1	up	up
Vlan20	20.1.1.1	up	down
Vlan23	23.1.1.2	up	down
Vlan24	24.1.1.2	up	down

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. 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.

RIB Support for ECMP Routes

RIB supports the adding of multiple next-hops to the same route. Such routes are called Equal Cost Multiple Paths (ECMP) routes.

The maximum number of different next-hops permitted for a route can be changed. To configure the maximum number of different paths for an ECMP route, use the following command:

```
G8272(config)# maximum-paths <1-32>
```

By default, the switch allows up to 32 paths for a single route. To reset this number to its default value, use the following command:

```
G8272(config)# no maximum-paths
```

ECMP Hashing

In case of an ECMP route, the next-hop address will be chosen based on a hash algorithm. The hashing parameters used are:

- source IP address
- destination IP address
- source and destination IP addresses
- Layer 4 source port
- Layer 4 destination port
- Layer 4 source and destination ports

To configure the hashing parameters, use the following command:

```
G8272(config)# ip load-sharing <hashing parameters> [universal-id <1-4294967295>]
```

Note: The universal ID is used to randomize the hashing result for each pair of source and destination addresses.

By default, the hash algorithm utilizes the all four hashing parameters and has a universal ID of 1431655765. To reset the hash algorithm to its default settings, use the following command:

```
G8272(config)# no ip load-sharing
```

To view the current hash settings, use the following command:

```
G8272# display ip load-sharing
```

When ECMP is configured in a multi-tier network, where multiple redundant paths are present, the hash algorithm on neighboring switches, that have the same hashing parameters, may calculate the same path for outgoing traffic. This results in a congestion of traffic on a upstream switch, while leaving other ECMP routes completely unused, thus negating the redundancy of the network.

To avoid such scenarios, a hash offset can be introduced in the hash algorithm calculations. The hash offset is just a number that is taken into account when calculating the hashing result. By default, the hash offset is set to 0, therefore having no influence on the hash calculation. When it is set to a non-zero value, the hash offset ensures that a different ECMP route is chosen.

Use the following command to set an ECMP hash offset:

```
G8272(config)# hardware ecmp hash-offset <0-15>
```

To reset it to its default value, use the following command:

```
G8272(config)# no hardware ecmp hash-offset
```

Note: ECMP hashing isn't available for IPv6 traffic.

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, prefix length and gateway).

```
G8272(config)# ip route 10.10.1.1/32 100.10.1.1
```

2. Continue adding static routes with the same IP address and prefix length, but a different gateway address.

```
G8272(config)# ip route 10.10.1.1/32 200.20.2.2  
...
```

3. Select an ECMP hashing method (optional).

```
G8272(config)# ip load-sharing source-dest-ip
```

You may add up to 32 gateways for each static route.

4. Use the following command to check the status of ECMP static routes:

```
G8272(config)# display ip route static
```

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 statically 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.

The G8272 support DHCP for both IPv4 and IPv6 addressing.

To enable or disable DHCP, use the following command:

```
G8272(config)# [no] feature dhcp
```

By default, DHCP is enabled on the management port, but disabled on all other switch Layer 3 interfaces.

For more details about DHCP, see [“DHCP IP Address Services”](#) on page 35.

Internet Control Message Protocol

The Internet Control Message Protocol (ICMP) is used to send error messages indicating, for example, that a network device is unreachable. ICMP messages are primarily used for diagnostic and control purposes.

ICMP errors are sent back to the source device of the IP packet. For example, when a packet crosses a network node, its TTL decreases by one. When the TTL reaches zero, the packet is discarded and an ICMP error message is sent to source device.

ICMP messages are used by many network utilities, such as Traceroute or Ping.

ICMP messages are identified by type and code, that is used to identify the type of message. Below are a few of the most common ICMP message types:

- 0 - echo reply
An echo reply is generated when the switch receives an echo request (ping) and it indicates that the switch can be reached.
- 3 - destination unreachable
It means that there is error along the path to the destination device. This type has multiple codes that indicate the cause.
- 5 - redirect
The switch detects that the packet isn't optimally forwarded to its destination.
- 11 - time exceeded
The TTL of the packet has reached zero.

ICMP can be configured independently on each switch interface.

To view the current settings regarding ICMP, use the following command:

```
G8272# display ip interface

IP Interface Status for VRF (default)
  loopback0, Interface Status: link up/admin up
  IP MTU:1500 bytes (using link MTU)
  IP icmp redirects: enabled
  IP icmp unreachable (except port): disabled
  IP icmp port-unreachable: enabled

  Vlan1, Interface Status: link up/admin up
  IP address: 205.21.17.1, IP subnet: 205.21.17.0/24
  IP MTU:1500 bytes (using link MTU)
  IP icmp redirects: enabled
  IP icmp unreachable (except port): disabled
  IP icmp port-unreachable: enabled
  ...
```

ICMP Redirects

The switch generates an ICMP redirect message when it detects that a packet isn't optimally forwarded to the destination and sends the message to inform the source device to send future packets along the optimal path across the network. This kind of ICMP messages are identified as type 5.

For example, the switch, a host device and a neighboring router are all directly connected with each other. The switch is configured as the default gateway for the host. The host sends the switch a packet with a destination that is reachable only through the router. The switch sends host an ICMP redirect error and informs the host to send subsequent packets directly to the router, thus reducing the number of hops to the destination by one.

By default, ICMP redirect errors are enabled on all interfaces.

To enable or disable ICMP redirect error messages, use the following command:

```
G8272(config-if)# [no] ip redirects
```

ICMP Port Unreachable

The switch generates an ICMP port unreachable message when the interface of destination device is unavailable and sends the message to inform the source device of this. This kind of ICMP messages are identified as type 3, having error code 3.

By default, ICMP port unreachable errors are enabled on all interfaces.

To enable or disable ICMP port unreachable error messages, use the following command:

```
G8272(config-if)# [no] ip port-unreachable
```

ICMP Unreachable (except Port)

The switch generates an ICMP unreachable message when the destination device isn't available. This kind of ICMP messages are identified as type 3 and have multiple error codes, ranging from 0 to 15. For example, an error code 1 means that the destination device couldn't be reached, while an error code 2 means that the protocol for which the packet was intended isn't available, but the destination device is reachable.

By default, ICMP unreachable errors, except port unreachable messages, are disabled on all interfaces.

To enable or disable ICMP unreachable error messages (except port unreachable), use the following command:

```
G8272(config-if)# [no] ip unreachable
```

Chapter 15. Routed Ports

By default, all ports on the Lenovo RackSwitch G8272 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, G8272 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.

This section discusses the following topics:

- [“Routed Ports Overview” on page 288](#)
- [“Configuring a Routed Port” on page 290](#)

Routed Ports Overview

When a switch port is configured as a routed port, it forwards Layer 3 traffic and no longer performs Layer 2 switching functions.

By default, all ethernet ports are configured as switch access ports. To configure the port to operate as a routed port, see [“Configuring a Routed Port” on page 290](#).

You can also assign an IP address to a routed port and configure OSPF to route IP traffic through the interface.

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 the Spanning Tree Protocol (STP).
- Is always in a forwarding state.
- Can participate in IPv4 or IPv6 routing.
- Can be configured with basic IP protocols, such as Internet Control Message Protocol (ICMP) and with Layer 3 protocols, such as Open Shortest Path First (OSPF) or Virtual Router Redundancy Protocol (VRRP).
- Layer 3 applications can be enabled, such as Network Time Protocol (NTP) or Dynamic Host Configuration Protocol (DHCP).
- Layer 3 configuration is saved even when the interface is shutdown.
- MAC address learning is always enabled.
- Native VLAN 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.

Notes:

- Ports on which Link Aggregation Control Protocol (LACP) is enabled or are part of a Link Aggregation Group (LAG) cannot be changed to routed ports.
- Ports with configured static MAC addresses 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 the first available internal VLAN on which flooding is disabled. The ID of this internal VLAN will be between 4000 - 4093, starting in descending order from 4093. The internal VLAN is assigned to the Common Internal Spanning Tree (CIST). You cannot change the VLAN number assigned to the routed port.
- STP is disabled and the port is set to a forwarding state.
- All the Layer 2 configuration is lost.
- The port will be deleted from the bridge.
- MAC Access Control Lists (ACL) from the Layer 2 port will be removed.

When a routed port is changed back to a switch port, the following changes take place:

- All relevant Layer 3 configuration is lost. Common Layer 3 and Layer 2 settings will be preserved during the transition.
- The ARP entry corresponding to the IP address is lost.
- The switch port is added to the default VLAN (VLAN 1).
- STP is turned on and the port is added to the default STG (STG 1).
- The switch port can participate in STG and VLAN flooding.
- The switch port can participate in bridging.
- LACP port attributes are set to default.
- Multicast threshold remains disabled.
- BPDU guard remains disabled.
- QoS configurations is lost.
- IP Access Control Lists (ACL) are removed.

Note: When you configure a routed port 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

Use only the ISCLI to configure routed ports. Configurations made using SNMP cannot be saved or applied.

You cannot configure a management interface to be a routed port.

Following are the basic steps for configuring a routed port:

1. Enter the interface configuration mode for the port (for this example, ethernet interface 1/12 is used).

```
G8272(config)# interface ethernet 1/12
```

2. Enable routing.

```
G8272(config-if)# no bridge-port
```

3. Assign an IP address.

a. an IPv4 address:

```
G8272(config-if)# ip address <IPv4 address>/<prefix length>
```

b. an IPv6 address:

```
G8272(config-if)# ipv6 address <IPv6 address>/<prefix length>
```

4. Optionally, make sure the interface is in uplink state:

```
G8272(config-if)# no shutdown
```

5. Optionally, you can set the maximum transmission unit (MTU) size in bytes for received or sent frames for the routed port:

```
G8272(config-if)# mtu <64-9216>
```

Note: The default value is 1500 bytes.

6. Check the interface configuration:

```
G8272(config-if)# display interface ethernet 1/12

Interface Ethernet1/12
Hardware is Ethernet Current HW addr: a897.dcde.2501
Physical:a897.dcde.250e Logical:(not set)
index 11 metric 1 MTU 1500 Bandwidth 10000000 Kbit
no bridge-port
arp ageing timeout 1500
<UP,BROADCAST,MULTICAST>
VRF Binding: Not bound
Speed 10000 Mb/s Duplex full
DHCP client is disabled.
...
```

Configuring OSPF on Routed Ports

The following OSPF configuration commands are supported on routed ports:

```
G8272(config-if)# ip ospf ?

A.B.C.D          Address of interface
authentication    Enable authentication
authentication-key Authentication password (key)
bfd              Bidirectional Forwarding Detection (BFD)
cost             Interface cost
database-filter   Filter OSPF LSA during synchronization and
                 flooding
dead-interval    Interval after which a neighbor is declared dead
hello-interval   Time between HELLO packets
message-digest-key Message digest authentication password (key)
mtu              OSPF interface MTU
mtu-ignore       Ignores the MTU in DBD packets
network          Network type
passive-interface Suppress routing updates on an interface or on all
                 interfaces
priority         Router priority
retransmit-interval Time between retransmitting lost link state
                 advertisements
shutdown         Shutdown OSPF
transmit-delay   Link state transmit delay

G8272(config-if)# ip router ospf 0 ?

area          Set the OSPF area ID
multi-area    Set the multi-area-adjacency
```

See [Chapter 20, "Open Shortest Path First,"](#) for details on the OSPF protocol and its configuration.

For a full description of the above OSPF commands, see the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

Note: OSPFv3 cannot be configured on routed ports.

OSPF Configuration Example

The following example includes the basic steps for configuring OSPF on a routed port:

1. Enable OSPF on the switch and configure the OSPF router ID:

```
G8272(config)# routing-protocol ospf
G8272(config-router)# router-id <IPv4 address>
G8272(config-router)# exit
G8272(config)#
```

2. Make the routed port part of OSPF area 0:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# ip router ospf 0 area 0
```

Chapter 16. Address Resolution Protocol

The Address Resolution Protocol (ARP) is a protocol that maps an IPv4 address to a MAC address. ARP uses a request and reply mechanism to determine the MAC address of the destination host.

This section discusses the following topics:

- [“ARP Overview” on page 294](#)
- [“ARP Aging Timer” on page 295](#)
- [“ARP Inspection” on page 296](#)
- [“Static ARP Entries” on page 297](#)
- [“ARP Entry States” on page 298](#)
- [“ARP Table Refresh” on page 299](#)

ARP Overview

When an IP packet is sent across the local network, from one device to another, the source host must know the MAC address of the destination host. If the address is known, the source device encapsulates the IP packet in a layer 2 frame containing the destination MAC address and then transmits the packet over the LAN. If the MAC address is unknown, the source device uses ARP to determine it.

The protocol sends an ARP broadcast message across the local network, requesting the MAC address of the device with the specified destination IP address. When the owner of the IP address receives the ARP broadcast, it will send the requesting device an ARP unicast message containing its MAC address. The source device will store the received MAC address in its ARP table and start sending traffic. The source MAC address is also kept in the ARP table of the destination device.

MAC addresses learned using ARP are kept in an ARP table, which is independent of the Forwarding Database (FDB). The ARP tables consists of MAC addresses and their associated IP addresses. The ARP table can contain up to a maximum of 48,000 entries.

To view the contents of the ARP table, use the following command:

```
G8272> display ip arp
```

ARP only operates in the boundaries of a single network and it is never routed across network nodes.

Note: ARP only works in IPv4 local networks. In IPv6 environments, the functionality of ARP is provided by the Neighbor Discovery Protocol (NDP). For more details, see [“Neighbor Discovery” on page 306](#).

Note: ARP can only be configured on routed ports or switch virtual interfaces (SVI). When a routed port using ARP changes to a switch port, any ARP entries (dynamic or static) associated with that port are removed from the ARP table.

ARP Aging Timer

Because the ARP table has a limited size, ARP periodically removes unused entries by assigning each entry an aging timer. When the timer expires, the entry is refreshed by sending an ARP request. If ARP doesn't receive any replies for three consecutive requests, the entry is removed from the ARP table.

The aging timer of an ARP entry is reset if the switch receives an IP packet from the device mapped to that entry before the timer expires.

To configure the aging time of ARP entries, in seconds, use the following command:

```
G8272(config)# ip arp timeout <60-28800>
```

By default, the aging time of ARP entries is 1500 seconds (25 minutes). To reset the aging time to its default value, use the following command:

```
G8272(config)# no ip arp timeout
```

Note: The aging time of ARP entries can be also individually configured for each routed port or switch virtual interface (SVI).

Note: Static ARP entries are permanent and aren't affected by aging.

To view the aging time of ARP entries, use the following command:

```
G8272> display ip arp
```

ARP Inspection

Access Control Lists (ACLs) can be configured to filter ARP packets received on VLANs. An ARP ACL is used to deny or permit ARP requests or replies based on various fields from the ARP packet. For more details on how to configure an ARP ACL, see [Chapter 6, “Access Control Lists”](#).

To enable or disable ARP inspection, use the following command:

```
G8272(config)# [no] ip arp inspection filter <ARP ACL name> vlan <VLAN number>
```

For example:

```
G8272(config)# ip arp inspection filter arp-acl-22 vlan 100-105
```

To view the configured ARP filters, use the following command:

```
G8272> display ip arp inspection
```

Static ARP Entries

Up to a maximum of 512 static ARP entries can be configured on the switch.

Static ARP entries don't have an aging timer associated with them. They are permanently kept in the ARP table until they are manually removed. A static ARP entry will overwrite any dynamic entries associated with the specified IP address.

Static ARP entries are configured individually on each routed port of the switch.

To add a static ARP entry, use the following command:

```
G8272(config-if)# ip arp <IPv4 address> <MAC address>
```

To remove a static ARP entry, use the following command:

```
G8272(config-if)# no ip arp <IPv4 address>
```

Static ARP Configuration Example

To add a static ARP entry, use the following steps:

1. Configure the ethernet interface as a routed port (for this example, ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12  
G8272(config-if)# no bridge-port
```

2. Configure the IPv4 address of the routed port:

```
G8272(config-if)# ip address <IPv4 address>/<prefix length>
```

For example:

```
G8272(config-if)# ip address 10.0.0.1/24
```

3. Configure the static ARP entry:

```
G8272(config-if)# ip arp <IPv4 address> <MAC address>
```

Note: The IP address used for the static ARP entry must be part of the same network as the routed port.

For example:

```
G8272(config-if)# ip arp 10.0.0.93 0080.c8e8.1efc  
G8272(config-if)# exit
```

4. Verify the ARP table:

```
G8272(config)# exit  
G8272# display ip arp static
```

ARP Entry States

ARP assigns a state to each entry in the ARP table. Depending on the state of each entry, ARP will take certain actions.

The following list describes each state and the actions taken by ARP:

- NOARP - is used for neighbors that don't require ARP, like interfaces connect using Point-to-Point Protocol (PPP).
- PERMANENT - is used for static ARP entries. ARP entries with this state aren't affected by aging and are not removed from the ARP table, unless they are manually deleted.
- INCOMPLETE - is used for unmapped IP addresses. An ARP request has been sent and the protocol is waiting for an ARP reply, containing the MAC address that needs to be assigned to the incomplete entry.
- REACHABLE - is used for mapped IP addresses. When an ARP reply is received, the MAC address is associated with the destination IP address and the ARP entry is now complete. The aging timer is started for that entry. If an IP packet is received from a device mapped to an ARP entry before its aging timer expires, the timer is reset.
- STALE - is used when the base reachable timer for an ARP entry has expired. Although the entry is still present in the ARP table, its reachability has not been confirmed for the duration of the timer. The next time a packet is sent to the owner of the MAC address associated with a stale ARP entry, the reachability verification process will be started.
- DELAY - is used when an IP packet needs to be sent to a device mapped to a stale ARP entry. This is temporary, until the switch sends an ARP request.
- PROBE - is used when the protocol rechecks an expired entry. An ARP request is sent to determine the MAC address of the destination IP address.
- FAILED - is used for ARP entries that couldn't be rechecked. If the switch doesn't receive any ARP replies after three consecutive requests, the entry is marked as failed and it is removed when the ARP table is refreshed.

To view the state of ARP entries, use the following command:

```
G8272> display ip arp
```

ARP Table Refresh

Each ARP entry is checked periodically to determine its state. Based on the entry's state, ARP undertakes certain actions, like refreshing the entry or removing it from the ARP table.

An ARP table refresh can be manually triggered. The switch will recheck each dynamic ARP entry by placing it in the probe state and sending an ARP request. If the switch doesn't receive any ARP replies after three consecutive requests, it places the ARP entry in the incomplete state. If the switch receives an ARP reply, the entry is placed in the reachable state. After all the ARP entries are checked, the switch removes the incomplete entries from the ARP table.

To manually trigger a refresh of the ARP table, use the following command:

```
G8272# remove ip arp
```

To delete all dynamic entries from the ARP table, use the following command:

```
G8272# remove ip arp force-delete
```

Chapter 17. 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 2460
- RFC 4291
- RFC 4429
- RFC 4861
- RFC 4862
- RFC 4443

The following topics are discussed in this section:

- ["IPv6 Address Format" on page 302](#)
- ["IPv6 Address Types" on page 303](#)
- ["IPv6 Interfaces" on page 305](#)
- ["Neighbor Discovery" on page 306](#)
- ["Supported Applications" on page 308](#)
- ["Configuration Guidelines" on page 309](#)
- ["IPv6 Configuration Examples" on page 310](#)
- ["IPv6 Limitations" on page 311](#)

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 Cloud NOS 10.1 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 Interfaces

Each IPv6 interface supports multiple IPv6 addresses. You can manually configure up to 1024 IPv6 addresses for each interface.

You can manually configure up to 1024 IPv6 addresses for each interface, as follows:

- Initial IPv6 address is a global unicast or anycast address.

```
G8272(config)# interface {<interface name>|ethernet <slot/chassis number>
|loopback <interface number>|mgmt <interface number>|port-aggregation <LAG number>
|vlan <vlan number>}
G8272(config-if)# ipv6 address <IPv6 address>
```

- Second IPv6 address can be a unicast or anycast address.

```
G8272(config-ip-if)# ipv6 address <IPv6 address> secondary
G8272(config-ip-if)# exit
```

Each IPv6 address can belong to only one VLAN interface. Each VLAN interface can support multiple IPv4/IPv6 addresses.

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.

You can configure up to 256 IPv6 static neighbors for each interface.

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:

```
G8272(config)# interface ethernet <chassis number/port number>
G8272(config)# interface vlan <VLAN number>
G8272(config-if)# [no] ipv6 nd <arguments>
G8272(config-if)# exit
```

To add or remove entries in the static neighbor cache, use the following command:

```
G8272(config-if)# [no] ipv6 neighbor <IPv6 address>
```

To view the neighbor discovery information for the specified interface, use the following command:

```
G8272(config)# display ipv6 nd interface <name>
```

To clear the neighbor cache, use the following command:

```
G8272# remove ipv6 neighbor
```

Router

Each IPv6 interface can be configured as a router node, as follows:

- A router node's IP address is configured manually. Router nodes can send Router Advertisements.

Note: When IP forwarding is turned on, all IPv6 interfaces configured on the switch can forward packets.

Supported Applications

The following applications have been enhanced to provide IPv6 support.

- **Ping**

The **ping6** command supports IPv6 addresses. Use the following format to ping an IPv6 address:

```
ping6 <IPv6 address> [vrf {default|management}] [interface <destination interface>] [source <IPv6 source address>] [count <number of pings>] [interval <delay time between packets>] [packet-size <length>] [timeout <interval>]
```

- **Traceroute**

The **traceroute6** command supports IPv6 addresses (but not link-local addresses). Use the following format to perform a traceroute to an IPv6 address:

```
traceroute6 <IPv6 address> [vrf {default|management}] [interface <destination interface>] [source <IPv6 source>]
```

- **Telnet**

The **telnet6** command supports IPv6 addresses (but not link-local addresses). Use the following format to Telnet to IPv6 address:

```
telnet6 <IPv6 address> [vrf {default|management}] [port <port number>]
```

- **SSH**

Secure Shell (SSH) connections over IPv6 are supported (but not link-local addresses). The following syntax is required from the client:

```
ssh6 <IPv6 address> [vrf {default|management}] [port <port number>]
```

- **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.

Configuration Guidelines

When you configure an interface for IPv6, consider the following guidelines:

- A single interface can accept multiple IPv6 addresses.
- Reserved IPv6 addresses are not supported on the interface (e.g. 1::1/64).
- 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

Use the following example to configure IPv6 neighbor discovery prefix settings on the router.

1. Enable IPv6 on an interface.

```
G8272(config)# interface ethernet 1/2
G8272(config-if)# no bridge-port
G8272(config-if)# ipv6 address 12ef:4533::0f30/64
```

Neighbor discovery is automatically enabled.

2. Advertise the IPv6 prefix in the router-advertisement messages.

```
G8272(config-if)# ipv6 nd prefix
```

3. Verify the interface configuration.

```
G8272(config-if)# display ipv6 nd interface
```

IPv6 Example 2

Use the following example to configure IPv6 neighbor discovery reachable time.

1. Enable IPv6 on an interface.

```
G8272(config)# interface ethernet 1/2
G8272(config-if)# no bridge-port
G8272(config-if)# ipv6 address ee33:a239:22::bcd3/32
```

Neighbor discovery is automatically enabled.

2. Configure the time when a node considers a neighbor to be up.

```
G8272(config-if)# ipv6 nd reachable-time 10
```

3. Verify the configuration.

```
G8272(config-if)# display ipv6 nd interface
```

IPv6 Limitations

Most other Cloud Network OS 10.1 features permit IP addresses to be configured using either IPv4 or IPv6 address formats. However, the following switch features support IPv4 only:

- Open Shortest Path First (OSPF) version 2.

Chapter 18. Internet Group Management Protocol

Internet Group Management Protocol (IGMP) is used by IPv4 Multicast routers to learn about the existence of host group members on their directly attached subnet. The IPv4 multicast routers 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 G8272 can perform IGMP Snooping and supports both static and dynamic IGMP groups and multicast routers. The G8272 can act as a Querier and participate in the IGMP Querier election process.

The following topics are discussed in this chapter:

- [“IGMP Terms” on page 314](#)
- [“How IGMP Works” on page 315](#)
- [“IGMP Capacity and Default Values” on page 316](#)
- [“IGMP Snooping” on page 317](#)
- [“IGMP Snooping Configuration Example” on page 325](#)
- [“Additional IGMP Features” on page 335](#)

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.
- 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.
- Fast Leave: 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 a specific multicast group.

How IGMP Works

When IGMP is not configured, switches forward multicast traffic through all ports, increasing network load. When IGMP is configured on a switch, multicast traffic flows as follows:

- A server sends multicast traffic to a multicast group.
- The multicast router 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 multicast router.
- The switch forwards multicast traffic only to hosts that have joined a group and to the multicast router.
- The multicast router 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.
- An IGMPv2 host can initiate the Leave process by sending an IGMPv2 Leave packet to the IGMP Snooper.
- When a host sends an IGMP Leave packet, the IGMP Snooper sends *Group-specific 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 multicast router.

The G8272 supports the following:

- IGMP version 1, 2 and 3;
- 64 static multicast routers;
- 64 dynamic multicast routers;
- 128 static IGMP groups;
- 3072 dynamic IGMP groups.

Note: Unknown multicast traffic is sent to all ports if the IPMC 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 are learned on a switch.

To enable or disable IP multicast (IPMC) flood, use the following commands:

```
G8272(config)# vlan <VLAN number>
G8272(config-vlan)# [no] flood [ipv4|ipv6]
```

For more details, see [“IPMC Flooding” on page 173](#).

IGMP Capacity and Default Values

The following table lists the maximum and minimum values of the G8272 variables.

Table 16. *G8272 Capacity Table*

Variable	Maximum
VLANs - Snooping	3999
Static multicast routers	64
Dynamic multicast routers	64
Static Group - Snooping	128
Dynamic Groups - Snooping	3072

The following table lists the default settings for IGMP features and variables.

Table 17. *IGMP Default Configuration Settings*

Field	Default Value
IGMP Snooping	Enabled for all VLANs
IGMP Snooping Version	3
IPMC Flood	Enabled for all VLANs
IGMP FastLeave	Disabled for all VLANs
IGMP multicast router Timeout	255 Seconds
IGMP Group Timeout	260 Seconds
IGMP Report Suppression	Enabled for all VLANs
IGMP Query-Interval Variable	125 Seconds
IGMP Query-Max-Response-Time Variable	10 Seconds
IGMP Last-Member-Query-Interval Variable	1 Second
IGMP Robustness Variable	2

Note: The IGMP multicast router Timeout is calculated as:

$$robustness \times query\ interval + \frac{query\ max\ response\ time}{2} = 2 \times 125 + \frac{10}{2} = 255\ seconds$$

Note: The IGMP Group Timeout is calculated as:

$$robustness \times query\ interval + query\ max\ response\ time = 2 \times 125 + 10 = 260\ seconds$$

IGMP Snooping

IGMP Snooping allows a switch to listen to the IGMP conversation between hosts and multicast routers. With IGMP Snooping enabled, the switch learns the ports interested in receiving multicast data and forwards it only to those ports. Thus, 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 multicast router. 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.

If the IGMP Snooping switch receives a Query, it forwards the Query to all interfaces members of the VLAN on which it was received. Any Reports sent in reply to that Query are forwarded only on interfaces with multicast routers present.

By default, IGMP Snooping is globally enabled on the switch. To globally enable or disable IGMP Snooping, use the following command:

```
G8272(config)# [no] ip igmp snooping
```

IGMP Snooping can also be enabled or disabled for each VLAN. To enable or disable IGMP Snooping on a VLAN, use the following command:

```
G8272(config)# vlan <VLAN number>  
G8272(config-vlan)# [no] ip igmp snooping
```

Note: If IGMP Snooping is disabled globally, it will also be disabled on any VLAN, regardless of the VLAN settings.

IGMP Snooping can monitor all version and types of IGMP messages (v1, v2 and v3). If the IGMP is configured to version 2, IGMPv3 packets will be ignored by the switch.

To change the IGMP version used on a VLAN, use the following command:

```
G8272(config)# vlan <VLAN number>  
G8272(config-vlan)# ip igmp snooping version {2|3}
```

The default IGMP version is 3. You can reset the IGMP version used by a VLAN using the following command:

```
G8272(config-vlan)# no ip igmp snooping version
```

To view the current IGMP Snooping settings, use the following command:

```
G8272> display ip igmp snooping
```

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 (RFC 3376). The registrations of other hosts for the same group on the same interface are not changed.

The G8272 supports the following IGMPv3 filter modes:

- **INCLUDE mode:** The host requests membership to a multicast group and provides 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.

IGMPv3 Snooping is compatible with IGMPv1 and IGMPv2 Snooping.

Spanning Tree Topology Change

If a Spanning Tree topology changes happens on an interface, the switch will receive a Topology Change Notification (TCN) message. After 30 seconds from the moment the TCN message is received, the IGMP Snooping switch will delete all dynamically learned IGMP multicast groups and multicast routers associated with that interface. After deleting the IGMP entries, the switch sends General Queries on all interface/VLAN pairs belonging to the Spanning Tree Group (STG) specified in the TCN message. The reports received in reply to the General Queries will enable the switch to re-learn IGMP multicast groups.

This mechanism happens even if the switch is not elected as an IGMP Querier. In this case, the source IPv4 address of the General Queries messages will be 0.0.0.0.

This mechanism can only be configured globally for the switch and it cannot be enabled or disabled only on a certain interface.

By default, this mechanism is enabled on the switch. To enable or disable it, use the following command:

```
G8272(config)# [no] ip igmp snooping tcn flood
```

IGMP Querier

For IGMP Snooping to function, you must have a multicast router on a VLAN to generate IGMP Membership 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 periodic IGMP Query packets for the VLAN. Hosts that want to receive multicast traffic will respond to the Membership Query packets with IGMP Report messages, which are used by the Querier to establish an appropriate forwarding list.

Querier Election

If multiple multicast routers exist on the VLAN, only one can be elected as a Querier. The multicast routers elect the one with the lowest source IPv4 address as the Querier. The Querier performs all periodic Membership Queries. All other multicast routers (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 a multicast router with a lower IP address, it transitions to a non-Querier state and initializes an *other querier present timer*. When this timer expires, the multicast router transitions back to the Querier state and sends a General Query packet.

While the Querier switch is in a non-Querier state, only Query packets that have a lower IP address than the elected Querier switch will be processed. This will cause an update of the Querier information.

By default, IGMP Snooping Querier is disabled on all VLANs.

Follow the steps below for a basic IGMP Querier configuration.

1. Configure the source IPv4 address for IGMP Querier on a VLAN.

```
G8272(config)# vlan <VLAN number>
G8272(config-vlan)# ip igmp snooping querier <Querier IP address>
```

2. Verify the configuration.

```
G8272> display ip igmp snooping querier
```

To disable IGMP Querier for a VLAN, use the following command:

```
G8272(config-vlan)# no ip igmp snooping querier
```

When IGMP Querier is enabled on a VLAN, the switch will send a number of consecutive General Query packets, waiting a certain period of time between each General Query. The number of General Queries is called Startup Query Count and time interval between each packet is called Startup Query Interval.

Note: The first General Query is sent with a random delay between one and five seconds.

To configure the Startup Query Count, use the following command:

```
G8272(config)# vlan <VLAN number>  
G8272(config-vlan)# ip igmp snooping startup-query-count <1-10>
```

The default value is the same as the configured robustness value, which by default is 2 (for more information about robustness, see [“Robustness Variable” on page 335](#)). You can reset the Startup Query Count to this value by using the following command:

```
G8272(config-vlan)# no ip igmp snooping startup-query-count
```

To configure the Startup Query Interval (in seconds), use the following command:

```
G8272(config-vlan)# ip igmp snooping startup-query-interval <1-18000>
```

The default value is a quarter of the Query Interval (which by default is 125 seconds, resulting in a default value of 31 seconds for the Startup Query Interval). You can reset the Startup Query Interval to this value by using the following command:

```
G8272(config-vlan)# no ip igmp snooping startup-query-interval
```

After the switch sends out the number of General Query packets as configured by the Startup Query Count and Interval, it will use Query Interval for sending further General Query packets if it is elected as Querier.

To configure the Query Interval (in seconds), use the following command:

```
G8272(config-vlan)# ip igmp snooping query-interval <1-18000>
```

Note: When setting values for the Query Interval that are higher than 128 seconds, some of these values will be automatically adjusted by the switch to be consistent with the value used within the generated query packets. When this occurs, the switch will display a logging message.

The default value is 125 seconds. You can reset the Query Interval to this value by using the following command:

```
G8272(config-vlan)# no ip igmp snooping query-interval
```

Note: This command is not the same as configuring the Query Interval to 125 seconds. For IGMPv3, the Non-Querier switches get their Robustness and Query Interval values from the elected Querier switch. When the Query Interval is configured on a Non-Querier switch, the specified value is enforced regardless if it is the same as the default value or not (in this case, the default value is the Query Interval configured on the elected Querier switch). To configure the Non-Querier switch to take on the same value as the one configured on the elected Querier, the above command must be used.

The Querier Timeout is a time interval used by Non-Querier switches. After the IGMP Querier election process is over, if a Non-Querier switch doesn't receive a Query during a time interval equal to twice the configured Querier Timeout, it will trigger a new IGMP Querier election.

To configure the Querier Timeout (in seconds), use the following command:

```
G8272(config-vlan)# ip igmp snooping querier-timeout <1-65535>
```

The default value is 255 seconds and you can reset the Querier Timeout to this value by using the following command:

```
G8272(config-vlan)# no ip igmp snooping querier-timeout
```

Multicast Router Discovery

For each VLAN, IGMP Snooping keeps track of switch ports that are connected to Multicast Routers (multicast routers) by checking IGMP queries sent by multicast routers.

An IGMP Snooping switch forwards multicast traffic and IGMP control messages (report and leave messages) by using information about the presence of multicast routers on switch ports.

You can also configure a static multicast router on a specific ethernet port or Link Aggregation Group (LAG). For more details, see [“Static Multicast Router” on page 337](#).

When a IGMP query packet with 0.0.0.0 as the source IPv4 address is received, the switch learns a multicast router on the receiving interface only if IGMP Snooping Querier is disabled.

If IGMP Querier is enabled, a multicast router is learned when receiving an IGMP query packet only if both of the following conditions are met:

- the source IP address is not 0.0.0.0;
- the query’s source IP address is strictly lower than the Querier’s IP address.

Note: multicast router cannot be discovered using Protocol Independent Multicast (PIM) Hello messages.

IGMP Query Messages

An IGMP Snooping switch periodically receives General Query messages from the elected IGMP Querier on a VLAN and learns that the switch port on which such queries are received is a multicast router port. The switch forwards IGMP General Query messages to all interfaces that belong to the VLAN on which the messages have been received.

The IGMP Snooping switch keeps track of the Querier version and other parameters, like the maximum response time. This information is gathered from IGMP General Query messages. After a Spanning Tree topology change happens, the switch will send General Queries using the same IGMP version as the one used by the Querier.

When a host sends an IGMP Report indicating its interest in receiving multicast traffic, the IGMP Snooping switch forwards the Report message only using ports that have a multicast router.

To configure the maximum response time (in seconds) during which a switch can reply after received IGMP Query message, use the following command:

```
G8272(config)# vlan <VLAN number>  
G8272(config-vlan)# ip igmp snooping query-max-response-time <1-25>
```

The default value is 10 seconds and you can reset the maximum response time to that value by using the following command:

```
G8272(config-vlan)# no ip igmp snooping query-max-response-time
```

IGMP Groups

One IGMP entry is allocated for each unique join multicast traffic 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.

The IGMP Snooping switch keeps track of group membership with group records as: (Group, group-timer, {Filter - Include/Exclude}, {Source-records}). For IGMPv1 and IGMPv2 group membership records, the filter and source records are set to the default values. The group membership state is updated when receiving IGMPv1, IGMPv2 and IGMPv3 Report messages that force the compatibility between different IGMP versions.

You can add or remove a static member of a multicast group for an ethernet port or LAG belonging to a VLAN (for this example, ethernet interface 1/12 is used):

```
G8272(config)# vlan <VLAN number>
G8272(config-vlan)# [no] ip igmp snooping static-group <multicast group IP
address> interface ethernet 1/12
```

You can remove all static multicast groups by using the following command:

```
G8272(config)# no ip igmp snooping static-group all
```

The IGMP Snooping switch uses group membership state information to configure the multicast forwarding interfaces.

Static and dynamic IGMP entries for the same or different multicast groups can be present on an interface at the same time. When a group is present, both statically and dynamically, on the same interface, only one IGMP entry will be shown in the group table with both the static (S) and dynamic (D) flags. A static multicast group will not interfere with the dynamic learning of the same group on a different interface or in the same or different VLAN.

To view all IGMP multicast groups, use the following command:

```
G8272> display ip igmp snooping groups
```

You can delete all dynamic entries from the multicast group table of an IGMP Snooping switch by using the following command:

```
G8272(config)# remove ip igmp snooping group *
```

Note: When clearing an IGMPv2 multicast group, the switch automatically generates Leave messages for those groups and sends the messages to all interface with multicast routers present.

You can delete only a certain multicast group by using the following command:

```
G8272# remove ip igmp snooping group <multicast group IPv4 address>
```

You can also delete all dynamic IGMP multicast group entries associated with a VLAN by using the following command:

```
G8272# remove ip igmp snooping group vlan <VLAN number>
```

IGMP Snooping Configuration Guidelines

Consider the following guidelines when you configure IGMP Snooping:

- When multicast traffic flood is disabled, the multicast traffic sent by the multicast server is discarded if no hosts are learned on the switch.
- The multicast router periodically sends IGMP Queries.
- The switch learns the multicast router on the interface connected to the router when it sees Query messages. The switch then floods the IGMP queries on all other interfaces, including Link Aggregation Groups (LAG).
- Multicast hosts send IGMP Membership Reports as a reply to the IGMP Queries sent by the multicast router.
- When an IGMP Snooping switch receives an IGMPv2 Leave message from a group member on a certain interface, it sends an IGMPv2 Group-specific Query on that interface to determine if there any more group members interested in receiving multicast traffic. If the switch doesn't receive any Membership Reports during a certain time interval, it removes the interface from the multicast group and forwards the IGMPv2 Leave message to all multicast routers.

IGMP Snooping Configuration Example

This section provides steps to configure IGMP Snooping on the G8272. For this example, ethernet interface 1/12 and VLAN 100 are used.

1. Create VLAN 100:

```
G8272(config)# vlan 100  
G8272(config-vlan)# exit  
G8272(config)#
```

2. Set ethernet interface 1/12 as a member of VLAN 100:

- a. If the interface is configured as an access switch port:

```
G8272(config)# interface ethernet 1/12  
G8272(config-if)# bridge-port access vlan 100  
G8272(config)# exit
```

- b. If the interface is configured as a trunk switch port:

```
G8272(config)# interface ethernet 1/12  
G8272(config-if)# bridge-port trunk allowed vlan add 100  
G8272(config)# exit
```

3. (Optional) Enable IGMP Snooping on VLAN 100:

```
G8272(config)# vlan 100  
G8272(config-vlan)# ip igmp snooping
```

4. Disable IP multicast (IPMC) flooding:

```
G8272(config-vlan)# no flood
```

5. (Optional) Enable IGMPv3 Snooping (optional).

```
G8272(config-vlan)# ip igmp snooping version 3
```

6. Check IGMP Snooping settings:

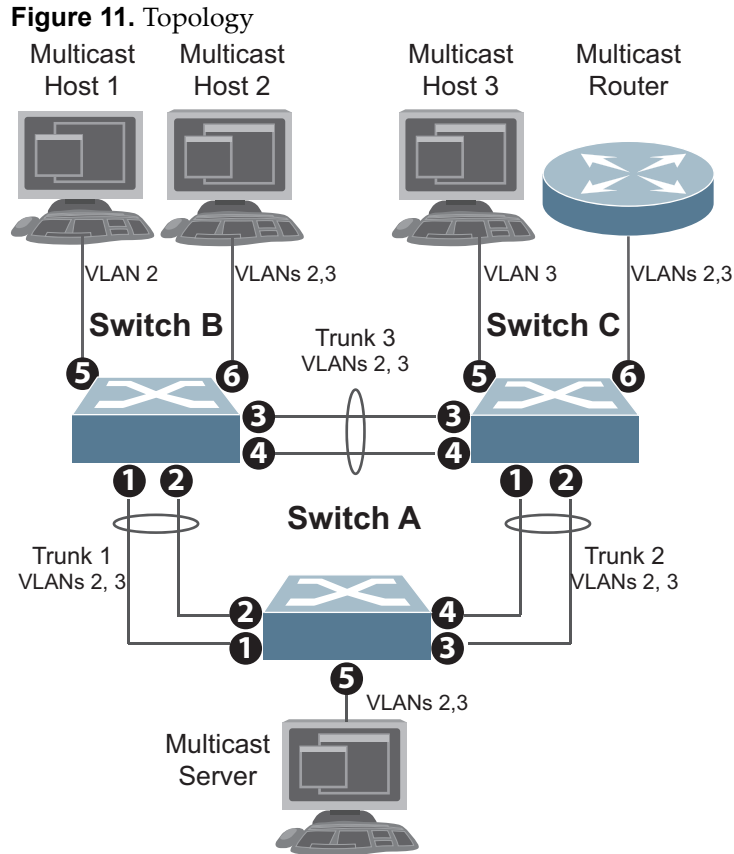
```
G8272> display ip igmp snooping

Global IGMP Snooping information
IGMP Snooping Enabled
IGMP Snooping V1/V2 Report Suppression Enabled
General query transmission on TCN Enabled

IGMP Snooping information for Vlan1
IGMP Snooping enabled
IGMP Snooping Version: 3
Robustness: 2 (operational: 2)
Query Interval: 125 seconds (operational: 125 seconds)
Group Membership Interval: 260 seconds
Query Response Interval: 10 seconds
Last Member Query Count: 2
Last Member Query Interval: 1000 milliseconds
IGMPv2 fast-leave: disabled
IGMPv1/v2 Report suppression: enabled
IGMPv3 Report suppression: disabled
Router port detection using: IGMP Queries
Snooping Querier disabled
Querier timeout: 255 seconds (default, operational: 255 seconds)
Querier Startup Query Count: 2
Querier Startup Query Interval: 31 seconds
Number of router-ports: 0
Number of Groups: 0
Number of Joins: 0
Number of Leaves: 0
Active Ports:
```

Advanced IGMP Snooping Configuration Example

Figure 11 shows an example topology. Switches B and C are configured with IGMP Snooping.



Devices in this topology are configured as follows:

- STG 2 includes VLAN 2; STG 3 includes VLAN 3.
- 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 multicast router sends IGMP Query packets in VLAN 2 and VLAN 3. The multicast router'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 multicast router receives all the multicast traffic.

Prerequisites

Before you configure IGMP Snooping, ensure you have performed the following actions:

- Configured VLANs.
- Enabled IGMP on the VLANs.
- Configured a switch or multicast router as the Querier.
- Identified the IGMP version(s) you want to enable.

Configuration

This section provides the configuration details of the switches shown in [Figure 11](#).

Switch A Configuration

1. Create VLANs 2 and 3.

```
G8272(config)# vlan 2,3  
G8272(config-vlan)# exit
```

2. Configure ethernet interfaces 1/1 - 1/5 as switch trunk ports and add VLANs 2 and 3 to their allowed VLAN lists:

```
G8272(config)# interface ethernet 1/1-5  
G8272(config-if-range)# bridge-port mode trunk  
G8272(config-if-range)# bridge-port trunk allowed vlan 2,3  
G8272(config-if-range)# exit
```

3. Assign a bridge priority lower than the default value to enable the switch to become the STP root in STG 2 and 3:

```
G8272(config)# spanning-tree mst 2 priority 4096  
G8272(config)# spanning-tree mst 3 priority 4096
```

4. Configure ethernet ports 1, 2, 3 and 4 as LACP LAG members:

```
G8272(config)# interface ethernet 1/1-2  
G8272(config-if-range)# aggregation-group 1 mode active  
G8272(config-if-range)# exit  
  
G8272(config)# interface ethernet 1/3-4  
G8272(config-if-range)# aggregation-group 2 mode active  
G8272(config-if-range)# exit
```

Switch B Configuration

1. Create VLANs 2 and 3:

```
G8272(config)# vlan 2,3  
G8272(config-vlan)# exit
```

2. Configure ethernet interfaces 1/1 - 1/4 and 1/6 as switch trunk ports and add VLANs 2 and 3 to their allowed VLAN lists:

```
G8272(config)# interface ethernet 1/1-4, ethernet 1/6  
G8272(config-if-range)# bridge-port mode trunk  
G8272(config-if-range)# bridge-port trunk allowed vlan 2,3  
G8272(config-if-range)# exit
```

3. Configure ethernet interface 1/5 as a switch access port and add VLAN 2 as its access VLAN:

```
G8272(config)# interface ethernet 1/5  
G8272(config-if)# bridge-port mode access  
G8272(config-if)# bridge-port access vlan 2  
G8272(config-if)# exit
```

4. Configure STP edge port:

```
G8272(config)# interface ethernet 1/5-6  
G8272(config-if-range)# spanning-tree port type edge  
G8272(config-if-range)# exit
```

5. Configure ethernet ports 1, 2, 3 and 4 as LACP LAG members:

```
G8272(config)# interface ethernet 1/1-2  
G8272(config-if)# aggregation-group 1 mode active  
G8272(config-if)# exit  
  
G8272(config)# interface ethernet 1/3-4  
G8272(config-if)# aggregation-group 3 mode active  
G8272(config-if)# exit
```

6. Configure IGMP Snooping:

```
G8272(config)# vlan 2,3  
G8272(config-vlan)# no flood  
G8272(config-vlan)# ip igmp snooping  
G8272(config-vlan)# ip igmp snooping version 3  
G8272(config-vlan)# exit
```

Switch C Configuration

1. Create VLANs 2 and 3:

```
G8272(config)# vlan 2,3  
G8272(config-vlan)# exit
```

2. Configure ethernet interfaces 1/1 - 1/4 and 1/6 as switch trunk ports and add VLANs 2 and 3 to their allowed VLAN lists:

```
G8272(config-vlan)# interface ethernet 1/1-4, ethernet 1/6  
G8272(config-if)# bridge-port mode trunk  
G8272(config-if)# bridge-port trunk allowed vlan 2,3  
G8272(config-if)# exit
```

3. Configure ethernet interface 1/5 as a switch access port and add VLAN 3 as its access VLAN:

```
G8272(config)# interface ethernet 1/5  
G8272(config-if)# bridge-port mode access  
G8272(config-if)# bridge-port access vlan 3  
G8272(config-if)# exit
```

4. Configure STP edge ports:

```
G8272(config)# interface ethernet 1/5-6  
G8272(config-if)# spanning-tree port type edge  
G8272(config-if)# exit
```

5. Configure ethernet ports 1, 2, 3 and 4 as LACP LAG members:

```
G8272(config)# interface ethernet 1/1-2  
G8272(config-if)# aggregation-group 2 mode active  
G8272(config-if)# exit  
  
G8272(config)# interface ethernet 1/3-4  
G8272(config-if)# aggregation-group 3 mode active  
G8272(config-if)# exit
```

6. Configure IGMP Snooping:

```
G8272(config)# vlan 2,3  
G8272(config-vlan)# no flood  
G8272(config-vlan)# ip igmp snooping  
G8272(config-vlan)# ip igmp snooping version 3  
G8272(config-vlan)# exit
```

Troubleshooting

This section provides the steps to resolve common IGMP Snooping configuration issues. The topology described in [Figure 11](#) is used as an example.

Multicast traffic from non-member groups reaches the host or multicast router

- Check if traffic is unregistered. For unregistered traffic, an IGMP entry is not displayed in the IGMP groups table.

```
G8272> display ip igmp snooping groups
```

- Ensure IPMC flooding is disabled.

```
G8272(config)# vlan <VLAN number>  
G8272(config-vlan)# no flood
```

- Check the egress port's VLAN membership. The ports to which the hosts and multicast router are connected must be used only for VLAN 2 and VLAN 3.

```
G8272> display vlan id 2,3
```

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.

```
G8272> display ip igmp snooping 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.

```
G8272> display interface ethernet <chassis number/port number> counters
```

- Ensure no static multicast MACs, static multicast groups or static multicast routers are 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.

```
G8272# display ip igmp snooping 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:

```
G8272# remove ip igmp snooping groups *
```

However, this will clear all the IGMP groups and will influence other hosts.

- 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 multicast router do not reach the host.

- Ensure the multicast router is learned on switches B and C.

```
G8272> display ip igmp snooping 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 multicast router

- Ensure IGMP Queries sent by the multicast router reach the hosts.
- Ensure the multicast router is learned on both switches. Note that the multicast router may not be learned on switch B immediately after a LAG failover/failback.

```
G8272> display ip igmp snooping mrouter
```

- Ensure the host's multicast application is started and is sending correct IGMP Reports/Leaves.

```
G8272> display ip igmp snooping groups
G8272> display ip igmp snooping statistics
```

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 multicast router 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

- Ensure a storm control is not configured on the LAGs.

```
G8272(config)# interface ethernet <chassis number/port number>  
G8272(config-if)# no storm-control multicast level
```

- Check link speeds and network congestion.

Additional IGMP Features

The following topics are discussed in this section:

- [“Report Suppression” on page 335](#)
- [“Robustness Variable” on page 335](#)
- [“Fast Leave” on page 336](#)
- [“Static Multicast Router” on page 337](#)

Report Suppression

An IGMP Snooping switch will forward to a multicast router all IGMPv1 and IGMPv2 reports. To reduce IGMP traffic on a VLAN, you can enable the suppression of such reports.

Report suppression will forward only the first occurrences of IGMPv1 and IGMPv2 reports for any multicast group. Any subsequent reports from the same multicast group are discarded by the IGMP Snooping switch and not forwarded to the multicast router.

By default, report suppression is enabled on the switch for all VLANs. Report suppression can be enabled or disabled globally or for each VLAN.

To globally enable or disable report suppression, use the following command:

```
G8272(config)# [no] ip igmp snooping report-suppression
```

To enable or disable report suppression on a VLAN, use the following command:

```
G8272(config-vlan)# [no] ip igmp snooping report-suppression
```

Note: If report suppression is disabled globally, any individual VLAN configuration is ignored and report suppression is disabled for all VLANs.

Note: IGMPv3 reports cannot be suppressed.

Robustness Variable

The Robustness Variable allows tuning for the expected packet loss on a network. In this scenario, the robustness variable can be increased to reduce the impact of traffic loss.

You can increase this value by using the following command:

```
G8272(config-vlan)# ip igmp snooping robustness-variable <1-7>
```

You can reset IGMP Robustness to its default value by using the following command:

```
G8272(config-vlan)# no ip igmp snooping robustness-variable
```

Fast Leave

When an IGMP Snooping switch receives an IGMPv2 Leave message from a group member on a certain interface, it sends an IGMPv2 Group-specific Query on that interface to determine if there any more group members interested in receiving multicast traffic. If the switch doesn't receive any Membership Reports during a time interval called last-member-query-time, it removes the interface from the multicast group and forwards the IGMPv2 Leave message to all multicast routers.

The IGMP Snooping switch will send a particular number of IGMPv2 Group-specific Queries waiting a certain period of time between each query. The number of queries is called last-member-query-count and the period of time between each query is called last-member-query-interval.

The last-member-query-time is defined as the last-member-query-interval multiplied by the last-member-query-count. While the last-member-query-time isn't directly configurable, you can configure its components.

To configure the last-member-query-interval (in seconds), use the following command:

```
G8272(config-vlan)# ip igmp snooping last-member-query-interval <1-25>
```

The default value is one second and you can reset the last-member-query-interval to this value by using the following command:

```
G8272(config-vlan)# no ip igmp snooping last-member-query-interval
```

The last-member-query-count is equal to the value of the robustness variable. To modify the robustness variable, see [“Robustness Variable” on page 335](#).

Note: When host generates an IGMPv2 Leave message, the source IPv4 address of the message will be 0.0.0.0.

If Fast Leave is enabled on a VLAN, the IGMP Snooping switch immediately removes an interface from the multicast group when receiving an IGMPv2 Leave message on that interface and forwards the Leave message to all multicast routers.

Note: Only IGMPv2 supports Fast Leave. Enable Fast Leave on ports that have only one host connected. If more than one host is connected to a port, you may lose some hosts unexpectedly.

To enable or disable Fast Leave on a VLAN, use the following command:

```
G8272(config-vlan)# [no] ip igmp snooping fast-leave
```


Static Multicast Router

A static multicast router can be configured for a particular ethernet interface or Link Aggregation Group (LAG) on a particular VLAN. Any static multicast routers don't have to be learned through IGMP Snooping, but they can be also learned this way.

An multicast router is kept as an entry in the multicast group table of a IGMP Snooping switch. If the multicast router is both statically configured and dynamically learned, it will appear only as one IGMP entry, but will have both static (S) and dynamic (D) flags.

To add or remove a static multicast router on an interface of a VLAN, use the following command (for this example, ethernet interface 1/12 is used):

```
G8272(config-vlan)# [no] ip igmp snooping mrouter interface ethernet 1/12
```

To remove all static multicast routers, use the following command:

```
G8272(config-vlan)# no ip igmp snooping mrouter all
```

To view all static and dynamic multicast routers, use the following command:

```
G8272> display ip igmp snooping mrouter
```

Chapter 19. Border Gateway Protocol

Border Gateway Protocol (BGP) is an routing protocol that enables routers on a network to advertise routing information with each other about the segments of the IP (IPv4 or IPV6) 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 4271.

The RackSwitch G8272 can advertise their IP interfaces and IP 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.

The following topics are discussed in this section:

- [“BGP Overview” on page 340](#)
- [“Internal Routing Versus External Routing” on page 341](#)
- [“Route Reflector” on page 343](#)
- [“Forming BGP Peer Routers” on page 346](#)
- [“Loopback Interfaces” on page 348](#)
- [“What is a Route Map?” on page 349](#)
- [“Aggregating Routes” on page 352](#)
- [“Redistributing Routes” on page 353](#)
- [“BGP Path Attributes” on page 359](#)
- [“Best Path Selection Logic” on page 361](#)
- [“BGP Failover Configuration” on page 375](#)
- [“Default Redistribution and Route Aggregation Example” on page 377](#)
- [“Designing a Clos Network Using BGP” on page 379](#)

BGP Overview

BGP is an inter-Autonomous System routing protocol. The primary function of BGP is to exchange network reachability information with other BGP systems. This information is sufficient for constructing a graph of Autonomous System (AS) connectivity for this reachability, from which routing loops may be pruned and, at the AS level, policy decisions may be enforced.

Routing information exchanged via BGP supports the destination-based forwarding paradigm, which assumes a router forwards a packet based solely on the destination address carried in the IP header of the packet. This, in turn, determines the policy decisions that can and cannot be enforced using BGP.

BGP supports only those policies conforming to the destination-based forwarding paradigm over IPv4 as described in RFC 4271.

BGP Router Identifier

To establish BGP sessions between peers, BGP must have a router ID, which is sent to BGP peers in the OPEN message when a BGP session is established. The BGP router ID is a 32-bit value that is often represented by an IPv4 address. You can configure the router ID.

By default, Cloud NOS can select a router ID among all the IP addresses of all interfaces automatically. If there are loopback interfaces with configured IPv4 address, the highest address among them is chosen as the router ID. If no loopback interfaces exist, the highest IPv4 address whose interface is in the up state will become the router ID. If that interface goes down, the current router ID is kept.

The BGP router ID must be unique to the BGP peers in a network. If BGP does not have a router ID, it cannot establish any peering sessions with BGP peers.

To configure the router ID, use the following command:

```
G8272(config)# router-id <IPv4 address>
```

To view the router ID, use the following command:

```
G8272# display router-id  
Router ID: 191.1.1.1 (automatic)
```

```
G8272# display router-id  
Router ID: 26.4.4.4 (config)
```

To reset the router ID to its default value (the one automatically chosen by the switch), use the following command:

```
G8272(config)# no router-id
```

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 OSPF.

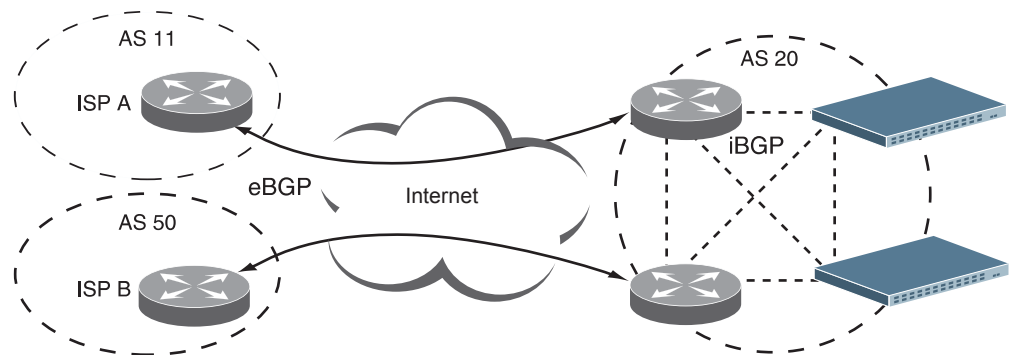
Static routes must have a higher degree of precedence than dynamic routing protocols. If the destination route is not in the route cache, the packets are forwarded to a default gateway which may be incorrect if a dynamic routing protocol is enabled.

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 12](#)), 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 G8272.

Figure 12. 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 343](#).

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 Cloud NOS implementation conforms to the BGP Route Reflection specification defined in RFC 4456.

As per RFC 4271 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.

Route reflection is an alternative to using a full mesh. This approach allows a BGP speaker known as a route reflector (RR) to advertise iBGP learned routes to certain iBGP peers. It represents a change in the commonly understood concept of iBGP, and the addition of two new optional non-transitive BGP attributes to prevent loops in routing updates.

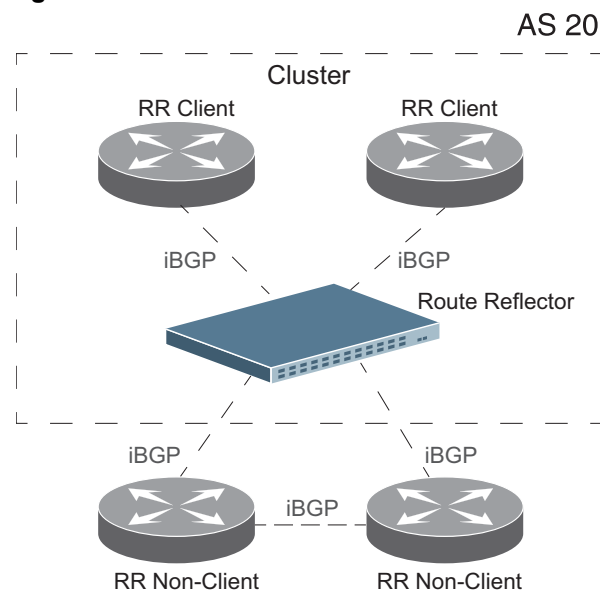
A route reflector has two groups of internal peers: clients and non-clients. A route reflector reflects between these groups and among the clients. The non-client peers must be fully meshed. The route reflector and its clients form a cluster.

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 13](#), the G8272 is configured as a route reflector. All clients and non-clients are in the same AS.

Figure 13. 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. When a router 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 can be configured with the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# cluster-id {<ID (1-4294967295)>|<IPv4 address>}
```

The value configured with the above command (or the router ID of the route reflector if the cluster-id is not configured) is prepended to the Cluster list attribute. When 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 Configuration Example

To configure BGP route reflection, follow the steps below:

1. Configure an AS.

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)#
```

2. Configure a route reflector client.

```
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>  
G8272(config-router-neighbor)# address-family ipv4 unicast  
G8272(config-router-neighbor)# exit
```

Note: When a client is configured on the G8272, the switch automatically gets configured as a route reflector.

3. Verify configuration.

```
G8272# display ip bgp neighbors <neighbor address>

BGP neighbor is 10.1.50.1, remote AS 22, local AS 3, external link
  BGP version 4, remote router ID 0.0.0.0
  BGP state = Idle
  Last read 01w6d03h, hold time is 180, keepalive interval is 60 seconds
  Received 0 messages, 0 notifications, 0 in queue
  Sent 0 messages, 0 notifications, 0 in queue
  Route refresh request: received 0, sent 0
  Minimum time between advertisement runs is 30 seconds
  Connections established 0; dropped 0
  TTL: 1, TTL Security hops: 0

Update packets: 0
Update packets dropped: 0
- Decode error drops: 0
- Internal error drops: 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:

```
G8272(config-router-af)# no client-to-client reflection
```

You can view the route reflector BGP attributes attached to a BGP route using the following command:

```
G8272# display ip bgp attribute-info
attr[3] nexthop 0.0.0.0
```

Restrictions

Consider the following restriction when configuring route reflection functionality:

- 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 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.

BGP Peers and Dynamic Peers

A BGP speaker does not discover another BGP speaker automatically. You must configure the relationships between BGP speakers. A BGP peer is a BGP speaker that has an active TCP connection to another BGP speaker.

Cloud NOS accepts an IP address range to establish BGP sessions. For example, if you configure BGP to use IPv4 prefix 192.168.1.0/24, BGP will establish a session with 192.168.1.2 but will reject a session from 192.168.2.2.

Static Peers

You can configure BGP static peers by using the following commands:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>
G8272(config-router-neighbor)# address-family {ipv4|ipv6} unicast
```

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 (in the following format: IP address/subnet mask). BGP waits to receive an open message initiated from BGP speakers within that range. Dynamic peers are automatically created when a peer from the same subnet initiates a TCP connection. Dynamic peers are passive. When they are not in the established state, they accept inbound connections but do not initiate outbound connections.

When the BGP speaker receives an open message from a dynamic peer, the AS number from the packet must match the configured remote AS number.

All static BGP peer attributes can also be configured for dynamic peers.

To set the maximum number of dynamic peers for a group that can simultaneously be in an established state, enter the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor address>/<prefix> remote-as <AS number>
G8272(config-router-neighbor)# maximum-peers <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 (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.

The following are the basic commands for configuring dynamic peers:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor address>/<prefix> remote-as <AS number>
```

You cannot remove dynamic peers manually. However, you can stop a dynamic peer using the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# no neighbor <neighbor address>/<prefix>
```

The stop command interrupts the BGP connection until the peer tries to reestablish 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:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# neighbor <neighbor address>  
G8272(config-router-neighbor)# update-source loopback <0-7>
```


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 learned 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.

```
G8272(config)# route-map <route map name>  
G8272(config-route-map)# apply ip next-hop peer-address <peer IP address>
```

Incoming and Outgoing Route Maps

A BGP peer router can be configured to support up to two simultaneous route maps; one incoming route map and one outgoing route map.

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):

```
G8272(config)# route-map <route map name> <sequence number (1-65535)>
```

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. Create the IP address prefix-list, specifying the IPv4 address and subnet mask of the network that you want to match.

```
G8272(config)# ip prefix-list <prefix list name> permit <IP address>/<prefix length>
```

2. Define a network filter.

```
G8272(config)# route-map <route map name>  
G8272(config-route-map)# match ip address prefix-list <prefix list name>
```

Enable the network filter. You can distribute up to 256 network filters among 64 route maps each containing 32 access lists.

3. (Optional) Configure the AS filter attributes.

```
G8272(config)# ip as-path access-list <ACL name> permit <Regular expression to match>  
G8272(config)# route-map <route map name>  
G8272(config-route-map)# match as-path <ACL name>
```

4. (Optional) Set up the BGP attributes.

If you want to overwrite the attributes that the peer router is sending, define the following BGP attributes:

- Specify AS paths that you want to prepend to a matched route and the local preference for the matched route.
- Specify the metric [Multi Exit Discriminator (MED)] for the matched route.

```
G8272(config-route-map)# apply as-path prepend <AS number>  
[<AS number>...]  
G8272(config-route-map)# apply local-preference <local preference value>  
G8272(config-route-map)# apply metric <metric value>  
G8272(config-route-map)# exit
```

5. Assign the route map to a peer router.

Select the peer router and then add the route map to the incoming route map list,

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# neighbor <neighbor address>  
G8272(config-router-neighbor)# address-family ipv4 unicast  
G8272(config-router-neighbor-af)# route-map <route map name> in
```

or to the outgoing route map list.

```
G8272(config-router-neighbor-af)# route-map <route map name> out
```

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:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# address-family ipv4 unicast
G8272(config-router-af)# aggregate-address <IP address>/<prefix length>
```

The origin code for the aggregated route is either IGP or INCOMPLETE.

An example of creating a BGP aggregate route is shown in [“Default Redistribution and Route Aggregation Example”](#) on page 377.

Redistributing Routes

BGP injects local routes in two different ways:

- Using network configuration commands. These commands list networks that are candidates if they appear in the routing table.
- Using redistribution from another routing protocol.

In addition to running multiple routing protocols simultaneously, CNOS 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 349](#). 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 377](#).

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

The origin code for the redistributed routes is INCOMPLETE.

The metrics of the non-BGP-learned routes are transferred to the metrics or MEDs of the new BGP route.

A route map can be used for each redistributed protocol to control which routes are redistributed. A route map can also be used to modify the BGP attributes of the redistributed routes. The switch can redistribute directly connected routes, static routes or routes learned through OSPF.

To configure route redistribution, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# address-family {ipv4|ipv6} unicast
G8272(config-router-af)# redistribute {direct|ospf|static} [route-map
<route map name>]
```

To view statistics for redistributed routes, use the following command:

```
G8272# display {ipv4|ipv6} {unicast|multicast} policy statistics
redistribute {direct|static|all}
```

This command displays the following information:

- Route-map: The route-map name (if route-map is applied at redistribution)
- Compared count: How many static and connected routes were considered for redistribution into BGP
- Matched count: How many of these routes actually matched the route-map
- Total compared count: How many static and connected routes can be found in the routing table (including management routes)
- Total matched count: How many static and connected routes were actually redistributed by BGP (all routes minus management routes)

To clear statistics for redistributed routes, use the following command:

```
G8272# remove {ipv4|ipv6} {unicast|multicast} policy statistics
redistribute {direct|static|all}
```

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 G8272 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 replace the existing communities. To set communities to prefixes that pass the route-map, use the following commands:

```
G8272(config)# route-map <route map name>
G8272(config-route-map)# apply community <AA:NN (autonomous system number:
community number)>
```

To remove all community tags from prefixes that pass the route-map, use the following command:

```
G8272(config-route-map)# apply community none
```

To propagate prefixes that pass the route-map with their original community settings, use the following command:

```
G8272(config-route-map)# no apply community <AA:NN>
```

To enable or disable community tags forwarding for specific neighbors or neighbor groups, use the following commands:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor address>[/<prefix>]
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# send-community
```

BGP Community

A *community* is a group of destinations which share some common property; this may be added to each prefix. Communities are transitive optional attributes. The community attribute can be thought of as a flat, 32-bit value that can be applied to any set of prefixes. It can be read as a 32-bit value or split into two portions, the first two bytes representing an ASN and the last two bytes as a value with a predetermined meaning. The community attribute is always used to do routing policy.

The values 0x00000000 through 0x0000FFFF and 0xFFFF0000 through 0xFFFFFFFF are reserved. Most modern router software displays communities as ASN:VALUE. In this format the communities 1:0 through 65534:65535 are available for use. The convention is to use the ASN of your own network as the leading 16 bits for your internal communities and communities with which you accept and send data.

Each autonomous system administrator may define which communities a destination belongs to. By default, all destinations belong to the general Internet community.

Three communities are defined in RFC 1997 and are standard within BGP implementations:

- **NO-EXPORT (0xFFFFF01)**
NO-EXPORT is commonly used within an AS to instruct routers not to export a prefix to eBGP neighbors.
- **NO-ADVERTISE (0xFFFFF02)**
NO-ADVERTISE instructs a BGP-speaking router not to send the tagged prefix to any other neighbor, including other iBGP routers.
- **NO-ADVERTISE-SUBCONFED (0xFFFFF03).**
NO-ADVERTISE-SUBCONFED is used to prevent a prefix from being advertised to external BGP peers (this includes peers in other members autonomous systems inside a BGP confederation).

BGP community attribute exchange between BGP peers is enabled when the `send-community` command is configured for the specified neighbor.

When multiple values are configured in the same community list statement, a logical AND condition is created. All community values must match to satisfy this AND condition. When multiple values are configured in separate community list statements, a logical OR condition is created. The first list that matches a condition is processed.

To add or remove a community list, use the following command:

```
G8272(config)# [no] ip community-list {standard|expanded} <community list name> {deny|permit} {<AA:NN>|internet|local-as|no-advertise|no-export}
```

For example, the following command adds a community list called 'myCommList', which permits the advertising of routes to the internet community.

```
G8272(config)# ip community-list standard myCommList internet
```

To view the configuration, use the following command:

```
G8272# display ip community-list
Named Community standard list myCommlist
  permit internet
```

BGP Extended Community

An upgrade to the standard community with a transitive-optional attribute (Type 16) is called the *extended community*. The current Internet RFC defines the extended community as an 8-octet value, while the community is only a 4-octet value. The first octet specifies the type; an optional second value can specify a subtype. This value dictates the structure given to the remaining octets.

The Type field gives the community some immediate flexibility. The first is the use of bit 0 to represent whether the community is registered with the IANA. The second bit gives the Extended Community a coarse scope, either Transitive, meaning it may be passed between ASs, or Non-Transitive, meaning it should be carried only within the local AS.

BGP extended community attribute exchange between BGP peers is enabled when the `send-community` command is configured for the specified neighbor.

When multiple values are configured in the same extended community list statement, a logical AND condition is created. All extended community values must match to satisfy this AND condition. When multiple values are configured in separate extended community list statements, a logical OR condition is created. The first list that matches a condition is processed.

To add or remove an extended community list, use the following command:

```
G8272(config)# [no] ip extcommunity-list {standard|expanded} <extended
community list name> {deny|permit} {<AA:NN>|rt|soo}
```

For example, the following commands adds an extended community list called 'denyRT-100-SOO-133', which denies routes from route target 34400:100 and site of origin 2500:133:

```
G8272(config)# ip extcommunity-list standard denyRT-100-SOO-133 deny rt
34400:100 soo 2500:133
```

To view the configuration, use the following command:

```
G8272# display ip extcommunity-list
Named extended community standard list denyRT-100-SOO-133
  deny 34400:100 2500:133
```

BGP Confederation

A fully meshed iBGP network becomes complex as the number of iBGP peers grows. You can reduce the iBGP mesh by dividing the autonomous system into multiple sub-autonomous systems and grouping them into a single confederation. A *confederation* is a group of iBGP peers that use the same autonomous system number (confederation identifier) to communicate to external networks. Each sub-autonomous system is fully meshed within itself and has a few connections to other sub-autonomous systems in the same confederation. The peers in sub-autonomous systems communicate with each other using sub-autonomous AS numbers.

To configure the confederation AS, use the following command:

```
G8272(config)# routing-protocol bgp <AS number>  
G8272(config-router)# confederation identifier <confederation AS (1-65535)>
```

To delete the confederation AS, use the following command:

```
G8272(config-router)# no confederation identifier
```

To configure the peer member AS number to be part of the BGP confederation, use the following command:

```
G8272(config)# routing-protocol bgp <AS number>  
G8272(config-router)# confederation peers <member AS (1-65535)> ...
```

Note: You can add multiple confederation members with a single command.

Note: The member AS number is only visible with the BGP confederation.

BGP Path Attributes

Path attributes are the characteristics of an advertised BGP route. BGP routing policy is set and communicated using the path attributes.

There are four kinds of BGP path attributes:

- Well-known mandatory
- Well-known discretionary
- Optional transitive
- Optional non-transitive

Well-Known Mandatory

Well-known, mandatory attributes must appear in every UPDATE message. It must be supported by all BGP software implementations. If a well-known, mandatory attribute is missing from an UPDATE message, a NOTIFICATION message must be sent to the peer.

Cloud NOS supports the following well-known, mandatory attributes:

- AS_PATH: the sequence of autonomous systems (AS) through the network from the switch to the destination;
- ORIGIN: the origin of the BGP route (IGP, EGP or incomplete);
- NEXT_HOP: the IP address of the next hop on the route.

Well-Known Discretionary

Well-known, discretionary attributes may or may not appear in an UPDATE message, but they must be supported by any BGP software implementation.

Cloud NOS supports the following well-known, discretionary attributes:

- LOCAL_PREF: the preference value of a certain path;
- ATOMIC_AGGREGATE: informs the neighboring AS that the originating routes has aggregated routes.

Optional Transitive

Optional, transitive attributes may or may not be supported in all BGP implementations. If an optional, transitive attribute is sent in an UPDATE message but is not recognized by the receiver, it must be passed on to the next AS.

Cloud NOS supports the following optional, transitive attributes:

- AS4_PATH: informs that the AS number is stored in four-octet format;
- AS4_AGGREGATOR: informs that the aggregator carries an AS number in four-octet format;
- AGGREGATOR: specifies the IP address and AS number of route aggregating BGP peer;

- **COMMUNITY:** a common policy used by BGP peers that are in the same group;
- **EXTENDED COMMUNITIES:** an extended range BGP community;

Optional Non-Transitive

Optional, non-transitive attributes may not be supported and will not be passed on if advertised. If an optional, non-transitive attribute is received, the router does not have to pass it on and may safely and quietly ignore the attribute.

Cloud NOS supports the following optional, non-transitive attributes:

- **MULTI_EXIT_DISC (MED):** indicates the preferred entry point into an AS, when multiple entry points are available;
- **ORIGINATOR_ID:** the router ID of route origin;
- **CLUSTER_LIST:** the sequence of the CLUSTER ID (the router ID) values representing the reflection path that the route has passed;
- **MP_REACH_NLRI:** Multiprotocol Reachable NLRI - advertises a feasible route to a peer or permits a router to advertise the Network Layer address of the BGP peer that should be used as the next hop address to the destinations listed in the Network Layer Reachability Information field;
- **MP_UNREACH_NLRI:** Multiprotocol Unreachable NLRI - withdraws multiple unfeasible routes from service.

Best Path Selection Logic

A BGP speaker may have several paths to the same destination. The local speaker can select the best paths, and then install it into the routing information base (RIB). The best path selection logic includes the following:

- Exclude paths that matches any of the following conditions.
 - Paths for which the NEXT_HOP is inaccessible
 - Paths for which it's punished by damping
 - Paths that are not synchronized if synchronization is configured
 - Paths from an external BGP (eBGP) neighbor if the local autonomous system (AS) appears in the AS_PATH
- If you enable `bgp enforce-first-as` and the UPDATE does not contain the AS of the neighbor as the first AS number in the AS_SEQUENCE, the router sends a notification and closes the session.
- Paths that are marked as `received-only` in the `display ip bgp longer-prefixes` output have been rejected by the policy. However, the router has stored the paths because you have configured `soft-reconfiguration inbound` for the neighbor that sent the path.

BGP Best Path Selection

If there are multiple paths, BGP compares the best path with the next path until BGP reaches the end of valid paths. The following rules determine the best path:

- Prefer the path with the highest WEIGHT.
- Prefer the path with the highest LOCAL_PREF.
- Prefer the path that was locally originated via a network command, then the path that was locally aggregated.
- Prefer the path with the shortest AS_PATH.
- Prefer the path with the lowest origin type. Interior Gateway Protocol (IGP) is lower than Exterior Gateway Protocol (EGP), and EGP is lower than INCOMPLETE.
- Prefer the path with the lowest multi-exit discriminator (MED).
- Prefer eBGP over Internal Border Gateway Protocol (iBGP) paths.
- Prefer the path with the lowest IGP metric to the BGP next hop.
- When both paths are external, prefer the path that was received first (the oldest one).
- Prefer the route that comes from the BGP router with the lowest router ID.
- If the originator or router ID is the same for multiple paths, prefer the path with the minimum cluster list length.
- Prefer the path that comes from the lowest neighbor address.

BGP Weight

The weight is assigned locally to each individual router, and the value is only understood by that particular router. The value is neither propagated nor carried through in route updates.

A weight can be any integer from 0 to 65535. Routes with a higher weight value have preference when there are multiple routes to the same destination. Routes learned through another BGP peer have a default weight of 0, and routes sourced by the local router (the paths that the router originated) have a default weight of 32768.

To configure the weight of a BGP peer, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>  
G8272(config-router-neighbor)# weight <0-65535>
```

To remove the weight associated with a BGP peer, use the following command:

```
G8272(config-router-neighbor)# no weight
```

Local Preference

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 following commands use the route map local preference method, which affects both inbound and outbound directions.

```
G8272(config)# route-map <route map name>  
G8272(config-route-map)# apply local-preference <0-4294967295>
```

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

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. The edge router receiving this next hop attribute, if it has any iBGP connection, will advertise this same IP address as the next hop attribute to other iBGP routers. However, this can sometimes cause routing path failures in Non-Broadcast Multiaccess Networks (NBMA). In some cases the routers inside this AS do not have a direct route to the external IP address in this next hop attribute, causing the route to not be installed in the hardware.

To avoid routing failures, you can manually configure the next hop IP address. When using NBMA networks, you can configure the external BGP speaker to advertise its own IP address as the next hop. With 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:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor IP address>
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# next-hop-self
```

- Next Hop for a BGP Peer Group:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor IP address>/<prefix>
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# next-hop-self
```

Best Path Selection Tuning

There are several ways of tuning the BGP path selection. They are available by using the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# bestpath ?

always-compare-med      Allow comparing MED from different
                           neighbors
as-path                 AS-path attribute
compare-confed-asp     Allow comparing confederation AS path
                           length
compare-routerid      Compare router-id for identical EBGP paths
dont-compare-origina   Don't Compare originator-id for BGP
med                    MED attribute
tie-break-on-age      Prefer old route when compare-route-id is
                           not set
```

- **bestpath always-compare-med**

In best path selection section, BGP speaker compare the MED of 2 paths only when they are from same AS. When customer enables this option, BGP speaker will compare MED even from different AS.

- **bestpath med**

Use this command to specify two MED (Multi Exit Discriminator) attributes:

- **bestpath med confed**

Enables MED comparison among paths learned from confederation peers. The MEDs are compared only if there is no external Autonomous System (an AS not within the confederation) in the path. If there is an external autonomous system in the path, the MED comparison is not made.

For example in the following paths, the MED is not compared with Path3 as it is not in the confederation. MED is compared for Path1 and Path2 only.

```
Path1 = 32000 32004, med=4
Path2 = 32001 32004, med=2
Path3 = 32003 1, med=1
```

- **bestpath med missing-as-worst**

Considers a missing MED attribute in a path as having a value of infinity, making the path without a MED value the least desirable path. If `missing-as-worst` is disabled, the missing MED is assigned the value of 0, making the path with the missing MED attribute the best path.

- **bestpath med non-deterministic**

When you enabled, BGP route selection process does not group routes from the same AS, the best route will depend on the order of routes receiving.

- **bestpath med remove-recv-med**

Removes the receive MED attribute from the MED attribute.

- **bestpath med remove-send-med**

Removes the send MED attribute from the MED attribute.

- **bestpath compare-routerid**

Use this command to compare router IDs for identical eBGP paths. When comparing similar routes from peers, the BGP router does not consider the router ID of the routes. By default, it selects the first received route. This command causes BGP to include the router ID in the selection process. Similar routes are compared, and the route with lowest router ID is selected.

- **bestpath tie-break-on-age**

Use this command to always select the older preferred route when the switch doesn't compare router IDs for identical eBGP paths.

- **bestpath compare-confed-aspath**

Use this command to allow comparison of the confederation AS path length. This command specifies that the AS confederation path length must be used when available in the BGP best path decision process.

- **bestpath dont-compare-originator-id**

Use this command to compare "not" an originator ID for an identical eBGP path.

- **bestpath as-path multipath-relax**

Use this command to have the BGP speaker handle the paths received from different autonomous systems for multipath if their AS-path lengths are the same and all other multipath conditions are met.

- **bestpath as-path ignore**

Ignores the AS path length when selecting a BGP route.

BGP ECMP

The BGP multipath feature is used for balancing forwarding loads on eBGP or iBGP routers when the destination route can be reached through multiple Equal Cost next hops. Equal cost is determined via the best path selection mechanism, and all paths that have equal cost are considered Multipath Candidates. To be candidates for multipath, paths to the same destination must have the following characteristics equal to the best-path characteristics:

- Weight
- Local preference
- AS path length
- Origin
- MED
- One of the following:
 - Neighboring AS or sub-AS
 - AS path

BGP Features and Functions

AS-Path Filter

When the BGP speaker uses the `as-path` parameter to filter inbound or outbound routes, it compares the regular express of AS-path to the AS path of the inbound outbound routes. If they match and the action is `permit`, the routes are installed or sent to peers. If they match and the action is `deny`, the routes are denied or not sent to peers.

To configure the AS path filter, use the following command:

```
G8272(config)# ip as-path access-list <ACL name> {deny|permit} <regular expression to match>
```

BGP Capability Code

A BGP speaker can negotiate capability by sending a capability code in an OPEN message when establishing the BGP session. CNOS supports the following capability codes:

Table 18. BGP Capability Codes

Capability Code	Capability Name
1	Multiprotocol Extensions for BGP-4
2	Route Refresh Capability for BGP-4
64	Graceful Restart Capability
65	Support for 4-octet AS number capability
67	Support for Dynamic Capability (capability specific)

Administrative Distance

An *administrative distance* is a rating of the trustworthiness of a routing information source. The administrative distance does not influence the BGP path selection algorithm, but it does influence whether BGP-learned routes are installed in the IP routing table.

TTL-Security Check

The BGP TTL Security Check feature protects the eBGP peering session by comparing the value in the TTL field of received IP packets against a hop count that is configured locally for each eBGP peering session:

- If the value in the TTL field of the incoming IP packet is greater than or equal to the locally configured value, the IP packet is accepted and processed normally.
- If the TTL value in the IP packet is less than the locally configured value, the packet is silently discarded, and no ICMP message is generated.

This is designed behavior; a response to a forged packet is unnecessary.

By default, eBGP doesn't support eBGP multi-hop. To enable or disable the exterior Border Gateway Protocol (eBGP) time-to-live (TTL) value to support eBGP multi-hop, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>
G8272(config-router-neighbor)# [no] ebgp-multihop <1-255>
```

By default, there is no limit to how many hops can separate two BGP peers. To configure the maximum number of hops that separate two peers, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>
G8272(config-router-neighbor)# ttl-security hops <1-254>
```

To reset this limit to its default value (0), use the following command:

```
G8272(config-router-neighbor)# no ttl-security hops
```

Note: When this feature is configured for a multi-hop peering session, the **ebgp-multihop** command cannot be used. If you attempt to use both commands for the same peering session, an error message will be displayed on the console.

Local-AS

The local AS feature allows a router to appear to be a member of a second autonomous system (AS) in addition to its real AS. Local AS allows two ISPs to merge without modifying peering arrangements. Routers in the merged ISP become members of the new autonomous system but continue to use their old AS numbers for their customers. You can only use the Local AS feature with true eBGP peers; you cannot use this feature for two peers that are members of different confederation sub-autonomous systems.

To enable or disable eBGP to prepend the local AS number to the AS path attribute, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>
G8272(config-router-neighbor)# [no] local-as <AS number (1-4294967295)>
```

BGP Authentication

You can configure MD5 authentication between two BGP peers; each segment sent on the TCP connection between the peers is verified. MD5 authentication must be configured with the same password on both BGP peers or the connection between them will not be made. Configuring MD5 authentication causes the BGP speaker to generate and check the MD5 digest of every segment sent on the TCP connection.

To enable or disable BGP to use MD5 authentication when communicating with a certain neighbor, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>  
G8272(config-router-neighbor)# [no] password <neighbor password>
```

Note: An unencrypted password can be configured by using the following command:

```
G8272(config-router-neighbor)# [no] password 0 <neighbor password>
```

Note: For dynamic BGP peers, MD5 authentication is not supported.

Originate Default Route

By default, BGP sends the known default routes to neighbors, but will not originate a default route. If a route is configured instead of learned, it will not be sent. To originate default routes to neighbors, the `default-originate` feature must be enabled.

To enable or disable a BGP routing process to distribute a default route, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>  
G8272(config-router-neighbor)# address-family {ipv4|ipv6} unicast  
G8272(config-router-neighbor-af)# [no] default-originate
```

Optionally, a route map can be specified to allow the default route to be injected conditionally. To do this, use the following command:

```
G8272(config-router-neighbor-af)# [no] default-originate route-map <route  
map name>
```


IP Prefix-List Filter

There are several ways to filter BGP routes:

- IP prefix-list filter
- as-path filter
- route-map filter

When the BGP speaker uses the IP prefix list to filter inbound routes or outbound routes, it compares the prefix from the prefix list to the prefix of the inbound routes or outbound routes. If they match and the action is **permit**, the routes are installed or sent to peers. If they match and the action is **deny**, the routes are denied or not sent to peers.

To create or delete an IP prefix list, use the following command:

```
G8272(config)# ip prefix-list <prefix list name> ?
deny          Specify packets to reject
description   Prefix-list specific description
permit       Specify packets to forward
seq          Sequence number of an entry
```

For more details about the above command, see the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

For example, the following command adds an IP prefix list called 'pref-list-35' that forwards BGP routes that match IPv4 address 35.0.0.0 with prefix length 8:

```
G8272(config)# ip prefix-list pref-list-35 permit 35.0.0.0/8
```

To view prefix list information, use the following command:

```
G8272# display ip prefix list
ip prefix-list pref-list-35: 1 entries
    seq 5 permit 35.0.0.0/8
```

To enable or disable using IP prefix lists to filter BGP routes, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>
G8272(config-router-neighbor)# address-family {ipv4|ipv6} unicast
G8272(config-router-neighbor-af)# [no] prefix-list <prefix list name> {in|out}
```

For example, to filter inbound BGP routes using the prefix list previously created (pref-list-35), use the following command:

```
G8272(config-router-neighbor-af)# prefix-list pref-list-35 in
```

Dynamic Capability

Dynamic Capability allows the dynamic update of capabilities over an established BGP session without disrupting the session.

By advertising Dynamic Capability to a peer in the open, a BGP speaker tells the peer that the speaker is capable of receiving and properly handling the CAPABILITY message from the peer after the BGP session has been established.

By default, Dynamic Capability is disabled. To enable or disable it on the switch, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>  
G8272(config-router-neighbor)# [no] dynamic-capability
```

BGP Graceful Restart

Usually when BGP on a router restarts, all the BGP peers detect that the session went down and then came up. This “down/up” transition is called *route flapping*. This causes BGP route re-computation, generation of BGP routing updates, and unnecessary churn to the forwarding tables. It can spread across multiple routing domains and create transient forwarding black holes and transient forwarding loops.

BGP graceful restart specifies a mechanism that allows a BGP speaker to preserve the forwarding state during BGP restart. BGP works in helper mode to help the connected router to perform a graceful restart.

The router can help the peer restart but it cannot preserve its own routing information when restarting. Thus, the peer neighbor cannot preserve the routing information advertised by a BGP router in only graceful restart helper mode.

By default, BGP graceful restart helper functionality is disabled. To enable or disable it on the switch, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# [no] graceful-restart-helper
```

To configure the maximum time (in seconds) to keep a restarting peer's stale routes, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# graceful-restart stalepath-time <1-3600>
```

By default, the maximum time the switch will keep a restarting peer's stale routes is 360 seconds. To reset this value to its default, use the following command:

```
G8272(config-router)# [no] graceful-restart stalepath-time
```

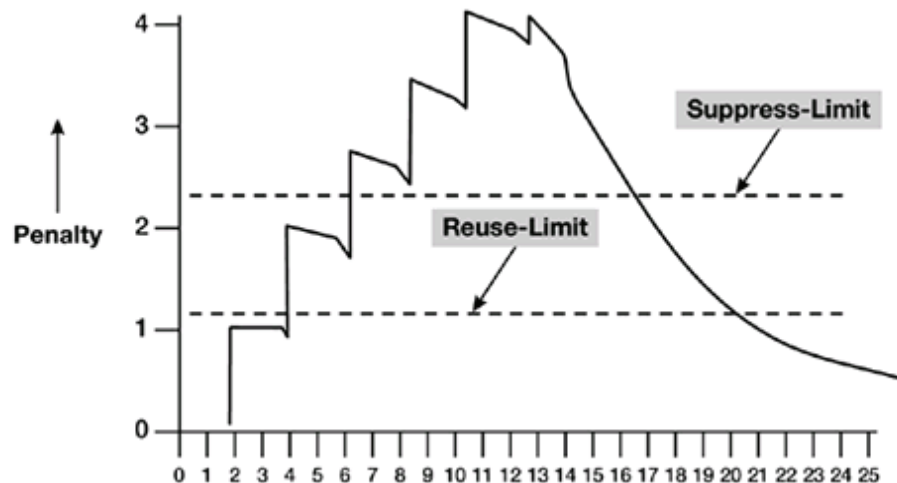
BGP Damping

Route flapping is a symptom of route instability. A route in the routing table may disappear and reappear intermittently when BGP sends a routing table, then withdraws it. A route that appears and disappears intermittently and repeatedly propagates BGP UPDATE and WITHDRAWN messages onto the Internet. If the prefix is constantly flapping, it can result in wasted CPU, bandwidth, and other network resources.

RFC 2439, *BGP Route Flap Dampening*, defines a mechanism by which the amount of routing state propagated via BGP can be reduced by reducing the scope of route flap propagation. BGP Route Flap Dampening categorizes routes as either well-behaved or ill-behaved. A well-behaved route shows a high degree of stability during an extended period of time (as determined by configurable limits). An ill-behaved route is no longer advertised; it is suppressed until there is an indication that the route has become stable.

Route flap dampening requires BGP speakers to maintain a history of the advertisement and withdraw of each BGP path implementations provide the capability to tune the suppress-limit, reuse-limit, and half-life values to control how long a misbehaving prefix is suppressed. The general operation of BGP dampening is depicted in [Figure 15](#).

Figure 15. BGP Route Flap Dampening



- A fixed penalty is assessed for each flap. A flap can be either a route withdraw or update with an attribute change.
- If the penalty exceeds the suppress-limit, the route is not advertised.
- The penalty is exponentially decayed based on the half-life value.
- Once the penalty is decayed below the reuse limit, the route is advertised.

The BGP flap dampening feature affects only external BGP routes (those in different ASes). You can also apply it within a confederation, between confederation member ASes. Because routing consistency within an AS is important, do not apply flap dampening to iBGP routes, for it will be ignored.)

A value for reachability half time and a different value for unreachability half time is supported. The unreachability half time value must be higher. In this case, when computing the maximum suppress penalty, BGP uses this higher unreachability value.

By default, BGP route flap dampening is disabled. To enable or disable it, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# address-family {ipv4|ipv6} unicast  
G8272(config-router-af)# [no] dampening [<half-life> [<reuse> <suppress> <maximum duration> [<unreachability half-life>]]|route-map <route map name>]
```

Note: If the **dampening** command is used without any additional parameters, the switch will assume the default values for those parameters. For more details about this command, see the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

Soft Reconfiguration Inbound

This feature triggers routing updates for the specified peer without resetting the session. You can use this option if you change an inbound route policy. Soft reconfiguration inbound saves a copy of all routes received from the peer before processing the routes through the inbound route policy. If you change the inbound route policy, Cloud NOS passes these stored routes through the modified inbound route policy to update the route table without tearing down existing peering sessions. Soft reconfiguration inbound can use significant memory resources to store the unfiltered BGP routes.

This feature is disabled by default. To enable or disable it for a BGP peer, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>  
G8272(config-router-neighbor)# address-family {ipv4|ipv6} unicast  
G8272(config-router-neighbor-af)# [no] soft-configuration inbound
```

BGP Route Refresh

BGP route refresh updates the inbound routing tables dynamically by sending route refresh requests to supporting peers when you change an inbound route policy. The remote BGP peer responds with a new copy of its routes that the local BGP speaker processes with the modified route policy.

The device supports both old and new route refresh capability. BGP peers advertise the route refresh capability as part of the BGP capability negotiation when establishing the BGP peer session.

To initiate a BGP route refresh, use the following command:

```
G8272# remove ip bgp <neighbor address> soft
```

BGP Multiple Address Families

You can manually configure the address families that BGP supports:

- Use the **address-family** command in BGP Configuration Mode to configure features for an address family.

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# address-family {ipv4|ipv6} unicast
G8272(config-router-af)#
```

- Use the **address-family** command in BGP Neighbor Configuration Mode to configure the specific address family for the neighbor.

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>
G8272(config-router-neighbor)# address-family {ipv4|ipv6} unicast
G8272(config-router-neighbor-af)#
```

You must configure address families if you are using route redistribution, address aggregation and other advanced features.

Note: Only IPv4 unicast is supported.

BGP BFD

Bidirectional forwarding detection (BFD) is a detection protocol designed to provide fast forwarding path failure detection times for media types, encapsulations, topologies, and routing protocols. You can use BFD to detect forwarding path failures at a uniform rate, rather than the variable rates for different protocol hello mechanisms. BFD makes network profiling and planning easier and convergence time consistent and predictable.

BFD provides subsecond failure detection between two BGP peers.

By default, BFD is disabled for BGP peers. To enable or disable it for a BGP peer, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>
G8272(config-router)# neighbor <neighbor address> remote-as <AS number>
G8272(config-router-neighbor)# [no] bfd
```

If the BGP peer is multiple hops away from the switch, use the following command to enable or disable BFD for that peer:

```
G8272(config-router-neighbor)# [no] bfd multihop
```

BGP Next Hop Tracking

Next hop tracking (NHT) is used to notify the BGP process asynchronously whenever there is any change in the IGP routes. NHT reduces the convergence time of BGP routes when IGP routes are changed.

The BGP process can wait for a configured or a default time after receiving a next hop change trigger from NSM, before it updates the BGP RIB. You can set and unset the time delay.

This feature is enabled by default. To configure BGP next hop tracking, use the following command:

```
G8272(config)# routing-protocol bgp <AS number (1-4294967295)>  
G8272(config-router)# address-family {ipv4|ipv6} unicast  
G8272(config-router-af)# nexthop trigger-delay critical <delay value  
(1-4294967295 milliseconds)> non-critical <delay value (1-4294967295 milliseconds)>
```

To disable BGP next hop tracking, use the following command:

```
G8272(config-router-af)# nexthop trigger-delay
```

BGP Tuning

The following commands enable you to tune BGP protocols in Routing-protocol BGP Neighbor Address Family Configuration Mode:

- **allowas-in**
- **as-origination-interval**
- **attribute-unchanged**
- **connection-retry-time**
- **description**
- **disallow-infinite-holdtime**
- **enforce-first-as**
- **fall-over**
- **fast-external-fallover**
- **log-neighbor-changes**
- **synchronization**
- **timer**
- **transport connection-mode passive**

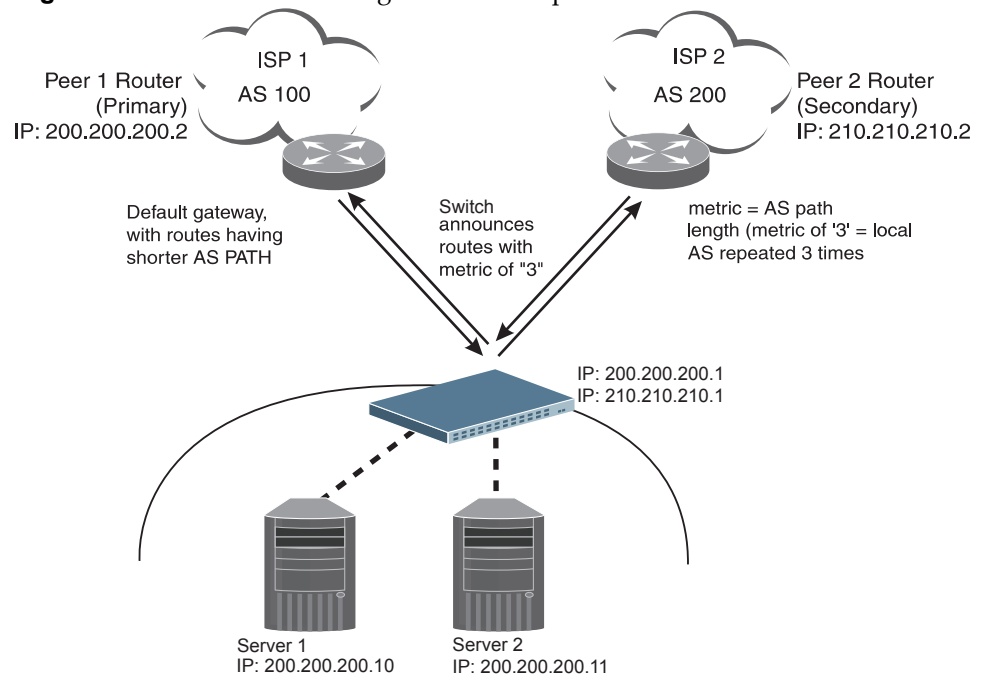
For more information about these commands, see the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

BGP Failover Configuration

Use the following example to create redundant default gateways for a G8272 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 16](#), 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 G8272.

Figure 16. BGP Failover Configuration Example



On the G8272, 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 and IP interfaces using 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.

```
G8272(config)# interface vlan 1
G8272(config-if)# ip address 200.200.200.1/24
G8272(config-if)# exit

G8272(config)# interface vlan 2
G8272(config-if)# ip address 200.200.200.1/24
G8272(config-if)# exit
```

2. Configure BGP peer router 1 and 2 with IPv4 addresses.

```
G8272(config)# routing-protocol bgp 1
G8272(config-router)# neighbor 200.200.200.2 remote-as 100
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit
G8272(config-router)# exit

G8272(config)# routing-protocol bgp 2
G8272(config-router)# neighbor 210.210.210.2 remote-as 200
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit
G8272(config-router)# exit
```

3. Verify the configuration:

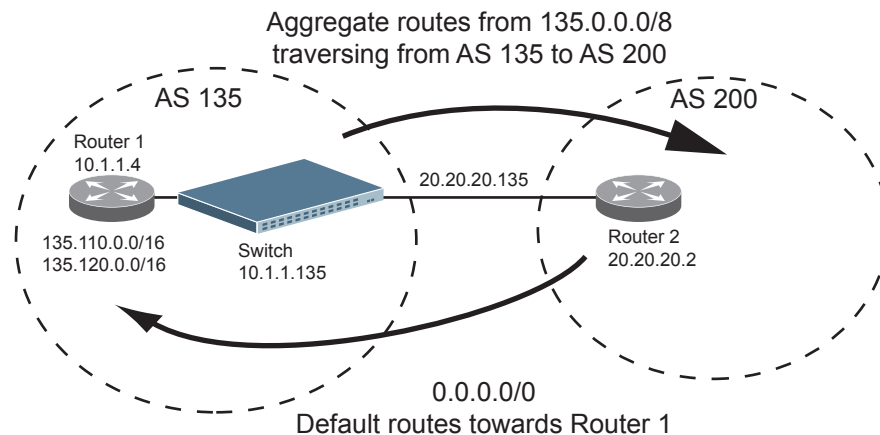
```
G8272# display ip bgp neighbors
```


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 17](#), you have two routers: one in AS 135 and the other in AS 200. Configure the G8272 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 17. Route Aggregation and Default Route Redistribution



1. Configure the IP interfaces.
2. Configure the AS number (AS 135) and router ID (10.1.1.135) on the switch.

```
G8272(config)# router-id 10.1.1.135
G8272(config)# routing-protocol bgp 135
G8272(config-router)#
```

3. Configure Router 1 with IPv4 addresses.

```
G8272(config)# router bgp 135
G8272(config-router)# router-id 10.1.1.4
```

4. Configure redistribution for Router 1.

```
G8272(config-router)# address-family ipv4 unicast
G8272(config-router-af)# redistribute static
```

5. Configure aggregation policy control on Router 1 by configuring the IPv4 routes that you want aggregated.

```
G8272(config-router-af)# aggregate-address 135.0.0.0/8
G8272(config-router-af)# exit
```

6. Configure Router 2 with IPv4 addresses.

```
G8272(config)# routing-protocol bgp 200  
G8272(config-router)# router-id 20.20.20.2
```

7. Configure redistribution for Router 2.

```
G8272(config-router)# address-family ipv4 unicast  
G8272(config-router-af)# redistribute static
```

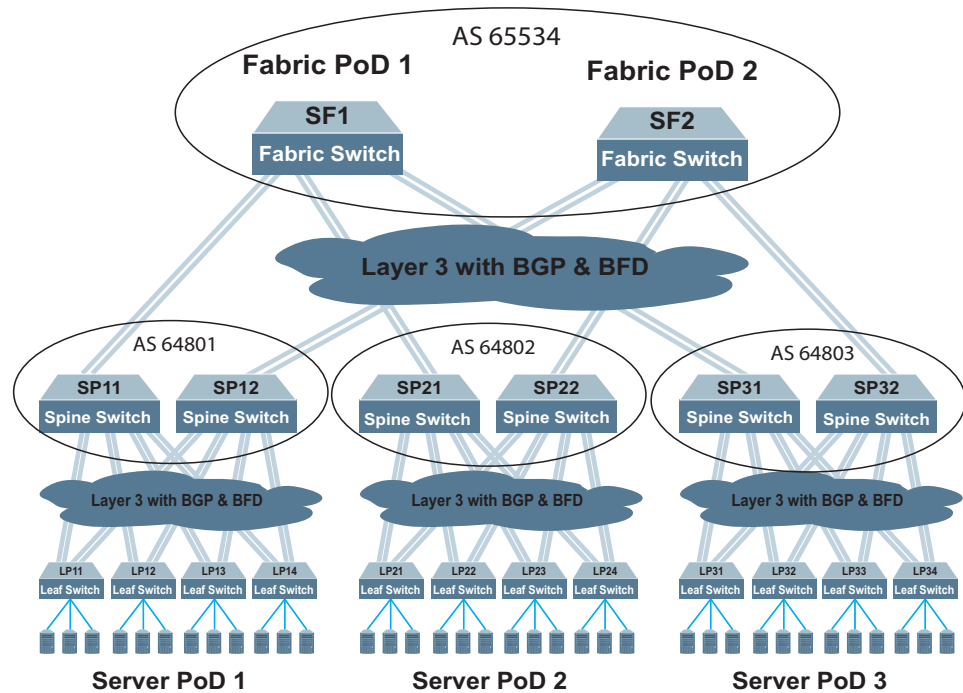
Designing a Clos Network Using BGP

A Clos network is a type of non-blocking, multistage switching architecture that reduces the number of ports required in an interconnected fabric. Instead of a hierarchically oversubscribed system of clusters, a Clos network turns your configuration into a high-performance network.

Clos networks are named after Bell Labs researcher Charles Clos, who determined that throughput is increased in a switching array or fabric if the switches are organized in a leaf-spine hierarchy.

In a leaf-spine topology, a series of leaf switches that form the access layer are fully meshed to a series of spine switches. [Figure 18](#) shows a sample leaf-spine topology with a three-stage multi-Point of Delivery (PoD) L3 Clos network.

Figure 18. Clos Network Topology



Each fabric switch (SF1 and SF2) serves as a fabric PoD that connects to one of the spine switches in each server PoD using BGP and Bidirectional Forwarding Detection (BFD). The spine switches each connect to four leaf switches using BGP and BFD. Using this kind of configuration, you can add capacity to your network by adding another server PoD instead of reconfiguring your existing infrastructure.

For more information on BFD, see [“Bidirectional Forwarding Detection” on page 272](#).

Clos Network BGP Configuration Example

In this example, from each tier of the network topology, a single switch configuration is covered:

- for the Fabric tier: fabric switch SF1
- for the Spine tier: spine switch SP11
- for the Leaf tier: leaf switch LP11

The configurations for the rest of the switches in each tier are similar to the those presented. Only the IP addresses of the switch interfaces, the neighbor addresses and AS numbers of the BGP peers are different.

Autonomous system (AS) membership is as follows:

- AS 65534 includes:
 - SF1
 - SF2
- AS 64801 includes:
 - SP11
 - SP12
- AS 64802 includes:
 - SP21
 - SP22
- AS 64803 includes:
 - SP31
 - SP32
- AS 64711 includes LP11
- AS 64712 includes LP12
- AS 64713 includes LP13
- AS 64714 includes LP14
- AS 64721 includes LP21
- AS 64722 includes LP22
- AS 64723 includes LP23
- AS 64724 includes LP24
- AS 64731 includes LP31
- AS 64732 includes LP32
- AS 64733 includes LP33
- AS 64734 includes LP34

To configure BGP for the Clos Network topology shown in [Figure 18 on page 379](#), follow the steps below.

Configure Fabric Switch SF1

1. Configure the switch ethernet interfaces that are connected to its BGP peers (SP11, SP21 and SP31) as routing ports and assign them IP addresses.

```
G8272(config)# interface ethernet 1/14
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 12.3.1.254/24
G8272(config-if)# exit

G8272(config)# interface ethernet 1/49                (connected to SP11)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 30.111.1.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/50                (connected to SP11)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 30.111.2.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/51                (connected to SP21)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 30.111.3.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/52                (connected to SP21)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 30.111.4.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/53                (connected to SP31)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 30.111.5.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/54                (connected to SP31)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 30.111.6.1/30
G8272(config-if)# exit
```

2. Configure a loopback interface on the switch as a routing port and define its control plane service policy:

```
G8272(config)# interface loopback 0
G8272(config-if)# no bridge-port
G8272(config-if)# service-policy input copp-system-policy
G8272(config-if)# exit
```

3. Configure the BGP AS number for the switch, the path selection method and the maximum eBGP paths:

```
G8272(config)# routing-protocol bgp 65534
G8272(config-router)# bestpath as-path multipath-relax
G8272(config-router)# address-family ipv4 unicast
G8272(config-router-af)# maximum-paths ebgp 32
G8272(config-router-af)# exit
```

4. Define its BGP peers:

```
G8272(config-router)# neighbor 12.3.1.1 remote-as 65123
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 30.111.1.2 remote-as 64801 (BGP peer SP11)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 30.111.2.2 remote-as 64801 (BGP peer SP11)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 30.111.3.2 remote-as 64802 (BGP peer SP21)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 30.111.4.2 remote-as 64802 (BGP peer SP21)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 30.111.5.2 remote-as 64801 (BGP peer SP31)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 30.111.6.2 remote-as 64801 (BGP peer SP31)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit
```

Configure Spine Switch SP11

1. Configure the switch ethernet interfaces that are connected to its BGP peers (SF2, LP11, LP12, LP13 and LP14) as routing ports and assign them IP addresses.

```
G8272(config)# interface ethernet 1/1 (connected to LP11)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.111.1.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/2 (connected to LP11)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.111.2.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/3 (connected to LP12)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.111.3.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/4 (connected to LP12)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.111.4.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/5 (connected to LP13)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.111.5.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/6 (connected to LP13)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.111.6.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/7 (connected to LP14)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.111.7.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/8 (connected to LP14)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.111.8.1/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/14
G8272(config-if)# no bridge-port
G8272(config-if)# exit

G8272(config)# interface ethernet 1/49 (connected to SF1)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 30.111.1.2/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/50 (connected to SF1)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 30.111.2.2/30
G8272(config-if)# exit
```

2. Configure a loopback interface on the switch as a routing port and define its control plane service policy:

```
G8272(config)# interface loopback 0
G8272(config-if)# no bridge-port
G8272(config-if)# service-policy input copp-system-policy
G8272(config-if)# exit
```

3. Configure the BGP AS number for the switch, the path selection method and the maximum eBGP paths:

```
G8272(config)# routing-protocol bgp 64801
G8272(config-router)# bestpath as-path multipath-relax
G8272(config-router)# address-family ipv4 unicast
G8272(config-router-af)# maximum-paths ebgp 32
G8272(config-router-af)# exit
```

4. Define its BGP peers:

```
G8272(config-router)# neighbor 20.111.1.2 remote-as 64711 (BGP peer LP11)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 20.111.2.2 remote-as 64711 (BGP peer LP11)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 20.112.3.2 remote-as 64712 (BGP peer LP12)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 20.112.4.2 remote-as 64712 (BGP peer LP12)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 20.113.5.2 remote-as 64713 (BGP peer LP13)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 20.113.6.2 remote-as 64713 (BGP peer LP13)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit
```



```
G8272(config-router)# neighbor 20.114.7.2 remote-as 64714 (BGP peer LP14)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 30.114.8.2 remote-as 64714 (BGP peer LP14)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 30.111.1.1 remote-as 65534 (BGP peer SF1)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 30.111.2.1 remote-as 65534 (BGP peer SF1)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit
```

Configure Leaf Switch LP11

1. Configure the switch ethernet interfaces that are connected to its BGP peers (SP11 and SP12) as routing ports and assign them IP addresses.

```
G8272(config)# interface ethernet 1/1 (connected to SP11)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.111.1.2/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/2 (connected to SP11)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.111.2.2/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/3 (connected to SP21)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.211.1.2/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/4 (connected to SP21)
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 20.211.2.2/30
G8272(config-if)# exit

G8272(config)# interface ethernet 1/14
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 9.3.11.254/24
G8272(config-if)# exit
```

2. Configure a loopback interface on the switch as a routing port and define its control plane service policy:

```
G8272(config)# interface loopback 0
G8272(config-if)# no bridge-port
G8272(config-if)# service-policy input copp-system-policy
G8272(config-if)# exit
```

3. Configure the BGP AS number for the switch, the path selection method and the maximum eBGP paths:

```
G8272(config)# routing-protocol bgp 64711
G8272(config-router)# bestpath as-path multipath-relax
G8272(config-router)# address-family ipv4 unicast
G8272(config-router-af)# maximum-paths ebgp 32
G8272(config-router-af)# exit
```

4. Define its BGP peers:

```
G8272(config-router)# neighbor 9.3.11.1 remote-as 65093
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 20.111.1.1 remote-as 64801 (BGP peer SP11)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 20.111.2.1 remote-as 64801 (BGP peer SP11)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 20.211.1.1 remote-as 64801 (BGP peer SP21)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit

G8272(config-router)# neighbor 20.211.2.1 remote-as 64801 (BGP peer SP21)
G8272(config-router-neighbor)# advertisement-interval 0
G8272(config-router-neighbor)# bfd
G8272(config-router-neighbor)# address-family ipv4 unicast
G8272(config-router-neighbor-af)# exit
G8272(config-router-neighbor)# exit
```

Chapter 20. Open Shortest Path First

Cloud Network OS supports the Open Shortest Path First (OSPF) routing protocol. The Cloud NOS implementation conforms to the OSPF version 2 specifications detailed in Internet RFC 2328. The following sections discuss OSPF support for the RackSwitch G8272:

- [“OSPFv2 Overview” on page 390](#). 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 Cloud NOS” on page 395](#). This section describes how OSPFv2 is implemented in CNOS, such as configuration parameters, electing the designated router, summarizing routes, defining route maps and so forth.
- [“OSPFv2 Configuration Examples” on page 404](#). This section provides step-by-step instructions on configuring different OSPFv2 examples:
 - Creating a simple OSPF domain
 - Creating virtual links
 - Summarizing routes

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.

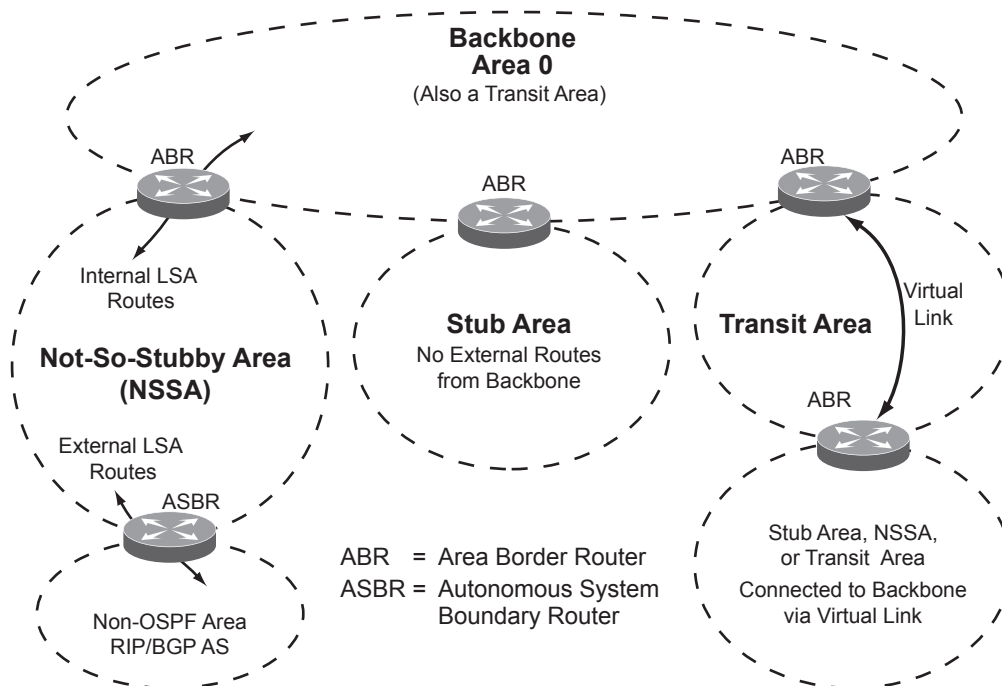
Types of OSPF Areas

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 19](#), 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 19. OSPF Area Types

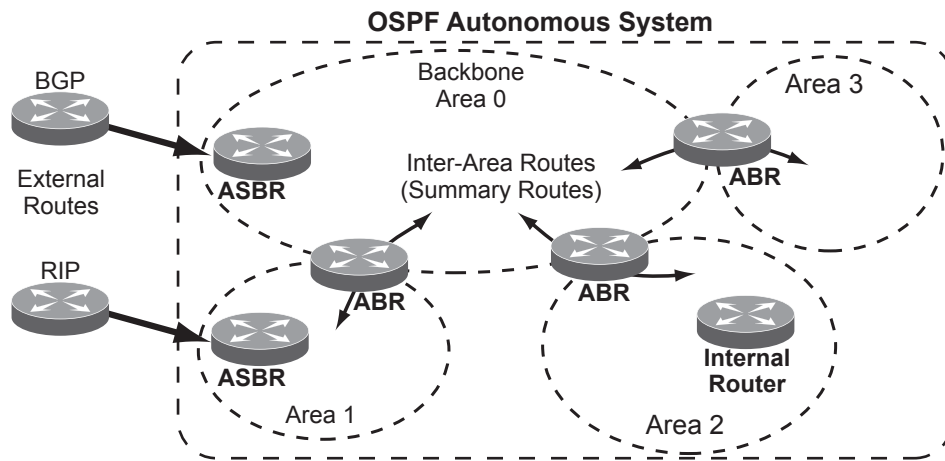


Types of OSPF Routing Devices

As shown in [Figure 20](#), 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 20. 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 Cloud NOS

CNOS supports a single instance of OSPF and up to 8k routes on the network. The following sections describe OSPF implementation in CNOS:

- [“Configurable Parameters” on page 395](#)
- [“Defining Areas” on page 396](#)
- [“Interface Cost” on page 397](#)
- [“Electing the Designated Router and Backup” on page 397](#)
- [“Summarizing Routes” on page 397](#)
- [“Default Routes” on page 398](#)
- [“Virtual Links” on page 399](#)
- [“Router ID” on page 400](#)
- [“Authentication” on page 400](#)

Configurable Parameters

In CNOS, OSPF parameters can be configured through the Industry Standard Command Line Interfaces (ISCLI), 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:

- OSPF traps—Traps produce messages upon certain events or error conditions, such as missing a hello, failing a neighbor, or recalculating the SPF.
- Link-State Database size—The size of the LSA database can be specified to help manage the memory resources on the switch.
- 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 G8272 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 399).

Up to 20 OSPF areas can be connected to the G8272 with CNOS 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 an *area ID*. The command to define and enable an OSPF area is as follows:

```
G8272(config)# ip router ospf 0 area <area-id>
```

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 G8272 as “`area 0.0.0.2`”.

On the G8272, using the last octet in the area ID, “area 1” is equivalent to “area 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 ID to an IP interface that participates in the area. The format for the command is as follows:

```
G8272(config)# interface <interface name or type>
G8272(config-if)# no bridge-port
G8272(config-if)# ip address <IP address/mask>
G8272(config-if)# ip router ospf 0 area <area-id> [secondaries none]
```

For example, the following commands could be used to add a network area 0.0.0.10 in OSPFv2 instance 200:

```
G8272(config)# interface ethernet 1/2
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 192.0.2.1/16
G8272(config-if)# ip router ospf 0 area 0.0.0.10
G8272(config-if)# exit
```

Note: OSPFv2 supports IPv4 only.

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):

```
G8272(config-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 than 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 G8272. The command is as follows:

```
G8272(config-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:

```
G8272(config-if)# ip ospf passive-interface
G8272(config-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:

```
G8272(config)# routing-protocol ospf
G8272(config-router)# area <area-id> range {<IP address> <netmask>|<IP address>/
<prefix length>}
```

where <area-id> is a number from 0 to 4294967295, <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 409.

A filter can be configured to advertise summary routes on an Area Border Router (ABR). This command suppresses incoming and outgoing summary routes between a specific area and the others:

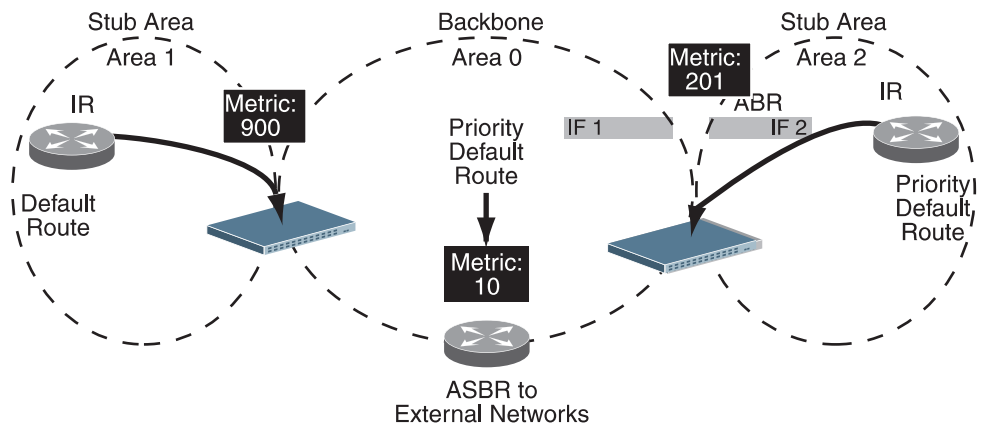
```
G8272(config-router)# area <area id> filter-list route-map <map name>
{in|out}
```

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 G8272 acting as an ABR automatically inserts a default route into stub. In simple OSPF stub areas with only one ABR leading upstream (see Area 1 in [Figure 21](#)), 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 21. 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 (Routing-protocol OSPF mode):

```
G8272(config-router)# default-information originate [always] [metric
<metric value>] [metric-type <1 or 2>] [route-map <map-name>]
```

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:

```
G8272(config-router)# no default-information originate
```

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 19 on page 391](#)).

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.

To set a link between two backbone areas that are physically separated through other non backbone area, use the following command:

```
G8272(config-router)# area <area-id> virtual link <IP address>
```

To configure an authentication type and key between virtual-link neighbors, use the following command:

```
G8272(config-router-vlink)# authentication [message-digest|null]
G8272(config-router-vlink)# authentication [0] <string>
G8272(config-router-vlink)# message-digest-key <key id> md5 [0] <string>
```

where *message-digest* enables MD5 authentication and *<string>* is the password string.

For a detailed configuration example on Virtual Links, see [“Example 2: Virtual Links” on page 406](#).

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 G8272 loopback interface.

The router ID can be configured in one of the following two ways:

- Dynamically—OSPF protocol configures the highest 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:

```
G8272(config-router)# 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 **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, delete the Router ID using the **no router-id** command and re-initialize the OSPF process.
- To view the router ID, use the following command:

```
G8272(config-router)# display router-id
```

Authentication

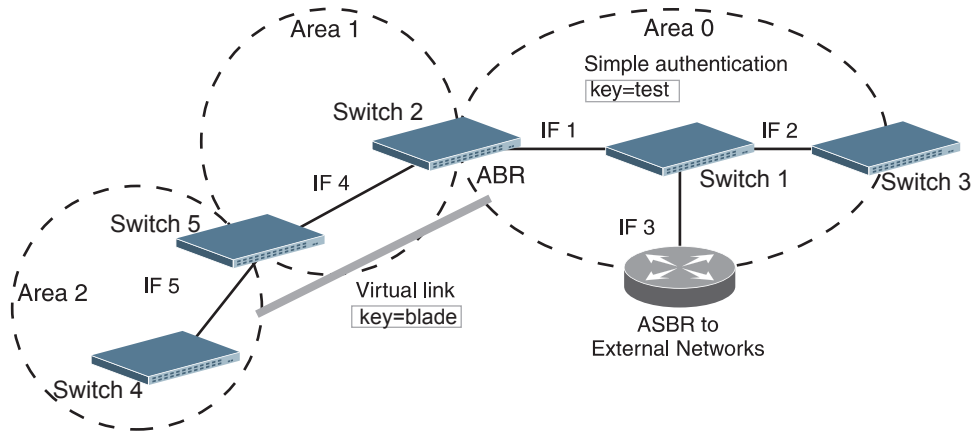
OSPF protocol exchanges can be authenticated so that only trusted routing devices can participate. This ensures less processing on routing devices that are not listening to OSPF packets.

OSPF allows packet authentication and uses IP multicast when sending and receiving packets. Routers participate in routing domains based on pre-defined passwords. CNOS 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 22 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 22. OSPF Authentication



Configuring Plain Text OSPF Passwords

To configure simple plain text OSPF passwords on the switches shown in Figure 22 use the following commands:

1. Enable OSPF authentication for Area 0 on switches 1, 2, and 3.

```
G8272(config-router)# area 0 authentication
G8272(config-router)# exit
```

2. Configure a simple text password up to eight characters for each OSPF interface in Area 0 on switches 1, 2, and 3.

```
G8272(config)# interface vlan 1
G8272(config-if)# ip ospf authentication-key 0 test
G8272(config-if)# exit

G8272(config)# interface vlan 2
G8272(config-if)# ip ospf authentication-key 0 test
G8272(config-if)# exit

G8272(config)# interface vlan 3
G8272(config-if)# ip ospf authentication-key 0 test
G8272(config-if)# exit
```

3. Enable OSPF authentication for Area 2 on switch 4.

```
G8272(config)# routing-protocol ospf
G8272(config-router)# area 2 authentication
```

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.

```
G8272(config)# routing-protocol ospf
G8272(config-router)# area 2 virtual-link 1.2.3.4
G8272(config-router-vlink)# authentication
G8272(config-router-vlink)# authentication-key 0 blade
```

Configuring MD5 Authentication

Use the following commands to configure MD5 authentication on the switches shown in [Figure 22](#):

1. Enable OSPF MD5 authentication for Area 0 on switches 1, 2, and 3.

```
G8272(config)# routing-protocol ospf
G8272(config-router)# area 0 authentication message-digest
G8272(config-router)# exit
```

2. Assign MD5 key ID to OSPF interfaces on switches 1, 2, and 3.

```
G8272(config)# interface vlan 1
G8272(config-if)# ip ospf message-digest-key 11 Authkey
G8272(config-if)# exit

G8272(config)# interface vlan 2
G8272(config-if)# ip ospf message-digest-key 22 Authkey
G8272(config-if)# exit

G8272(config)# interface vlan 3
G8272(config-if)# ip ospf message-digest-key 33 Authkey
G8272(config-if)# exit
```

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 400](#)) 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:

```
G8272(config)# interface loopback <0-7>
G8272(config-if)# ip router ospf 0 area <area ID>
G8272(config-if)# exit
```

Graceful Restart Helper

The N/OS design provides a complete separation of its control plane from the forwarding plane, thus allowing the restart or upgrade of control plane software without disturbing forwarding. Such a restart/upgrade is called *graceful restart*.

In graceful restart, a restarting router is helped by neighbors by announcing links in their LSAs. These neighbors are considered to be in *helper mode* for the duration (grace period) of graceful restart.

The feature is enabled by default. To disable it, use the following command:

```
G8272(config)# graceful-restart ospf helper never
```

OSPFv2 Configuration Examples

A summary of the basic steps for configuring OSPF on the G8272 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.

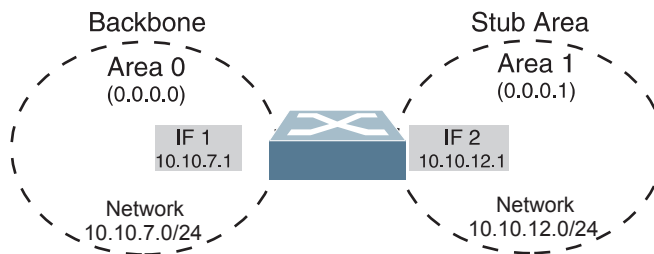
7. (Optional) Configure virtual links.

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 23. A Simple OSPF Domain



Follow this procedure to configure OSPF support as shown in [Figure 23](#):

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

```
G8272(config)# interface vlan 1
G8272(config-if)# ip address 10.10.7.1/24
G8272(config-if)# exit

G8272(config)# interface vlan 2
G8272(config-if)# ip address 10.10.12.1/24
G8272(config-if)# exit
```

Note: OSPFv2 supports IPv4 only.

2. Enable OSPF.

```
G8272(config)# routing-protocol ospf
G8272(config-router)#
```

3. Define the stub area.

```
G8272(config-router)# area 1 stub
G8272(config-router)# exit
```

4. Attach the network interface to the backbone.

```
G8272(config)# interface vlan 1
G8272(config-if)# ip router ospf 0 area 0
G8272(config-if)# exit
```

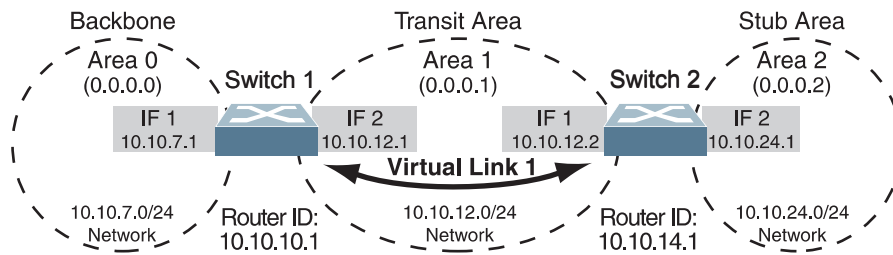
- Attach the network interface to the stub area.

```
G8272(config)# interface vlan 2
G8272(config-if)# ip router ospf 0 area 1
G8272(config-if)# exit
```

Example 2: Virtual Links

In the example shown in [Figure 24](#), 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 24. Configuring a Virtual Link



Note: OSPFv2 supports IPv4 only.

Configuring OSPF for a Virtual Link on Switch #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

```
G8272(config)# interface vlan 1
G8272(config-if)# ip address 10.10.7.1
G8272(config-if)# exit

G8272(config)# interface vlan 2
G8272(config-if)# ip address 10.10.12.1
G8272(config-if)# exit
```

- 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.

```
G8272(config)# routing-protocol ospf
G8272(config-router)# router-id 10.10.10.1
```

3. Attach the network interface to the backbone.

```
G8272(config)# interface vlan 1
G8272(config-if)# ip router ospf 0 area 0
G8272(config-if)# exit
```

4. Attach the network interface to the transit area.

```
G8272(config)# interface vlan 2
G8272(config-if)# ip router ospf 0 area 1
G8272(config-if)# exit
```

5. 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 [Step 2 on page 407](#).

```
G8272(config)# routing-protocol ospf
G8272(config-router)# area 1 virtual-link 10.10.14.1
```

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
- Interface 2 for the stub area network on 10.10.24.0/24

```
G8272(config)# interface vlan 1
G8272(config-ip-if)# ip address 10.10.12.2/24
G8272(config-ip-if)# vlan 10
G8272(config-ip-if)# exit
G8272(config)# interface vlan 2
G8272(config-ip-if)# ip address 10.10.24.1/24
G8272(config-ip-if)# vlan 20
G8272(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 [Step 5 on page 407](#).

```
G8272(config-router)# router-id 10.10.14.1
```

3. Enable OSPF.

```
G8272(config)# routing-protocol ospf
```

4. Define the stub area.

```
G8272(config-router)# area 2 stub
G8272(config-router)# exit
```

5. Attach the network interface to the transmit area.

```
G8272(config)# interface vlan 1
G8272(config-ip-if)# ip router ospf 0 area 1
G8272(config-ip-if)# exit
```

6. Attach the network interface to the stub area.

```
G8272(config)# interface vlan 2
G8272(config-ip-if)# ip router ospf 0 area 2
G8272(config-ip-if)# exit
```

7. 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 406](#).

```
G8272(config)# routing-protocol ospf
G8272(config-router)# area 1 virtual-link 10.10.10.1
```

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.

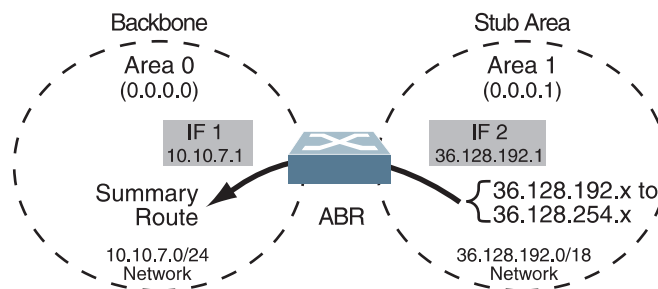
Example 3: Summarizing Routes

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 25. Summarizing Routes



Notes:

- OSPFv2 supports IPv4 only.
- 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 25](#):

1. Configure IP interfaces for each network which will be attached to OSPF areas.

```
G8272(config)# interface vlan 1
G8272(config-if)# ip address 10.10.7.1
G8272(config-if)# vlan 10
G8272(config-if)# exit

G8272(config)# interface vlan 2
G8272(config-if)# ip address 36.128.192.1
G8272(config-if)# vlan 20
G8272(config-if)# exit
```

2. Enable OSPF.

```
G8272(config)# routing-protocol ospf
```

3. Attach the network interface to the backbone.

```
G8272(config)# interface vlan 1
G8272(config-if)# ip router ospf 0 area 0
G8272(config-if)# exit
```

4. Attach the network interface to the stub area.

```
G8272(config)# interface vlan 2
G8272(config-if)# ip router ospf 0 area 1
G8272(config-if)# exit
```

5. Configure route summarization by specifying the starting address and mask of the range of addresses to be summarized.

```
G8272(config)# routing-protocol ospf
G8272(config-router)# area 1 range 36.128.192/18
G8272(config-router)# exit
```

6. Use the hide command to prevent a range of addresses from advertising to the backbone.

```
G8272(config)# routing-protocol ospf
G8272(config-router)# area 1 range 36.128.200.0/24 not-advertise
G8272(config-router)# exit
```

Verifying OSPF Configuration

Use the following commands to verify the OSPF configuration on your switch:

- **display ip ospf**
- **display ip ospf border-routers**
- **display ip ospf neighbor**
- **display ip ospf database**

Refer to the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1* for information on the preceding commands.

Chapter 21. Route Maps

Route maps are used to define route policy by permitting or denying certain routes based on configured set of rules. Each route map consists of multiple clauses that are ordered by their sequence number.

This section discusses the following topics:

- [“Route Maps Overview” on page 412](#)
- [“Permit and Deny Rules” on page 413](#)
- [“Match and Apply Clauses” on page 414](#)
- [“Route Maps Configuration Example” on page 417](#)

Route Maps Overview

Route maps define conditions for redistributing routes from one routing protocol to another. Also, route maps are widely used by routing protocols such as BGP or OSPF, which apply route maps to filter traffic.

The more common applications of route maps are listed below:

- route filtering during redistribution between routing protocols;
- route control and attribute modification of BGP neighbors;
- route metric modification during redistribution between routing protocols.

Route maps can be thought as conditional statements that the switch will consider before examining its routing table. They take precedence over routing protocols.

For example, you can define and use a route map when you want traffic to follow a certain path across the network. This path can be different from the usual path that the routing protocol forwards traffic on. The switch will examine the route map and if traffic meets the specified conditions, it will forward traffic across the selected network path. If traffic doesn't meet the conditions, it will be forwarded according to the routing table.

To create or modify a route map, use the following command:

```
G8272(config)# route-map <route map name>  
G8272(config-route-map)#
```

Note: After creating a route map, you will enter Route Map Configuration mode.

Note: By default, if the rule (permit or deny) isn't specified, the switch will configure the route map as a permit rule. Also, if the sequence number isn't specified, the switch will configure sequence 10.

To view information about a route map, use the following command:

```
G8272(config)# display route-map <route map name>
```

To delete a route map, use the following command:

```
G8272(config)# no route-map <route map name>
```

Permit and Deny Rules

Route maps contain conditional clauses (match and apply) that can be used to permit or deny routes.

A permit rule means that a route must meet all the match clauses to be allowed. A deny rule means that a route that meets at least one of the match clauses will be dropped.

For example, route map is configured to use a prefix-list as a match clause. If the route map is set as a permit rule, then routes that are allowed by the prefix-list will be redistributed.

If the route map is set as a deny rule, then routes that are allowed by the prefix-list will not be redistributed.

To configure a route map permit rule, use the following command:

```
G8272(config)# route-map <route map name> permit
G8272(config-route-map)#
```

Note: After running the above command, you will enter Route Map Configuration mode, where you can create match and apply clauses that will be used to check if a route will be redistributed or not.

Rules are configured with a sequence number. The switch will check routes against permit or deny rules in ascending order of their sequence number. Rules with a lower sequence number will be checked before rules with a higher sequence number. A route map can have up to a maximum of 65535 sequences.

Note: If the sequence number isn't specified, the switch will configure match and apply clauses for sequence 10.

Note: It is recommended that you use multiples of 10 when configuring sequences, in case you require to add more rules between already existing ones.

To configure a route map permit rule and specify its sequence number, use the following command:

```
G8272(config)# route-map <route map name> permit <1-65535 (sequence number)>
```

To configure a route map deny rule and specify its sequence number, use the following command:

```
G8272(config)# route-map <route map name> deny <1-65535>
```

Note: A route map cannot have two rules that share the same sequence number. The newer rule will replace the older one.

For example, sequence 20 is configured as a deny rule. If you then set it up as a permit rule, the new configuration will overwrite the old one. Thus, sequence 20 will become a permit rule.

To delete a rule (permit or deny), use the following command:

```
G8272(config)# no route-map <route map name> {permit|deny} <1-65535>
```

Match and Apply Clauses

A rule consists of two types of clauses: match and apply. You can consider these clauses in terms of 'IF' and 'THEN' statements from many computer programming languages. If a match clause is true, then an apply clause is executed.

To configure a clause, you must enter Route Map Configuration mode. To achieve this, use the following command:

```
G8272(config)# route-map <route map name> {deny|permit} <1-65535 (sequence number)>
G8272(config-route-map)#
```

Match clauses or conditions are used to create permit or deny rules. Only routes that meet the specified match clauses will be redistributed or rejected, depending on the type of the rule.

If multiple rules are created, the switch checks every rule in ascending order of their sequence number. If the first match clause isn't met by the route, the clause is not applicable to the route and the switch checks the route against the next match clause.

This goes on until the route meets the conditions of a rule or until the switch has checked all sequences. If a route didn't meet all the conditions of a single permit rule, it will be rejected. If a route didn't meet any of the conditions of any deny rule, it will be redistributed.

Note: Multiple match or apply clauses can be configured for each rule. Each clause will be checked in the order they were configured.

To add or remove a match clause, use the following command:

```
G8272(config-route-map)# [no] match <condition>
```

You can create a match clause for the conditions:

- match BGP AS path list;
- match BGP community list;
- match BGP extended community list;
- match next hop interface;
- match destination IP address;
- match next hop IP address;
- match route metric;
- match BGP origin code;
- match route type;
- match tag value.

For example, the following command will create a match clause for routes that have a metric of 1500:

```
G8272(config-route-map)# match metric 1500
```

For more details about the match clause commands, consult the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

Note: Multiple match clauses of the same type can be configured for a single rule. When being checked against the match clauses, the route must meet only one of the configured match clauses of the that type.

For example, you if configure a permit rule with two match clauses for route metric (one for 1500 and another for 3000), a route must have either a metric of 1500 or 3000 to met the condition of the rule and be redistributed.

Note: If no match clauses are configured for a rule, all routes will meet the conditions.

Apply clauses modify the information of route when the match clauses are met.

To add or remove a apply clause, use the following command:

```
G8272(config-route-map)# [no] apply <condition>
```

When the route matches the conditions of the rule, you can modify the following parameters:

- set the route aggregator attribute;
- set the BGP AS path;
- set the route atomic aggregate attribute;
- delete the BGP community list;
- set the BGP community list;
- set the BGP extended community list;
- set the BGP route dampening attributes;
- set the IP address of the next hop;
- set the BGP local preference attribute;
- set the route metric;
- set the route metric type;
- set the BGP origin code;
- set the BGP originator ID;
- set the tag value;
- set the BGP route weight attribute.

For example, the following command will create a set clause that will modify the originator ID of a BGP route:

```
G8272(config-route-map)# apply originator-id 10.10.125.60
```

For more details about the match clause commands, consult the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

Note: If no apply clauses are configured for a permit rule, the route is redistributed without any modifications.

Route Maps Configuration Example

Following are the basic steps and commands for configuring route maps.

1. Configure a route map rule (for example, a permit rule for route map 'rmap'):

```
G8272(config)# route-map rmap permit 10
G8272(config-route-map)#
```

2. Configure a match clause for the rule (for example, match the first hop interface to ethernet port 1/10):

```
G8272(config-route-map)# match interface ethernet 1/10
```

3. Optionally, you can configure an apply clause for any matching routes (for example, set the routes next hop address):

```
G8272(config-route-map)# apply ip next-hop 120.78.33.78
```

4. Optionally, you can configure additional rules (for example, deny routes with destination IP addresses included in prefix list 'PrefDeny' or have a metric of 200):

```
G8272(config-route-map)# exit
G8272(config)# route-map rmap deny 20
G8272(config-route-map)# match ip address prefix-list PrefDeny
G8272(config-route-map)# match metric 200
G8272(config-route-map)# exit
```

5. Check the route map configuration:

```
G8272(config)# display route-map rmap
route-map rmap, permit, sequence 10
  Match clauses:
    interface: Ethernet1/10
  Set clauses:
    ip next-hop 120.78.33.78
route-map rmap, deny, sequence 20
  Match clauses:
    ip address prefix-list: PrefDeny
    metric: 200
  Set clauses:
```


Part 5: 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.

This section discusses the following topics:

- [“Basic Redundancy” on page 421](#)
- [“Virtual Router Redundancy Protocol” on page 425](#)

Chapter 22. Basic Redundancy

Cloud Network OS 10.1 includes various features for providing basic link or device redundancy:

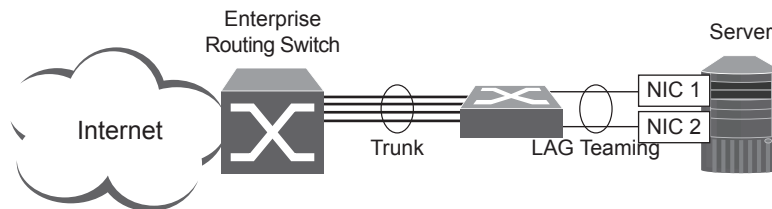
- [“Aggregating for Link Redundancy” on page 422](#)
- [“Virtual Link Aggregation” on page 423](#)

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 26](#), 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 ports on the server are in LAG teaming, allowing the secondary NIC to take over in the event that the primary NIC link fails.

Figure 26. Aggregating Ports for Link Redundancy



For more information on aggregation, see [Chapter 10, “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 12, “Virtual Link Aggregation Groups.”](#)

Chapter 23. Virtual Router Redundancy Protocol

The Lenovo RackSwitch G8272 supports high-availability network topologies through an enhanced implementation of the Virtual Router Redundancy Protocol (VRRP).

The following topics are discussed in this chapter:

- [“VRRP Overview” on page 426](#). This section discusses VRRP operation and Cloud NOS redundancy configurations.
- [“Failover Methods” on page 429](#). This section describes the three modes of high availability.
- [“Cloud NOS Extensions to VRRP” on page 430](#). This section describes VRRP enhancements implemented in CNOS.
- [“Virtual Router Deployment Considerations” on page 436](#). This section describes issues to consider when deploying virtual routers.
- [“Basic VRRP Configuration” on page 437](#). This section offers a simple example of a basic virtual router configuration consisting of two VRRP routers.
- [“High Availability Configuration” on page 439](#). This section discusses the more useful and easily deployed redundant configurations.

VRRP Overview

Typically, end devices in a LAN are connected to a larger network using a single router (first-hop router). They have the router's IP address configured as their default gateway. While, this method of configuration has its advantages, its main problem is that it creates a single point-of-failure if the first-hop router goes down.

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 offers a higher availability default path without the need for configuring dynamic routing or router discovery protocols on end devices.

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 IP address and ID number. One of the VRRP routers is elected as the master, based on a number of priority criteria, and assumes control of the shared virtual router IP address. If the master fails, one of the backup virtual routers will take control of the virtual router IP address and actively process traffic addressed to it.

With VRRP, Virtual Interface Routers (VIR) allow two VRRP routers to share an IP address across the routers. VIRs provide a single Destination IP (DIP) address for upstream routers to reach various servers and provide a virtual default Gateway for the servers. Also, end devices are configured to the virtual router's IP address as their default gateway.

Also, while using VRRP, you can configure two VRRP routers to evenly share traffic during normal running conditions. To achieve this, you assign the first VRRP router's virtual IP address as the default gateway for half of the end devices, and the second VRRP router's IP address as the default gateway for the other half of the end devices. In this scenario, both VRRP routers are members of two different virtual routers. Each VRRP router acts both as the master virtual router in one of the two virtual routers and the backup virtual router in the other virtual router.

Cloud NOS uses VRRP version 3 (VRRPv3), which implements support for IPv6 addresses. For more details, see RFC 5798.

Note: VRRP can be enabled only on Layer 3 interfaces, which on the G8272 are ethernet interfaces configured as routed ports or Switch Virtual Interfaces (SVIs).

By default, all ports are configured as access switch ports. To configure a switch port as a routed port, use the following command (for this example, ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# no bridge-port
```

For more details on routed ports, see [“Configuring a Routed Port” on page 290](#).

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 5798). 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 IP address. You can configure up to 255 virtual routers.

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) as defined in RFC 5798. 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 IP address owner. This router has the virtual router's IP address as its real interface address. This router responds to packets addressed to the virtual router's IP address for ICMP pings, TCP connections and so on.

There is no requirement for any VRRP router to be the IP address owner. Most VRRP installations choose not to implement an IP address owner. For the purposes of this chapter, VRRP routers that are not the IP 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 428](#) for an explanation of the selection process.

Note: If the IP 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 IP 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 interface across the routers. VIRs provide a single Destination IP (DIP) address for upstream routers to reach various destination networks and provide a virtual default Gateway.

Note: Every VIR must be assigned to an interface and every interface must be assigned to a VLAN. If no port in a VLAN has link up, the interface of that VLAN is down and if the interface of a VIR is down, that VIR goes into INIT state.

Assigning VRRP Virtual Router ID

When configuring a VRRP virtual router, you must assign it a virtual router identifier (ID). The virtual router ID may be configured as any number between 1 and 255. Use the following command to configure the virtual router ID:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# vrrp <1-255>
G8272(config-if-vrrp)#
```

Note: After creating a virtual router with the specified ID, you will enter the VRRP Configuration menu for that virtual router. Also, you can use the same command to configure any existing virtual routers.

Note: This is the configuration method for IPv4 addressing. For IPv6 addressing, see [“IPv6 VRRP” on page 434](#).

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 if it is the owner of the pinged IP address or has accept mode enabled. 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 IP 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 hold-off 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.

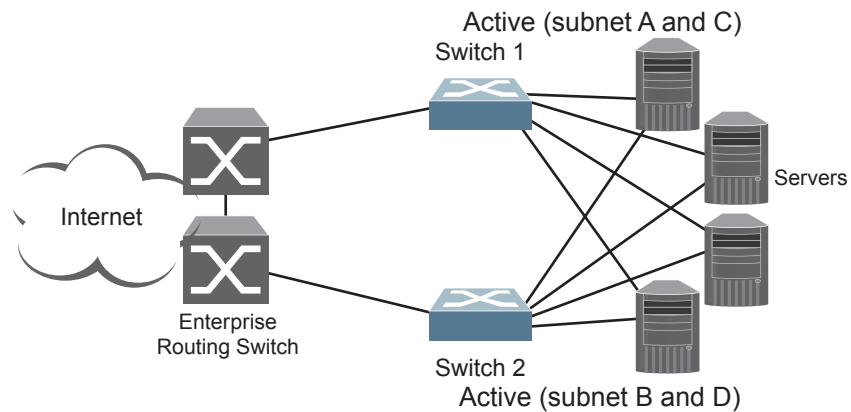
Failover Methods

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. Cloud NOS high availability configurations are based on VRRP. The implementation of VRRP includes proprietary extensions.

Active-Active Redundancy

In an active-active configuration, shown in [Figure 27](#), 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 more details and a configuration example, see [“High Availability Configuration”](#) on page 439.

Figure 27. Active-Active Redundancy



Cloud NOS Extensions to VRRP

This section describes VRRP enhancements that are implemented in CNOS:

- “VRRP Advertisement Interval and Sub-second Failover” on page 430
- “Interface Tracking” on page 431
- “Switch Back Delay” on page 431
- “Backward Compatibility with VRRPv2” on page 432
- “VRRP Accept Mode” on page 432
- “VRRP Preemption” on page 433
- “VRRP Priority” on page 433
- “IPv6 VRRP” on page 434

VRRP Advertisement Interval and Sub-second Failover

The master virtual router periodically sends VRRP advertisements to the backup virtual routers to keep them in their backup state. The master virtual router waits a certain amount of time between each advertisement. This period of time is called advertisement interval.

To configure the advertisement interval (in centiseconds or hundredths of a second), use the following command (for this example, interface ethernet 1/12 is used):

```
G8272(config)# interface ethernet 1/12  
G8272(config-if)# vrrp <1-255>  
G8272(config-if-vrrp)# advertisement-interval <5-4095>
```

The default value of the advertisement interval is 100 centiseconds or 1 second. To reset the advertisement interval to this value, use the following command:

```
G8272(config-if-vrrp)# no advertisement-interval
```

VRRPv3 supports the sub-second detection of master virtual router failure. To mitigate the potential loss of packets that happens in the time interval between when the master virtual router fails and when the backup virtual router takes its place, the advertisement interval can be configured to lower values than one second.

A backup virtual router detects that the master virtual router is down when three advertisement intervals have passed without the reception of any advertisements from the master. After the backup virtual router detects the failure, it will assume the role of master virtual router. During this time, any traffic that is sent to the downed master virtual router will be lost.

To lessen the volume of lost traffic, you can configure the advertisement interval to sub-second values, thus reducing the time required for the backup virtual router to detect the failure and take the master virtual router's place.

Interface Tracking

VRRP interface tracking allows the tracking of interface states which can modify the priority level of a VRRP router member of a virtual router group. If the tracked interface goes down or if it doesn't have a configured virtual IP address, the priority level of the virtual router group is lowered by a given value. This allows another VRRP router to become the new group master virtual router.

Only one interface can be tracked at a time for every virtual router group and the tracked interface must be a Layer 3 interface. VRRP doesn't support Layer 2 interface tracking.

Note: The interface on which you configure VRRP tracking can't be the interface that is being tracked. For example, ethernet interface 1/1 can be configured to track ethernet interface 1/2, 1/10 or 1/20, but not itself.

You can configure the value that is subtracted from a virtual router group's priority level when the state of the tracked interface changes to down. When the tracked interface returns to an up state, VRRP restores the original priority for the virtual router group.

By default, the tracking of interfaces is disabled.

To configure the subtracted value, use the following command (for this example, interfaces ethernet 1/10 and 1/12 are used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# vrrp <1-255>
G8272(config-if-vrrp)# track interface ethernet 1/10 priority <1-253>
```

Note: VRRP preempt mode must be enabled. For more details, see [“VRRP Preemption” on page 433](#).

To disable the tracking of an interface, use the following command:

```
G8272(config-if-vrrp)# no track interface ethernet 1/10 priority <1-253>
```

Switch Back Delay

Other Layer 3 routing protocols, like OSPF, have a slower convergence time than VRRP. When these routing protocols are enabled on the switch, there can be a situation where the route table of the master virtual router isn't completely populated when the VRRP router exchanges roles from master to backup virtual router and then vice versa. Remote network routes can be lost during the convergence time of these routing protocols, resulting in discarded traffic.

To avoid such a scenario, a time delay is used. The delay prevents the original master virtual router from becoming the master after it comes back online until a configured amount of time passes. This is called switch back delay.

By default, the switch back delay is disabled.

To configure the switch back delay (in milliseconds), use the following command (for this example, ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# vrrp <1-255>
G8272(config-if-vrrp)# switch-back-delay <1-50000>
```

To disable the switch back delay, use the following command:

```
G8272(config-if-vrrp)# no switch-back-delay
```

Backward Compatibility with VRRPv2

The G8272 uses VRRPv3, which supports IPv6 addresses, but can be configured to be compatible with VRRPv2 as well.

Note: This is only compatible with IPv4 addresses, since VRRPv2 doesn't have support for IPv6 addressing.

When VRRPv2 compatibility is enabled on the master virtual router, it sends both VRRPv2 and VRRPv3 advertisements at the configured advertisement interval. For more details about the advertisement interval, see [“VRRP Advertisement Interval and Sub-second Failover” on page 430](#).

To enable or disable VRRPv2 backward compatibility, use the following command (for this example, ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# vrrp <1-255>
G8272(config-if-vrrp)# [no] v2-compatible
```

A backup virtual router having VRRPv2 backward compatibility enabled will ignore VRRPv2 advertisements from the master virtual router if it also receives VRRPv3 advertisements.

If a backup virtual router has VRRPv2 backward compatibility enabled and it receives only VRRPv2 advertisements from the master virtual router, it will transform the received advertisement interval into centiseconds.

If the master virtual router has VRRPv2 backward compatibility enabled and the advertisement interval set to a sub-second value, it will ignore this setting and send advertisements to the backup virtual routers every one second.

VRRP Accept Mode

The G8272 supports VRRP Accept mode. It makes the master virtual router to accept packets with IP addresses not own by it. Thus, such packets are accepted by the master virtual router if it is the owner of the IP address or accept mode is enabled. The master virtual router will discard any traffic if it isn't the owner of the IP address and accept mode is disabled.

If Accept mode is disabled, the master virtual router accepts packets sent to the virtual router only if it is the IP address owner. Any packets sent to the master virtual router with an IP address that it doesn't own will be discarded.

This restriction causes problems when troubleshooting network connectivity issues. When a host is unreachable, it is common practice to ping the first-hop router to determine if the problem is related to the default gateway of the host. If the host's default gateway is a virtual router that doesn't respond to pings (or ICMP Echo Requests), this method of troubleshooting is unavailable for use.

Accept mode configures the master virtual router to respond only to ping, traceroute and telnet packets with an IP address not owned by the router. The master virtual router will not respond to any other type of packets. When it responds, it uses the virtual router's IP and MAC addresses.

By default, VRRP accept mode is enabled.

To enable or disable VRRP accept mode, use the following command:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# vrrp <1-255>
G8272(config-if-vrrp)# [no] accept-mode
```

VRRP Preemption

When VRRP preemption is enabled on a backup virtual router, if the router has a higher priority than the current master virtual router, it will preempt the lower priority master and assume control.

Note: Even when VRRP preemption is disabled, a backup virtual router will always preempt any other master if the router is the owner of the IP address associated with the virtual router.

By default, VRRP preemption is enabled.

To enable or disable VRRP preemption, use the following command:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# vrrp <1-255>
G8272(config-if-vrrp)# [no] preempt
```

VRRP Priority

Priority is used by VRRP in the master virtual router election process. Each VRRP router participating in the election process is assigned a priority level between 0 and 255 and the VRRP router with the highest priority becomes the master virtual process.

You can configure a VRRP router with a priority between 1 and 254. The value of 255 is reserved for the VRRP router that owns the IP address associated with the virtual router. The value of 0 is reserved for the master virtual router to indicate that it has stopped participating in the VRRP. This triggers the backup virtual routers to quickly start a new master election process without waiting for the current master virtual router to timeout.

To configure the priority of a VRRP router, use the following command:

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# vrrp <1-255>
G8272(config-if-vrrp)# priority <1-254>
```

By default, the priority of a VRRP router is 100. To reset the priority to its default value, use the following command:

```
G8272(config-if-vrrp)# no priority
```

Note: If you configure the virtual IP address of the virtual router to have the same IP address as the interface you set up a VRRP router, the previous priority configuration of the VRRP router will be deleted and its new priority will be 255.

Note: If a VRRP router is the owner of the virtual IP address of the virtual router, having a priority of 255, and you configure the virtual IP address to an IP address different than the one owned by the VRRP router, the priority of that VRRP router will be reset to the default.

IPv6 VRRP

VRRPv3 supports both IPv4 and IPv6 addressing. The type of IP addressing used by the VRRP router must be the same as used by the interface on which the VRRP router is configured. For example, if an ethernet interface is configured as an IPv4 address, the VRRP router created on that interface must also be configured with an virtual IPv4 address.

To create an IPv6 VRRP session, use the following command (for this example, ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12
G8272(config-if)# vrrp <1-255> ipv6
G8272(config-if-vrrp)#
```

Note: If there is a mismatch between the type of the virtual IP address and the type of the IP address used by the interface or by the VRRP session, you will be prompted with an error message.

When using IPv6 addressing, each VRRP session must be configured with a set of two IPv6 addresses: a link-local address as the primary virtual address and a global IPv6 address.

Note: The link-local IPv6 address must be different than the global IPv6 address.

The link-local address must be the first address in the IPv6 address list used in the VRRP packet. If this address is the same as the virtual IP address associated with the virtual router, then the VRRP router is the owner of the virtual IP address and its priority is set to 255. The VRRP router is advertised with priority 255 in VRRP packets regardless of the global IPv6 address.

Note: The same link-local IPv6 address cannot be used in other VRRP sessions on the same interface.

Even if an interface is configured with the same IPv6 address as the virtual global IPv6 address of a VRRP router, the VRRP router is not considered the owner of the IPv6 address. The link-local address takes precedence over the global address.

Note: Only one global IPv6 address and one link-local address are supported for each VRRP IPv6 session.

To configure a global IPv6 address, use the following command:

```
G8272(config)# interface ethernet 1/12  
G8272(config-if)# vrrp <1-255> ipv6  
G8272(config-if-vrrp)# address <IPv6 address>
```

To delete the global IPv6 address of a VRRP session, use the following command:

```
G8272(config-if-vrrp)# no address <IPv6 address>
```

To configure the link-local IPv6 address of the VRRP session, use the following command:

```
G8272(config-if-vrrp)# address <IPv6 address> primary
```

To delete the link-local IPv6 address of a VRRP session, use the following command:

```
G8272(config-if-vrrp)# no address <IPv6 address> primary
```

To delete all configured IP addresses of a VRRP session, use the following command:

```
G8272(config-if-vrrp)# no address
```

Virtual Router Deployment Considerations

Configuring the Switch for Tracking

Tracking configuration largely depends on user preferences and network environment. Consider the configuration shown in [Figure 27 on page 429](#). Assume the following behavior on the network:

- Upon initialization, switch 1 is the Master Virtual Router and switch 2 is the Backup Virtual Router.
- Switch 1 and 2 are both connected to each of the four servers on the local network.
- If switch 1 is the master and one of its interfaces goes down, meaning switch 1 is still connected to three out of the four servers, then switch 1 remains the Master Virtual Router.

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 two or more of its interfaces go down, meaning switch 1 is not connected to at least three of the four servers, then switch 2 becomes the new Master Virtual Router and switch 1 becomes the Backup Virtual Router.
- If switch 2 is the master, it will retain its role as the Master Virtual Router, even when the links between switch 1 and all the servers are restored, as long as switch 2 has at least the same number of actively connected servers as switch 1.
- If switch 2 is the master and the number of actively connected servers is lower than the ones connected to switch 1, it loses its role as the Master Virtual Router. Switch 1 regains its role as the Master Virtual Router and switch 2 becomes the Backup Virtual Router.

You can implement this behavior by configuring the switch for tracking as follows:

1. Set the priority for switch 1 to 101. For more details, see [“VRRP Priority” on page 433](#).
2. Leave the priority for switch 2 at the default value of 100.
3. On both switches, enable tracking based on ethernet interfaces or VLANs. For more details, see [“Interface Tracking” on page 431](#).

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.

Basic VRRP Configuration

The following steps are an example of a basic VRRP configuration:

Configuring Switch 1:

1. Enter Global Configuration command mode:

```
G8272> enable
G8272# configure device
G8272(config)#
```

2. Configure the ethernet port to be part of the virtual router:

```
G8272(config)# interface ethernet 1/1
G8272(config-if)# no bridge-port
```

3. Assign an IP address to the interface:

- a. either an IPv4 address:

```
G8272(config-if)# ip address 192.168.1.10/24
```

- b. or an IPv6 address:

```
G8272(config-if)# ipv6 address 2001:db8:a0b:12f0::10/32
```

4. Create a virtual router with an assigned ID and configure its virtual IP address:

- a. either using IPv4 addressing:

```
G8272(config-if)# vrrp 1
G8272(config-if-vrrp)# address 192.168.1.1
```

- b. or IPv6 addressing:

```
G8272(config-if)# vrrp 1 ipv6
G8272(config-if-vrrp)# address fe80::2001 primary
G8272(config-if-vrrp)# address 2001:db8:a0b:12f0::76e
```

Note: The virtual IP address of the VRRP router must be same type as the IP address of the interface on which VRRP is configured.

5. Set the priority of the VRRP router and enable VRRP:

```
G8272(config-if-vrrp)# priority 110
G8272(config-if-vrrp)# no shutdown
```

Note: Switch 1 has a higher priority and will become the master virtual router.

6. Check the VRRP configuration:

```
G8272# display vrrp detail
```

Configuring Switch 2:

1. Enter Global Configuration command mode:

```
G8272> enable
G8272# configure device
G8272(config)#
```

2. Configure the ethernet port to be part of the virtual router:

```
G8272(config)# interface ethernet 1/1
G8272(config-if)# no bridge-port
```

3. Assign an IP address to the interface:

- a. either an IPv4 address:

```
G8272(config-if)# ip address 192.168.1.20/24
```

- b. or an IPv6 address:

```
G8272(config-if)# ipv6 address 2001:db8:a0b:12f0::20/32
```

4. Create a virtual router with an assigned ID and configure its virtual IP address:

- a. either using IPv4 addressing:

```
G8272(config-if)# vrrp 1
G8272(config-if-vrrp)# address 192.168.1.1
```

- b. or IPv6 addressing:

```
G8272(config-if)# vrrp 1 ipv6
G8272(config-if-vrrp)# address fe80::2001 primary
G8272(config-if-vrrp)# address 2001:db8:a0b:12f0::76e
```

Note: The virtual IP address of the VRRP router must be same type as the IP address of the interface on which VRRP is configured.

5. Enable VRRP:

```
G8272(config-if-vrrp)# no shutdown
```

6. Check the VRRP configuration:

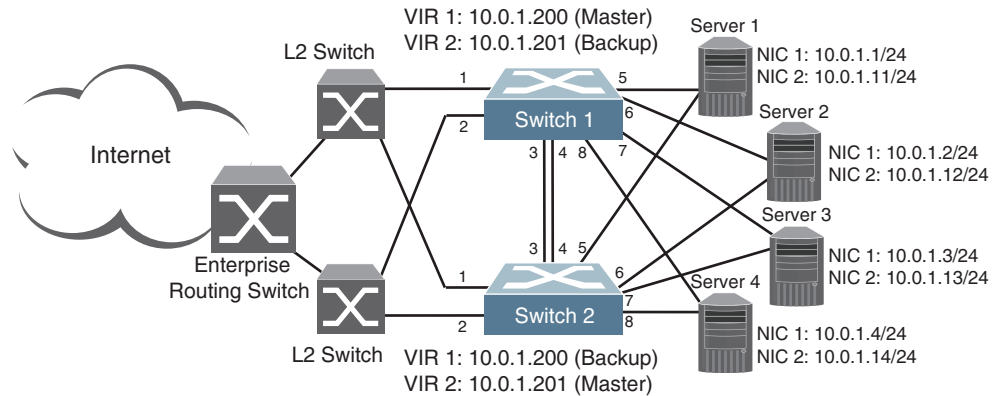
```
G8272# display vrrp detail
```

High Availability Configuration

VRRP High-Availability Using Multiple VIRs

Figure 28 shows an example configuration where two G8272s are used as VRRP routers in an active-active configuration. In this configuration, both switches respond to packets.

Figure 28. 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.

In the scenario illustrated in Figure 28, traffic destined for IPv4 address 10.0.1.1 is forwarded through the Layer 2 switch at the top of the drawing and ingresses Switch 1 on port 1. Return traffic uses default gateway 1 (192.168.1.1).

If the link between Switch 1 and a Layer 2 switch fails, Switch 2 becomes the Master because it has a higher priority.

Switches 1 and 2 have Servers 1, 2, 3 and 4 connected to them through ports 5, 6, 7 and 8, respectively. These ports are all members of the same VLAN.

To implement the active-active example, perform the following switch configuration.

Task 1: Configure Switch 1

1. Configure routing ports.

```
G8272(config)# interface ethernet 1/1
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 10.0.1.20/24
G8272(config-if)# exit

G8272(config)# interface ethernet 1/2
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 10.0.1.21/24
G8272(config-if)# exit

G8272(config)# interface ethernet 1/3
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 10.0.1.100/24
G8272(config-if)# exit

G8272(config)# interface ethernet 1/4
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 10.0.1.101/24
G8272(config-if)# exit
```

2. Configure the default gateway, which points to a Layer 3 router.

```
G8272(config)# ip route 0.0.0.0/0 ethernet 1/1 10.0.1.20
G8272(config)# ip route 0.0.0.0/0 ethernet 1/2 10.0.1.21
```

3. Configure two Virtual Interface Routers.

```
G8272(config)# interface ethernet 1/1
G8272(config-if)# vrrp 1
G8272(config-if-vrrp)# address 192.168.1.200
G8272(config-if-vrrp)# exit
G8272(config-if)# exit

G8272(config)# interface ethernet 1/2
G8272(config-if)# vrrp 2
G8272(config-if-vrrp)# address 192.168.1.201
G8272(config-if-vrrp)# exit
G8272(config-if)# exit
```


4. Enable tracking on ethernet ports 1/1 and 1/2. Set the priority of Virtual Router 1 to 101, so that it becomes the Master.

```
G8272(config)# interface ethernet 1/1
G8272(config-if)# vrrp 1
G8272(config-if-vrrp)# priority 101
G8272(config-if-vrrp)# track interface ethernet 1/2 priority 3
G8272(config-if-vrrp)# no shutdown
G8272(config-if-vrrp)# exit
G8272(config-if)# exit

G8272(config)# interface ethernet 1/2
G8272(config-if)# vrrp 2
G8272(config-if-vrrp)# track interface ethernet 1/1 priority 3
G8272(config-if-vrrp)# no shutdown
G8272(config-if-vrrp)# exit
G8272(config-if)# exit
```

5. Create VLAN 10 and configure server interfaces (ports 1/5 - 1/8).

```
G8272(config)# vlan 10
G8272(config-vlan)# exit

G8272(config)# interface ethernet 1/5-8
G8272(config-if-range)# bridge-port mode trunk
G8272(config-if-range)# bridge-port trunk allowed vlan 10
G8272(config-if-range)# exit
```

Task 2: Configure Switch 2

1. Configure routing ports.

```
G8272(config)# interface ethernet 1/1
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 10.0.1.22/24
G8272(config-if)# exit

G8272(config)# interface ethernet 1/2
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 10.0.1.23/24
G8272(config-if)# exit

G8272(config)# interface ethernet 1/3
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 10.0.1.102/24
G8272(config-if)# exit

G8272(config)# interface ethernet 1/4
G8272(config-if)# no bridge-port
G8272(config-if)# ip address 10.0.1.103/24
G8272(config-if)# exit
```

2. Configure the default gateways. Each default gateway points to a Layer 3 router.

```
G8272(config)# ip route 0.0.0.0/0 ethernet 1/1 10.0.1.22
G8272(config)# ip route 0.0.0.0/0 ethernet 1/2 10.0.1.23
```

3. Configure two Virtual Interface Routers.

```
G8272(config)# interface ethernet 1/1
G8272(config-if)# vrrp 1
G8272(config-if-vrrp)# address 192.168.1.200
G8272(config-if-vrrp)# exit
G8272(config-if)# exit

G8272(config)# interface ethernet 1/2
G8272(config-if)# vrrp 2
G8272(config-if-vrrp)# address 192.168.1.201
G8272(config-if-vrrp)# exit
G8272(config-if)# exit
```

4. Enable tracking on ethernet ports 1/1 and 1/2. Set the priority of Virtual Router 2 to 101, so that it becomes the Master.

```
G8272(config)# interface ethernet 1/1
G8272(config-if)# vrrp 1
G8272(config-if-vrrp)# track interface ethernet 1/2 priority 3
G8272(config-if-vrrp)# no shutdown
G8272(config-if-vrrp)# exit
G8272(config-if)# exit

G8272(config)# interface ethernet 1/2
G8272(config-if)# vrrp 2
G8272(config-if-vrrp)# priority 101
G8272(config-if-vrrp)# track interface ethernet 1/1 priority 3
G8272(config-if-vrrp)# no shutdown
G8272(config-if-vrrp)# exit
G8272(config-if)# exit
```

5. Create VLAN 10 and configure server interfaces (ports 1/5 - 1/8).

```
G8272(config)# vlan 10
G8272(config-vlan)# exit

G8272(config)# interface ethernet 1/5-8
G8272(config-if-range)# bridge-port mode trunk
G8272(config-if-range)# bridge-port trunk allowed vlan 10
G8272(config-if-range)# exit
```

Part 6: Network Management

This section discusses the following topics:

- [“Link Layer Discovery Protocol” on page 445](#)
- [“Simple Network Management Protocol” on page 459](#)

Chapter 24. Link Layer Discovery Protocol

Cloud Network OS supports Link Layer Discovery Protocol (LLDP). This chapter discusses the use and configuration of LLDP on the switch:

- [“LLDP Overview” on page 446](#)
- [“Enabling or Disabling LLDP” on page 447](#)
- [“LLDP Transmit Features” on page 448](#)
- [“LLDP Receive Features” on page 452](#)
- [“LLDP Example Configuration” on page 456](#)

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 G8272 can advertise the presence of its ports, their major capabilities, and their current status to adjacent LLDP stations. 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.

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 G8272 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

Follow the directions in this section to enable and disable LLDP.

Transmit and Receive Control

The G8272 can be configured to transmit or receive LLDP information on a port-by-port basis.

LLDP is enabled by default on the switch. When LLDP is enabled on the switch, G8272 ports transmit and receive LLDP information.

To enable or disable the transmission of LLDP information, use the following command (for this example ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12  
G8272(config-if)# [no] lldp transmit
```

To enable or disable the reception of LLDP information, use the following command (for this example ethernet interface 1/12 is used):

```
G8272(config)# interface ethernet 1/12  
G8272(config-if)# [no] lldp receive
```

To view the LLDP transmit and receive status, use the following commands:

```
G8272# display lldp interface all
```

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 G8272 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 transmit interval represents the number of seconds between two consecutive LLDP transmissions. To configure the global transmit interval (in seconds), use the following command:

```
G8272(config)# lldp timer <5-32768>
```

The default is 30 seconds. To reset the transmit interval to its default value, use the following command:

```
G8272(config)# no lldp timer
```

Minimum Interval

In addition to sending LLDP information at scheduled intervals, LLDP information is also sent when the G8272 detects relevant changes to its configuration or status (such as when ports are enabled or disabled). To prevent the G8272 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.

To configure the minimum transmit interval (in seconds), use the following command:

```
G8272(config)# lldp transmit-delay <1-8192>
```

where the range is from 1 second to one-quarter of the scheduled transmit interval (**lldp timer** <value>), up to a maximum of 8192 seconds.

The default is 2 seconds. To reset the minimum transmit interval to its default value, use the following command:

```
G8272(config)# no lldp transmit-delay
```


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.

```
G8272(config)# lldp holdtime-multiplier <2-10>
```

The default value is 4, meaning that remote systems will hold the port's LLDP information for four times the 30 second transmit interval/scheduled interval (`msgtxint`) value, or 120 seconds, before removing it from their MIB.

Note: If the transmission interval has a different value from the default value, the resulting timeout value will be different.

To reset the hold time multiplier to its default value, use the following command:

```
G8272(config)# no lldp holdtime-multiplier
```

Trap Notifications

If SNMP is enabled on the G8272 (see [“Using Simple Network Management Protocol” on page 34](#)), each port can be configured to send SNMP trap notifications whenever an update occurs in the LLDP database. By default, trap notification is disabled for each port. The trap notification state can be changed using the following commands:

```
G8272(config)# interface ethernet 1/1
G8272(config-if)# [no] lldp trap-notification
G8272(config-if)# exit
```

LLDP information is also sent when neighbors to the G8272 are added, removed, or changed. To prevent the G8272 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.

To configure the minimum trap notification interval (in seconds), use the following command:

```
G8272(config)# lldp trap-interval <5-3600>
```

The default is 5 seconds. To reset the minimum trap notification interval to its default value, use the following command:

```
G8272(config)# no lldp trap-interval
```

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:

```
G8272(config)# [no] logging level lldp <logging level>
```

where *logging level* is one of the following:

Value	Meaning
0	emergencies only
1	alerts
2	critical
3	errors
4	warnings
5	notifications
6	information
7	debugging

Changing the LLDP Transmit State

When the port is disabled, or when LLDP transmit is turned off for the port (see [“Transmit and Receive Control” on page 447](#)), 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 G8272 port from their MIB.

In addition, if LLDP is fully disabled on a port and then later re-enabled, the G8272 will temporarily delay resuming LLDP transmissions on the port to allow the port LLDP information to stabilize.

To globally configure the reinitialization delay interval (in seconds), use the following command:

```
G8272(config)# lldp reinit <1-10>
```

The default is 2 seconds. To reset the reinitialization delay interval to its default value, use the following command:

```
G8272(config)# no lldp reinit
```

Types of Information Transmitted

When LLDP transmission is permitted on the port (see [“Enabling or Disabling LLDP” on page 447](#)), 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 Ethernet Mode):

```
G8272(config)# interface ethernet 1/1
G8272(config-if)# [no] lldp tlv-select <type>
G8272(config-if)# exit
```

where *type* is one of the following values:

Value	Definition
link-aggregation	Link Aggregation TLV
mac-phy-status	MAC/PHY Configuration/Status TLV
management-address	Management Address TLV
max-frame-size	Maximum frame size TLV
port-description	Port description TLV
port-protocol-vlan	Port and Protocol VLAN ID TLV
port-vlan	Port VLAN ID TLV
power-mdi	Power via MDI TLV
protocol-identity	Protocol identity TLV
system-capabilities	System capabilities TLV
system-description	System description TLV
system-name	System name TLV
vid-management	Version ID Management TLV
vlan-name	VLAN Name TLV

LLDP Receive Features

The following LLDP receive features are supported by CNOS.

Types of Information Received

When the LLDP receive option is enabled on a port (see [“Enabling or Disabling LLDP” on page 447](#)), the port may receive the following TLV information from LLDP-capable remote systems:

- Chassis ID
- Port ID
- LLDP Time-to-Live
- Link Aggregation
- MAC/PHY Configuration/Status
- Management Address
- Maximum Frame Size
- Port Description
- Port and Protocol VLAN ID
- Port VLAN ID
- Power via MDI
- Protocol Identity
- System Capabilities
- System Description
- System Name
- VID Management
- VLAN Name

The G8272 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 G8272 MIB
- Using commands on the G8272

The following command displays LLDP neighbor information:

```
G8272# display lldp neighbors
```

For more information on LLDP commands, see the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

Note: Received LLDP information can change very quickly. When using display 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:

```
G8272# display lldp neighbors detail

Chassis ID MAC Address: a897.dcde.530d
Port ID Locally Assigned: Ethernet1/11
Local Port ID: Ethernet1/11
Port Description: Ethernet1/11
System Name: LENOVO G8272
System Description: LENOVO RackSwitch G8272, LENOVO NOS version 10.1.1.0
Time remaining: 106 seconds
System Capabilities: B, R
Enabled Capabilities: B, R
Management Address: 192.168.50.50
VLAN ID: 10

Chassis ID MAC Address: a897.dcde.530f
Port ID Locally Assigned: Ethernet1/13
Local Port ID: Ethernet1/13
Port Description: Ethernet1/13
System Name: LENOVO G8272
System Description: LENOVO RackSwitch G8272, LENOVO NOS version 10.1.1.0
Time remaining: 106 seconds
System Capabilities: B, R
Enabled Capabilities: B, R
Management Address: 192.168.50.50
VLAN ID: 10

Chassis ID MAC Address: a897.dcde.5311
Port ID Locally Assigned: Ethernet1/15
Local Port ID: Ethernet1/15
Port Description: Ethernet1/15
System Name: LENOVO G8272
System Description: LENOVO RackSwitch G8272, LENOVO NOS version 10.1.1.0
Time remaining: 106 seconds
System Capabilities: B, R
Enabled Capabilities: B, R
Management Address: 192.168.50.50
VLAN ID: 10

Total entries displayed: 3
```

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 timer 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.

Debugging LLDP

LLDP debugging is disabled by default. To enable or disable LLDP debugging, enter:

```
G8272(config)# [no] dbg lldp <type>
```

where *type* is one of the following:

Parameter	Description
all	Enable all debugging flags.
decode	Enable decode debugging.
encode	Enable encode debugging.
error	Enable error debugging.
event	Enable event debugging.
interface all interface ethernet <chassis number/port number>	Enable debugging on all interfaces or on the specified interface.
message	Enable Network Security Monitoring (NSM) message debugging.
rx	Enable reception debugging.
trace	Enable trace debugging.
tx	Enable transmission debugging.

LLDP Debugging Types

The following types of debugging are available:

- all
 - Enables all debug flags.
- event
 - Messages appear when the following events occur:
 - Timer expiration
 - Admin status changes
 - TLV selection changes
 - Timer configuration changes
 - Transmit initialization
 - Receive initialization

- **interface**
Debugging occurs on all physical interfaces or on the specified interface.
- **message**
Prints messages between Network Security Monitoring (NSM) and the Linux LLDP process.
- **rx**
Prints the packet contents when reception occurs.
- **trace**
Prints the process stage of LLDP packets. For example:
 - Encode/Decode management address
 - TLV/TX packet enters delay timer
 - Begin transmission
 - Begin receptionWhen an error is received on the LLDP PDU, the packet content is logged to the internal debug buffer.
- **tx**
Prints the packet contents when transmission occurs.
- **Error**
Prints a message when an error occurs.

To view LLDP debugging information, enter:

```
G8272# display dbg lldp
```

LLDP Example Configuration

The following is an example LLDP configuration.

1. Set the global LLDP timer features.

- a. Configure the transmit interval:

```
G8272(config)# lldp timer 8
```

- b. Configure the minimum transmit interval:

```
G8272(config)# lldp transmit-delay 2
```

- c. Configure the hold time multiplier:

```
G8272(config)# lldp holdtime-multiplier 4
```

- d. Configure the reinitialization interval:

```
G8272(config)# lldp reinit 2
```

- e. Configure the minimum trap notification interval:

```
G8272(config)# lldp trap-notification-interval 5
```

2. Set LLDP options for each port.

- a. Enter Interface Configuration mode for each ethernet port:

```
G8272(config)# interface ethernet 1/12
```

- b. Enable LLDP transmission and reception:

```
G8272(config-if)# lldp receive  
G8272(config-if)# lldp transmit
```

- c. Enable SNMP trap notifications:

```
G8272(config-if)# lldp trap-notification
```

- d. Configure the port to include port description in its outgoing LLDP messages:

```
G8272(config-if)# lldp tlv-select port-description  
G8272(config-if)# exit
```

3. Enable syslog reporting.

```
G8272(config)# logging level lldp 6
```


4. View remote device information.

```
G8272# display lldp neighbors

Capability codes:
  (R) Router, (B) Bridge, (T) Telephone, (C) DOCSIS Cable Device
  (W) WLAN Access Point, (P) Repeater, (S) Station, (O) Other
Device ID           Local Intf      Hold-time  Capability  Port ID
MA_49               Ethernet1/14   120       BR          Ethernet1/14
G8052-71            mgmt0          120       BR          3

Total entries displayed: 2
```

```
G8272# display lldp timers

LLDP Timers:

Holdtime in seconds: 120
Reinit-time in seconds: 2
Transmit interval in seconds: 30
Transmit delay in seconds: 2
Trap interval in seconds: 5
```

```
G8272# display lldp interface ethernet 1/12

Interface Name: Ethernet1/12
-----
Interface Information
Enable (tx/rx/trap): Y/Y/N   Port Mac address: a8:97:dc:de:02:0e
```

Chapter 25. Simple Network Management Protocol

Cloud Network 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.

The following topics are discussed in this section:

- [“SNMP Version 1 & Version 2” on page 460](#)
- [“SNMP Version 3” on page 461](#)
- [“Default Configuration” on page 462](#)
- [“Configuration Examples” on page 463](#)
- [“SNMP MIBs” on page 465](#)

SNMP Version 1 & Version 2

To access the SNMP agent on the G8272, 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:

```
G8272(config)# snmp-server community <1-32 characters> [ro|rw|group <word>|view <view name>] version {v1|v2c} {ro|rw}]
```

where `ro` provides read-only access and `rw` applies read-write access with the specified community string.

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

SNMP Notifications

Lenovo N/OS generates SNMP notifications as either traps or informs.

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:

```
G8272(config)# snmp-server enable traps [bfd|link [linkDown|linkUp]]
G8272(config)# snmp-server host <IPv4 or IPv6 address> traps version {1|2c|3}
<user name> [udp-port <165535>]
```

Note: The SNMP agent will not send traps on the management interface.

In the case of informs, the SNMP manager that receives an inform message will acknowledge the message with an SNMP response PDU. If the sender does not receive a response for an inform message, the inform can be sent again. To configure informs, use the following command:

```
G8272(config)# snmp-server host <IPv4 or IPv6 address> informs version {1|2c|3}
<user name> [udp-port <165535>]
```

SNMP Device Contact and Location

You can assign the device contact information and device location by using the following commands:

```
G8272(config)# snmp-server contact <contact name>
G8272(config)# snmp-server location <location name>
```

One-Time Authentication for SNMP over TCP

You can enable a one-time authentication for SNMP over a TCP session by using the following command:

```
G8272(config)# snmp-server tcp-session
```

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:

```
G8272(config)# snmp-server host <SNMP notification target host IPv4 or IPv6 address>
version 3 ?
```

For more information on SNMP MIBs and the commands used to configure SNMP on the switch, see the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

Default Configuration

Up to 16 SNMP users can be configured on the switch. To modify an SNMP user, enter the following commands:

```
G8272(config)# snmp-server user <1-32 characters>
```

Users can be configured to use the authentication/privacy options. The G8272 supports:

- two authentication algorithms:

- o MD5

```
G8272(config)# snmp-server user <username> auth md5 <password>
```

- o SHA

```
G8272(config)# snmp-server user <username> auth sha <password>
```

- two privacy protocols:

- o des (CBC-DES Symmetric Encryption Protocol)

```
G8272(config)# snmp-server user <username> auth {md5|sha} <password> priv des <privilege password>
```

- o aes (AES-128 Advanced Encryption Standard)

```
G8272(config)# snmp-server user <username> auth {md5|sha} <password> priv aes <privilege password>
```

Configuration Examples

Following are examples of the SNMP operations supported by Lenovo N/OS.

Basic SNMP Configuration Example

1. Enable SNMP feature.

```
G8272(config)# snmp-server enable snmp
```

2. Set a SNMP view with the name 'all', OID Tree 'OID-TREE' and type 'included'.

```
G8272(config)# snmp-server view all OID-TREE included
```

3. Configure the community parameters.

```
G8272(config)# snmp-server community public group network-operator  
G8272(config)# snmp-server community private group network-admin
```

4. View SNMP configuration.

```
G8272(config)# display snmp
```

User Configuration Example

1. Enable SNMP feature.

```
G8272(config)# snmp-server enable snmp
```

2. Set a SNMP view with the name 'all', OID Tree 'OID-TREE' and type 'included'.

```
G8272(config)# snmp-server view all OID-TREE included
```

3. Configure a SNMP user with authentication password 'password1' and privilege password 'password2'.

```
G8272(config)# snmp-server user summer network-admin auth sha password1  
priv des password2
```

4. View SNMP user configuration.

```
G8272(config)# display snmp user
```

Configuring SNMP Trap Hosts

1. Enable SNMP traps.

```
G8272(config)# snmp-server enable traps
```

2. Send traps messages to a SNMP recipient:

```
G8272(config)# snmp-server host 1.0.0.2 traps version 3 authpriv test  
udp-port 162
```

3. View SNMP trap configuration.

```
G8272(config)# display snmp trap
```

SNMP MIBs

The CNOS 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 CNOS SNMP agent are contained in the CNOS enterprise MIB document.

The CNOS SNMP agent supports the following standard MIBs:

- rfc1213.mib
- rfc1573.mib
- rfc1643.mib
- rfc2011.mib
- rfc2012.mib
- rfc2013.mib
- rfc2037.mib
- rfc2233.mib
- rfc2790.mib
- rfc2863.mib
- rfc3411.mib
- rfc3412.mib
- rfc3413.mib
- rfc3414.mib
- rfc3415.mib
- rfc3418.mib
- rfc4363.mib
- lag-mib.mib
- lldp-mib.mib
- rfc4022.mib
- rfc4113.mib
- rfc4133.mib
- rfc4188.mib
- rfc4363.mib
- rfc4273.mib
- rfc4293.mib
- rfc4750.mib

The CNOS SNMP agent supports the following generic traps as defined in:

- RFC 1215:
 - linkDown
 - linkUp
 - egpNeighborLoss
- RFC 4188:
 - newRoot
 - topologyChange
- RFC 4273:
 - bgpEstablishedNotification
 - bgpBackwardTransNotification
- RFC 4750:
 - ospfVirtIfStateChange
 - ospfNbrStateChange
 - ospfVirtNbrStateChange
 - ospfIfConfigError
 - ospfVirtIfConfigError
 - ospfIfAuthFailure
 - ospfVirtIfAuthFailure
 - ospfIfRxBadPacket
 - ospfVirtIfRxBadPacket
 - ospfTxRetransmit
 - ospfVirtIfTxRetransmit
 - ospfOriginateLsa
 - ospfMaxAgeLsa
 - ospfLsdbOverflow
 - ospfLsdbApproachingOverflow
 - ospfIfStateChange
 - ospfNssaTranslatorStatusChange
 - ospfRestartStatusChange
 - ospfNbrRestartHelperStatusChange
 - ospfVirtNbrRestartHelperStatusChange

Part 7: Monitoring

The ability to monitor traffic passing through the G8272 can be invaluable for troubleshooting some types of networking problems. This section covers the following monitoring features:

- [“Port Mirroring” on page 469](#)

Chapter 26. Port Mirroring

The Cloud Network 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.

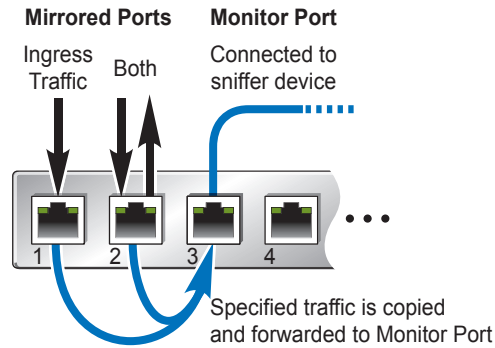
The following topics are discussed in this section:

- [“Port Mirroring Overview” on page 470](#)
- [“SPAN Configuration” on page 471](#)
- [“ERSPAN Configuration” on page 473](#)
- [“Limitations” on page 477](#)

Port Mirroring Overview

The G8272 supports a “many to one” mirroring model. As shown in [Figure 29](#), 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 29. Mirroring Ports



The G8272 supports four monitor sessions for mirroring rx or tx traffic and two monitor sessions for mirroring both rx and tx traffic. Each monitor port can receive mirrored traffic from any number of target ports.

Cloud NOS 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 before processing.

SPAN Configuration

The Switched Port Analyzer (SPAN) feature selects network traffic for analysis by a network analyzer. The network analyzer can be any Remote Monitoring mechanism with the capability to perform packet captures.

SPAN feature is applicable for the following types of ports:

- Ethernet ports (but not sub-interfaces);
- Link Aggregation Groups (LAGs).

Sources

SPAN sources refer to the interfaces from which traffic can be monitored. The Cloud NOS supports Ethernet and LAGs as SPAN sources. The SPAN traffic can be performed in the ingress direction, the egress direction or both directions:

- Ingress source (Rx)—Traffic entering the device through this source port is copied to the SPAN destination port.
- Egress source (Tx)—Traffic exiting the device through this source port is copied to the SPAN destination port.

The command to add or remove a source interface for the current SPAN session is as follows:

```
G8272(config)# monitor session {<1-18>|all}
G8272(config-monitor)# [no] source interface {ethernet <chassis number/port
number>|port-aggregation <LAG number>} {both|rx|tx}
```

Destinations

SPAN destinations refer to the interfaces that monitors source ports.

The command to add or remove a destination interface for the current SPAN session is as follows:

```
G8272(config)# monitor session {<1-18>|all}
G8272(config-monitor)# [no] destination interface {ethernet <chassis
number/port number>|port-aggregation <LAG number>}
```

Sessions

Lenovo N/OS supports SPAN sessions to designate sources and destinations to monitor.

The command to add or remove a SPAN session is as follows:

```
G8272(config)# [no] monitor session {<1-8>|all}
```

Configuration Example

Follow this procedure to configure SPAN feature:

1. Enter the monitor configuration mode.

```
G8272(config)# monitor session 3  
G8272(config-monitor)
```

2. Configure sources and the traffic direction in which to copy packets. Source interface is Ethernet 1/1 and both incoming and outgoing traffic is mirrored.

```
G8272(config-monitor)# source interface ethernet 1/1 both
```

3. Configure a destination for copied source packets.

```
G8272(config-monitor)# destination interface ethernet 1/3
```

4. Optionally, configure a description for the session. By default, no description is defined.

```
G8272(config-monitor)# description span_session_3
```

5. Enable the SPAN session. By default, the session is created in the down state.

```
G8272(config-monitor)# no shutdown
```

6. Display the SPAN configuration.

```
G8272# display monitor session 3
```

To clear the configuration of the specified SPAN session, use the following command:

```
G8272(config)# no monitor session 3
```

ERSPAN Configuration

Lenovo N/OS supports the Encapsulated Remote Switching Port Analyser (ERSPAN) feature on both source and destination ports. ERSPAN transports mirrored traffic over an IP network. The traffic is encapsulated using generic routing encapsulation (GRE) at the source router and is transferred across the network. The packet is decapsulated at the destination router and then sent to the destination interface.

SPAN feature is applicable for the following types of ports:

- Ethernet ports (but not sub-interfaces);
- LAGs.

Session Types

Lenovo N/OS supports two types of ERSPAN session: source sessions and destination sessions.

ERSPAN source sessions are created on a switch for which traffic needs to be monitored from a remote device across an IP network. The ERSPAN source session mirrors traffic from an interface, encapsulates that traffic and sends it across the network to the specified IPv4 address.

ERSPAN destination sessions are created on a switch to receive mirrored traffic from another switch (on which an ERSPAN source session was created) and then forward that traffic to a monitoring device for analysis. The ERSPAN destination session receives GRE encapsulated packets from an IPv4 address, decapsulates the packets and sends them to a monitoring device reachable through one of its interfaces.

ERSPAN source sessions can be configured to directly send encapsulated GRE packets to a monitoring device, if that device is reachable. If the target host isn't reachable from the source switch, an ERSPAN destination session must be created on the switch to which the monitoring device is directly connected.

Creating an ERSPAN destination session is also required when the mirrored traffic needs to be received by the monitoring device without GRE/ERSPAN headers. The ERSPAN destination switch will decapsulate the packets and then forward them to the monitoring device in the same format in which they were received on the ERSPAN source switch.

The command to select an ERSPAN session type is as follows:

```
G8272(config)# [no] monitor session <1-8> type {erspan-destination|  
erspan-source}
```

To configure the global origin IPv4 address of ethernet ERSPAN sessions, use the following command:

```
G8272(config)# monitor erspan origin ip-address <IPv4 address> global
```

Sources

For ERSPAN source sessions, sources refer to the interfaces from which traffic can be monitored. The Cloud NOS supports ethernet ports and LAGs as ERSPAN sources. The ERSPAN traffic can be performed in the ingress direction, the egress direction or both directions:

- Ingress source (Rx)—Traffic entering the device through this source port is copied to the SPAN destination port.
- Egress source (Tx)—Traffic exiting the device through this source port is copied to the SPAN destination port.

In the case of an ERSPAN source session, there is a source interface and a destination IPv4 address. The command to add or remove a source interface for the current ERSPAN source session is as follows:

```
G8272(config)# monitor session <1-8> type erspan-source
G8272(config-erspan-src)# [no] source interface {ethernet <chassis number/
port number>|port-aggregation <LAG number>} [both|rx|tx]
```

For an ERSPAN destination session, there is a source IPv4 address and a destination interface. The command to add or remove a source IPv4 address for the current ERSPAN destination session is as follows:

```
G8272(config)# monitor session <1-8> type erspan-destination
G8272(config-erspan-dst)# [no] source ip <IP address>
```

For a detailed list of the ERSPAN source configuring commands, see the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

Destinations

For ERSPAN source sessions, destinations refer to the IPv4 addresses of the hosts that monitor the mirrored traffic.

In the case of an ERSPAN source session, there is a source interface and a destination IPv4 address. The command to configure the destination IPv4 address for the current ERSPAN source session is as follows:

```
G8272(config)# monitor session <1-8> type erspan-source
G8272(config-erspan-dst)# destination ip <IP address>
```

For an ERSPAN destination session, there is a source IPv4 address and a destination interface. The command to configure the destination interface for the current ERSPAN destination session is as follows:

```
G8272(config)# monitor session <1-8> type erspan-destination
G8272(config-erspan-dst)# destination interface {ethernet <chassis number/
port number>|port-aggregation <LAG number>}
```

For a detailed list of the ERSPAN destination configuring commands, see the *Lenovo Network Command Reference for Lenovo Cloud Network Operating System 10.1*.

ERSPAN Source Session Configuration Example

Follow this procedure to configure an ERSPAN source session:

1. Configure the ERSPAN global origin IP address.

```
G8272(config)# monitor erSPAN origin ip-address 10.1.1.1 global
```

2. Clear the configuration of the specified ERSPAN session.

The new session configuration is added to the existing session configuration.

```
G8272(config)# no monitor session 3
```

3. Configure a monitor session of type ERSPAN source.

```
G8272(config)# monitor session 3 type erSPAN-source  
G8272(config-erSPAN-src)#
```

4. Optionally, configure a description for the session. By default, no description is defined.

```
G8272(config-erSPAN-src)# description erSPAN_source_session_3
```

5. Configure the source interface and the traffic direction in which to copy packets.

```
G8272(config-erSPAN-src)# source interface port-aggregation 2 both
```

6. Configure the destination IP address (to which the traffic is mirrored) in the ERSPAN session.

Only one destination IP address is supported per ERSPAN source session.

```
G8272(config-erSPAN-src)# destination ip 10.1.0.0
```

7. Enable the ERSPAN session. By default, the session is created in the shutdown state.

```
G8272(config-erSPAN-src)# no shutdown
```

8. Verify the ERSPAN configuration.

```
G8272# display monitor session 3
```

ERSPAN Destination Session Configuration Example

Follow this procedure to configure an ERSPAN destination session:

1. Clear the configuration of the specified ERSPAN session.

The new session configuration is added to the existing session configuration.

```
G8272(config)# no monitor session 4
```

2. Configure a monitor session of type ERSPAN source.

```
G8272(config)# monitor session 4 type erspan-destination  
G8272(config-erspan-dst)#
```

3. Optionally, configure a description for the session. By default, no description is defined.

```
G8272(config-erspan-dst)# description erspan_destination_session_4
```

4. Configure the source IPv4 address.

```
G8272(config-erspan-dst)# source ip 10.1.0.0
```

5. Configure the destination interface (to which the traffic is forwarded for monitoring) in the ERSPAN session.

```
G8272(config-erspan-dst)# destination interface ethernet 1/23
```

6. Enable the ERSPAN session. By default, the session is created in the shutdown state.

```
G8272(config-erspan-dst)# no shutdown
```

7. Verify the ERSPAN configuration.

```
G8272# display monitor session 4
```

Limitations

- Entering additional monitor session commands does not clear previously configured SPAN parameters. You must enter the **no monitor session** command to clear configured SPAN parameters.
- If you specify multiple ingress source ports, the ports can belong to different VLANs.
- When you specify sources and do not specify a traffic direction (ingress, egress, or both), both is used by default.
- For LAGs, only the first 8 members are supported, if that LAG is configured as a sniffer port.
- Up to two ERSPAN sessions are allowed to run simultaneously on the switch.

Part 8: Appendices

This section discusses the following topics:

- [“Getting help and technical assistance” on page 481](#)
- [“Notices” on page 483](#)

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.

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	<ul style="list-style-type: none"> The room air must be continuously filtered with 40% atmospheric dust spot efficiency (MERV 9) according to ASHRAE Standard 52.2¹. 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	<ul style="list-style-type: none"> Copper: Class G1 as per ANSI/ISA 71.04-1985³ 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 - Compliance to the Electromagnetic Compatibility Directive

This product is in conformity with the protection requirements of EU Council Directive 2004/108/EC (until April 19, 2016) and EU Council Directive 2014/30/EU (from April 20, 2016) 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 equipment according to European Standards harmonized in the Directives in compliance. 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



Warning: 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.

Germany Class A Statement

Deutschsprachiger EU Hinweis:

Hinweis für Geräte der Klasse A EU-Richtlinie zur Elektromagnetischen Verträglichkeit

Dieses Produkt entspricht den Schutzanforderungen der EU-Richtlinie 2014/30/EU (früher 2004/108/EC) zur Angleichung der Rechtsvorschriften über die elektromagnetische Verträglichkeit in den EU-Mitgliedsstaaten und hält die Grenzwerte der Klasse A der Norm gemäß Richtlinie.

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 Betriebsmitteln

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 2014/30/EU (früher 2004/108/EC) 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 EU Richtlinie 2014/30/EU (früher 2004/108/EC), 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, Meitnerstr. 9, 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。

Index

Symbols

[] 23

A

Access Control List (ACL) 244
Access Control Lists. *See* ACLs.
accessing the switch
 security 93, 105
 TACACS+ 106
ACLs 117, 244
active-active redundancy 429
administrator account 40
aggregating routes 352
 example 377
anycast address, IPv6 304
assistance, getting 481
Australia Class A statement 490
authenticating, in OSPF 400
autonomous systems (AS) 394

B

Border Gateway Protocol (BGP) 339
 failover configuration 375
 route aggregation 352
 route maps 349
Bridge Protocol Data Unit (BPDU) 200
broadcast domains 157

C

Canada Class A electronic emission statement 490
China Class A electronic emission statement 493
CIST 199
Class A electronic emission notice 490
command conventions 23
Command Line Interface 395
configuring
 BGP failover 375
 IP routing 279
 OSPF 404
contamination, particulate and gaseous 488

D

default gateway 278
 configuration example 280
default password 40
default route
 OSPF 398
Differentiated Services Code Point (DSCP) 246
downloading software 70
DSCP 246

E

- electronic emission Class A notice 490
- End user access control
 - configuring 99
- EtherChannel 183
- European Union EMC Directive conformance statement 490
- external routing 341, 394

F

- failover
 - overview 429
- FCC Class A notice 490
- FCC, Class A 490

G

- gaseous contamination 488
- gateway. *See* default gateway.
- Germany Class A statement 491
- getting help 481

H

- help
 - sources of 481
- help, getting 481
- high-availability 425
- HP-OpenView 459

I

- IBM DirectorSNMP
 - IBM Director 459
- IEEE standards
 - 802.1p 245
 - 802.1s 198
- IGMP 313
 - Querier 319
- IGMPv3 318
- image
 - downloading 70
- internal routing 341, 394
- Internet Group Management Protocol (IGMP) 313
- IP address
 - routing example 279
- IP interfaces
 - example configuration 279
- IP routing
 - cross-subnet example 277
 - default gateway configuration 280
 - IP interface configuration 279
 - IP subnets 277
 - subnet configuration example 278
 - switch-based topology 278
- IP subnets 278
 - routing 277, 278
 - VLANs 157

IPv6 addressing 301, 302
ISL aggregation 183

J

Japan Class A electronic emission statement 492
Japan Electronics and Information Technology Industries Association statement 492
JEITA statement 492

K

Korea Class A electronic emission statement 493

L

Link Layer Discovery Protocol 445
LLDP 445
logical segment. *See* IP subnets.
LSAs 393

M

manual style conventions 23
mirroring ports 469
monitoring ports 469
MSTP Multiple Spanning Tree Protocol (MSTP) 198
multi-links between switches
 using port aggregation 179
Multiple Spanning Tree Protocol 198

N

Neighbor Discovery, IPv6 306
network management 28, 34, 459
New Zealand Class A statement 490
notes, important 486
notices 483

O

OSPF
 area types 391
 authentication 400
 configuration examples 405
 default route 398
 link state database 393
 neighbors 393
 overview 390
 redistributing routes 353
 route maps 349, 351
 route summarization 397
 router ID 400
 virtual link 399

P

particulate contamination 488

- password
 - administrator account 40
 - default 40
- passwords 40
- People's Republic of China Class A electronic emission statement 493
- Per Hop Behavior (PHB)PHB 247
- port aggregation
 - configuration example 185
 - EtherChannel 183
- port mirroring 469
- ports
 - monitoring 469

Q

- QoS 241
- Quality of Service 241
- Querier (IGMP) 319

R

- RADIUS
 - port 1812 and 1645 120
 - port 1813 120
- redistributing routes 353, 377
- redundancy
 - active-active 429
- route aggregation 352, 377
- route maps 349
 - configuring 351
- Routed Ports 287
- Router ID
 - OSPF 400
- routers 277, 280
 - border 394
 - peer 394
 - port aggregation 183
 - switch-based routing topology 278
- routes, advertising 394
- routing 341
 - internal and external 394
- RSA keys 95
- Russia Class A electronic emission statement 493

S

- security
 - port mirroring 469
 - TACACS+ 106
 - VLANs 157
- segmentation. *See* IP subnets.
- segments. *See* IP subnets.
- service and support
 - before you call 481
- SNMP 28, 34, 395, 459
 - HP-OpenView 459
- SNMP Agent 459
- Source-Specific MulticastSSM 318

- SSH/SCP
 - configuring 96
 - RSA host and server keys 95
- summarizing routes 397
- switch failover 429

T

- TACACS+ 106
- Taiwan Class A electronic emission statement 493
- technical assistance 481
- text conventions 23
- trademarks 485
- typographic conventions 23

U

- United States FCC Class A notice 490
- upgrade, switch software 69

V

- virtual interface router (VIR) 427
- virtual link, OSPF 399
- Virtual Local Area Networks. *See* VLANs.
- virtual router ID numbering 428
- VLANs
 - broadcast domains 157
 - example showing multiple VLANs 175
 - ID numbers 159
 - routing 279
 - security 157
 - tagging ?? to 176
 - topologies 174
- VRRP (Virtual Router Redundancy Protocol)
 - active-active redundancy 429
 - overview 426
 - virtual interface router 427
 - virtual router ID numbering 428
 - vrid 427

