JUNIPER NETWORKS | Engineering Simplicity

Interfaces User Guide for Switches

Published
2025-04-04

JUNOS

Juniper Networks, Inc.
1133 Innovation Way
Sunnyvale, California 94089
USA
408-745-2000
www.juniper.net

Juniper Networks, the Juniper Networks logo, Juniper, and Junos are registered trademarks of Juniper Networks, Inc. in the United States and other countries. All other trademarks, service marks, registered marks, or registered service marks are the property of their respective owners.

Juniper Networks assumes no responsibility for any inaccuracies in this document. Juniper Networks reserves the right to change, modify, transfer, or otherwise revise this publication without notice.

Interfaces User Guide for Switches
Copyright © 2025 Juniper Networks, Inc. All rights reserved.

The information in this document is current as of the date on the title page.

YEAR 2000 NOTICE

Juniper Networks hardware and software products are Year 2000 compliant. Junos OS has no known time-related limitations through the year 2038. However, the NTP application is known to have some difficulty in the year 2036.

END USER LICENSE AGREEMENT

The Juniper Networks product that is the subject of this technical documentation consists of (or is intended for use with) Juniper Networks software. Use of such software is subject to the terms and conditions of the End User License Agreement ("EULA") posted at https://support.juniper.net/support/eula/. By downloading, installing or using such software, you agree to the terms and conditions of that EULA.

Table of Contents

About This Guide | xiv

1 **Configuring Interfaces**

Understanding Interfaces | 2

- Interfaces Overview for Switches | 2
- Understanding Interface Naming Conventions | 11
- Understanding Management Interfaces | 30

Physical Interface Properties | 32

- Configure Damping of Shorter Physical Interface Transitions | 33
- Accounting for Physical Interfaces | 34
  - Overview | 35
  - Configure an Accounting Profile for a Physical Interface | 35
  - How to Display the Accounting Profile | 37
- Enable SNMP Notifications on Physical Interfaces | 38
- Configuring Ethernet Loopback Capability | 39
- Configuring Short Reach Mode on QFX5100-48T | 40
- Configuring Flow Control | 41
- Setting the Mode on an SFP+ or SFP+ MACSec Uplink Module | 42
- Setting the Operating Mode on a 2-Port 40-Gigabit Ethernet QSFP+/100-Gigabit Ethernet QSFP28 Uplink Module | 43
- Configuring the Media Type on Dual-Purpose Uplink Ports | 45
- Disable a Physical Interface | 46
  - How to Disable a Physical Interface | 46
  - Example: Disable a Physical Interface | 47

Logical Interface Properties | 49

- Assign the Interface Address | 49

Add a Logical Unit Description to the Configuration | 51

Media MTU Overview | 52

Configure the Media MTU | 53

Protocol MTU | 54

Configure the Interface Bandwidth | 55

Enable or Disable SNMP Notifications on Logical Interfaces | 56

Overview of Accounting for the Logical Interface | 57

- Accounting Profiles Overview | 57
- Configure Accounting for the Logical Interface | 57
- Introduction to Displaying the Accounting Profile for the Logical Interface | 59

Disable a Logical Interface | 60

Damping Interfaces | 62

Physical Interface Damping Overview | 62

Configure Damping of Shorter Physical Interface Transitions | 70

Configure Damping of Aggregated Ethernet Interface Transitions | 71

Configure Damping of Longer Physical Interface Transitions | 72

Example: Configure Physical Interface Damping | 74

- Requirements | 74
- Overview | 74
- Configuration | 75
- Verification | 76

Interface Ranges | 77

Understanding Interface Ranges for Switches | 78

Configuring Interface Ranges for EX Series Switches with ELS | 81

- Configuring Interface Ranges on Switches | 82
- Expanded Interface Range Statements | 85
- Configuration Inheritance for Member Interfaces | 86
- Configuration Group Inheritance | 88
- Common Configuration Inheritance | 90
- Configuration Inheritance Priority | 90

Configuration Expansion Where Interface Range Is Used | **91**

Gigabit Ethernet Interface | 92

Speed and Autonegotiation | **93**

Configure Interface Speed on Switches | **93**

Configure Speed on EX2300-48MP and EX2300-24MP Switches | **94**

Configure Speed and Autonegotiation on QFX5100-48T Switches | **94**

Autonegotiation Support for EX4300-48MP Switches | **96**

Autonegotiation Support for EX4400 Switches | **97**

Autonegotiation Support for EX4100 Switches | **98**

Autonegotiation Support for EX4600-40F, QFX5110-48S and QFX5100-48S with JNP-SFPP-10GE-T Transceiver | **99**

Autonegotiation Support for QFX5120-48Y with JNP-SFPP-10GE-T Transceiver | **100**

Operating Speed of Interfaces on EX Switches | **101**

Configuring Gigabit and 10-Gigabit Ethernet Interfaces for EX4600 and QFX Series Switches | **105**

Configuring Port Mode on QFX5100-48S, QFX5100-48T, QFX5100-24Q, and EX4600 Switches | **106**

Configuring the Link Settings for Gigabit Ethernet Interfaces on QFX5100-48S, QFX5100-96S, and EX4600 Switches | **106**

Configuring Gigabit Ethernet Interfaces on QFX5100-48T Switches | **107**

Configuring the Link Settings for 10-Gigabit Ethernet Interfaces on QFX5100-48S, QFX5100-24Q, QFX5100-96S, and EX4600 Switches | **108**

Configuring the Link Settings for 10-Gigabit Ethernet Interfaces on QFX5100-48T Switches | **108**

Configuring the Link Settings for 10-Gigabit Ethernet Interfaces on QFX5120-48T Switches | **109**

Configuring the IP Options on QFX5100-48S, QFX5100-48T, QFX5100-24Q, and EX4600 Switches | **109**

Configuring the Link Settings for 1-Gigabit Ethernet Interfaces on EX4100 EX4100-F Multigigabit Switches | **110**

Configuring Gigabit Ethernet Interfaces for EX Series Switches with ELS support | **111**

Configuring VLAN Options and Interface Mode | **112**

Configuring the Link Settings | **112**

Configuring the IP Options | **115**

Configuring Gigabit and 10-Gigabit Ethernet Interfaces for OCX Series Switches | **116**

Configuring the Link Settings for Gigabit Ethernet and 10-Gigabit Ethernet Interfaces | **116**

Configuring the IP Options | **117**

Optical Transport Network (OTN) Interfaces | 118

Understanding the QFX10K-12C-DWDM Line Card | **118**

Configuring OTN Interface Options on QFX10K-12C-DWDM | **121**

Support for 400G-ZR Optics on QFX5220-32CD and QFX5130 | **127**

Energy Efficient Ethernet Interfaces | 129

Reduce Power Consumption on Interfaces using Energy Efficient Ethernet | **130**

Configure Energy Efficient Ethernet on Interfaces | **130**

Enable EEE on an EEE-Capable Base-T Copper Ethernet Port | **131**

Disable EEE on a Base-T Copper Ethernet Port | **131**

Verify EEE-Enabled Ports | **131**

Uplink Failure Detection | 134

Overview of Uplink Failure Detection | **134**

Configuring Interfaces for Uplink Failure Detection | **137**

Example: Configuring Interfaces for Uplink Failure Detection | **139**

Requirements | **139**

Overview and Topology | **139**

Configuring Uplink Failure Detection on Both Switches | **141**

Verification | **144**

Verifying That Uplink Failure Detection Is Working Correctly | **145**

Targeted Broadcast | 147

Understanding Targeted Broadcast | **147**

Understanding IP Directed Broadcast | **148**

Configure Targeted Broadcast | **150**

Configure Targeted Broadcast and Its Options | **150**

Display Targeted Broadcast Configuration Options | **152**

Configuring IP Directed Broadcast (CLI Procedure) | **153**

Example: Configuring IP Directed Broadcast on a Switch | **155**

Requirements | **155**

Overview and Topology | **156**

Configuring IP Directed Broadcast for non-ELS Switches | **157**

Configuring IP Directed Broadcast for Switches with ELS Support | **160**

Verifying IP Directed Broadcast Status | **164**

ARP | 165

Static ARP Table Entries Overview | **165**

Configuring Static ARP Table Entries For Mapping IP Addresses to MAC Addresses | **165**

Restricted and Unrestricted Proxy ARP Overview | **168**

Configuring Restricted and Unrestricted Proxy ARP | **171**

Configuring Gratuitous ARP | **172**

Use of Resillient Hashing to Minimize Flow Remapping | 173

Limitations and Caveats for Resilient Hashing | **176**

Configuring Resilient Hashing for ECMP | **176**

Configuring Resilient Hashing for Aggregated Ethernet Interfaces | **177**

Generic Routing Encapsulation (GRE) | 178

Understanding Generic Routing Encapsulation | **178**

Configuring Generic Routing Encapsulation Tunneling | **183**

Configuring a GRE Tunnel | **183**

Verifying That Generic Routing Encapsulation Tunneling Is Working Correctly | **185**

Understanding Per-Packet Load Balancing | 186

Understanding ECMP Groups | 189

Configuring Consistent Load Balancing for ECMP Groups | **189**

Understanding Consistent Load Balancing Through Resilient Hashing on ECMP Groups | **192**

2 **Port Speed for Switches**

Port Speed Overview | 194

Configure Port Speed at Chassis Level and Interface Level | 196

Port Speed on EX Switches | 199

Port Speed on EX4650-48Y Switches | **200**

Port Speed on EX4400 Switches | **203**

Port Speed on EX4100 Switches | **216**

Port Speed on EX4100-H Switches | **229**

Port Speed on EX4000-12MP, EX4000-24MP, and EX4000-48MP Switches | **231**

Port Speed on QFX Switches | 234

Port Speed on QFX5100-24Q Switches | **234**

Port Speed on QFX5110-48S Switches | **236**

Port Speed on QFX5120-32C Switches | **237**

Port Speed on QFX5120-48T Switches | **238**

Port Speed on QFX5120-48Y Switches | **239**

Port Speed on QFX5120-48YM Switches | **240**

Port Speed on QFX5130-32CD Switches | **241**

Port Speed on QFX5130-48C/QFX5130-48CM Switches | **243**

Port Speed on QFX5200-32C Switches | **250**

Port Speed on QFX5210-64C Switches | **250**

Port Speed on QFX5230-64CD Switches | **251**

Port Speed on QFX5240 Switches | **255**

Port Speed on QFX5700 Switches | **259**

3 **Configuring Aggregated Ethernet Interfaces**

Aggregated Ethernet Interfaces | 261

Understanding Aggregated Ethernet Interfaces and LACP for Switches | **262**

Forcing LAG Links or Interfaces with Limited LACP Capability to Be Up | **268**

Configuring an Aggregated Ethernet Interface | **268**

Configuring Tagged Aggregated Ethernet Interfaces | **270**

Configuring Untagged Aggregated Ethernet Interfaces | **270**

Configuring the Number of Aggregated Ethernet Interfaces on the Device (Enhanced Layer 2 Software) | **271**

Example: Configuring Aggregated Ethernet Interfaces | **272**

Deleting an Aggregated Ethernet Interface | **274**

Understanding Local Link Bias | **274**

Configuring Local Link Bias | **276**

Understanding Local Minimum Links | **277**

Troubleshooting an Aggregated Ethernet Interface | **280**

- Show Interfaces Command Shows the LAG is Down | **281**
- Logical Interface Statistics Do Not Reflect All Traffic | **281**
- IPv6 Interface Traffic Statistics Are Not Supported | **282**
- SNMP Counters ifHCInBroadcastPkts and ifInBroadcastPkts Are Always 0 | **282**

Configuring Link Aggregation | **283**

- Creating an Aggregated Ethernet Interface | **284**
- Configuring the VLAN Name and VLAN ID Number | **285**
- Configuring Aggregated Ethernet LACP (CLI Procedure) | **285**

Aggregated Ethernet Link Protection | **288**

- Configuring Link Protection for Aggregated Ethernet Interfaces | **289**
- Configuring Primary and Backup Links for Link Aggregated Ethernet Interfaces | **289**
- Reverting Traffic to a Primary Link When Traffic is Passing Through a Backup Link | **290**
- Disabling Link Protection for Aggregated Ethernet Interfaces | **291**

Configure the Aggregated Ethernet Link Speed | **291**

Configuring Periodic Rebalancing of Subscribers in an Aggregated Ethernet Interface | **294**

Example: Configuring Aggregated Ethernet High-Speed Uplinks Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch | **295**

- Requirements | **296**
- Overview and Topology | **296**
- Configuration | **299**
- Verification | **303**
- Troubleshooting | **304**

Example: Configuring Link Aggregation Between a QFX Series Product and an Aggregation Switch | **305**

- Requirements | **306**
- Overview and Topology | **306**
- Configuration | **307**
- Verification | **311**
- Troubleshooting | **312**

Configuring Aggregated Ethernet LACP | **313**

- Configuring the LACP Interval | **315**
- Configuring LACP Link Protection | **315**
- Configuring LACP System Priority | **317**
- Configuring LACP System Identifier | **317**
- Configuring LACP administrative Key | **317**
- Configuring LACP Port Priority | **318**
- Tracing LACP Operations | **318**
- LACP Limitations | **319**
- Example: Configuring Aggregated Ethernet LACP | **319**

Configuring LACP Link Protection of Aggregated Ethernet Interfaces for Switches | **323**

- Configuring LACP Link Protection for a Single Link at the Global Level | **325**
- Configuring LACP Link Protection for a Single Link at the Aggregated Interface Level | **325**
- Configuring Subgroup Bundles to Provide LACP Link Protection to Multiple Links in an Aggregated Ethernet Interface | **326**

Configuring LACP Hold-UP Timer to Prevent Link Flapping on LAG Interfaces | **329**

Verifying That LACP Is Configured Correctly and Bundle Members Are Exchanging LACP Protocol Packets | **330**

- Verifying the LACP Setup | **330**
- Verifying That LACP Packets Are Being Exchanged | **331**

Example: Configuring Aggregated Ethernet High-Speed Uplinks with LACP Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch | **332**

- Requirements | **333**
- Overview and Topology | **333**
- Configuring LACP for the LAGs on the Virtual Chassis Access Switch | **334**
- Configuring LACP for the LAGs on the Virtual Chassis Distribution Switch | **335**
- Verification | **336**

Troubleshooting | 339

Example: Configuring Link Aggregation with LACP Between a QFX Series Product and an Aggregation Switch | 340

Requirements | 340

Overview and Topology | 341

Configuring LACP for the LAG on the QFX Series | 341

Verification | 342

Troubleshooting | 345

Understanding Independent Micro BFD Sessions for LAG | 346

Configuration Guidelines for Micro-BFD Sessions | 347

Configuring Micro BFD Sessions for LAG | 348

Understanding the Algorithm Used to Hash LAG Bundle and Egress Next-Hop ECMP Traffic | 355

Configuring the Fields in the Algorithm Used To Hash LAG Bundle and ECMP Traffic (CLI Procedure) | 363

Configuring the Hashing Algorithm to Use Fields in the Layer 2 Header for Hashing | 364

Configuring the Hashing Algorithm to Use Fields in the IP Payload for Hashing | 365

Configuring the Hashing Algorithm to Use Fields in the IPv6 Payload for Hashing | 365

Configuring Other Hashing Parameters | 366

Load Balancing for Aggregated Ethernet Interfaces | 368

Load Balancing and Ethernet Link Aggregation Overview | 368

Configuring Load Balancing Based on MAC Addresses | 369

Configuring Load Balancing on a LAG Link | 371

Example: Configuring Load Balancing on a LAG Link | 371

Understanding Multicast Load Balancing on Aggregated 10-Gigabit Links for Routed Multicast Traffic on EX8200 Switches | 372

Example: Configuring Multicast Load Balancing for Use with Aggregated 10-Gigabit Ethernet Interfaces on EX8200 Switches | 377

Requirements | 378

Overview and Topology | 378

Configuration | 380

Verification | 383

Dynamic Load Balancing (DLB) | 385

Dynamic Load Balancing Overview | **385**

Configuring Dynamic Load Balancing | **388**

Example: Configure Dynamic Load Balancing | **390**

- Requirements | **390**
- Overview | **391**
- Configuration | **392**
- Verification | **396**

Selective Dynamic Load Balancing (DLB) | **398**

- Selective DLB Overview | **398**
- Selective DLB in AI-ML Data Centers | **399**
- Configuration | **399**
- Example: Selectively Enable DLB with a Firewall Filter Match Condition | **403**

Customize Egress Port Link Quality Metrics for DLB | **404**

- Overview | **404**
- Configuration | **405**

Configure Flowset Table Size in DLB Flowlet Mode | **406**

- Overview | **407**
- Configuration | **408**

Reactive Path Rebalancing | **408**

- Overview | **409**
- Configuration | **409**

Global Load Balancing (GLB) | 413

GLB Overview | **413**

GLB in AI-ML Data Centers | **414**

Configuration | **414**

- Considerations | **414**
- Configure GLB | **415**

4 **Flexible Ethernet Services Encapsulation**

Flexible Ethernet Services Encapsulation | 418

Understanding Flexible Ethernet Services Encapsulation on Switches | **418**

Configuring Flexible Ethernet Services Encapsulation to Support the Service Provider and Enterprise Styles of Configuration | **421**

Configure Flexible Ethernet Services Encapsulation to Include Layer 2 Interface Support with Other Encapsulations | **424**

Configure Flexible Ethernet Services Encapsulation to Support Multiple Logical Interfaces on the Same Physical Interface Mapped to the Same Bridge Domain | **426**

5 **Monitoring and Troubleshooting Information**

Monitoring Interfaces | 431

Monitoring Interface Status and Traffic | **431**

Monitoring System Process Information | **432**

Monitoring System Properties | **433**

Monitor Statistics for a Fast Ethernet or Gigabit Ethernet Interface | **436**

Trace Operations of the Interface Process | **438**

Troubleshooting Interfaces | 440

Troubleshooting Network Interfaces | **440**

Statistics for logical interfaces on Layer 2 interfaces are not accurate | **441**

The interface on the port in which an SFP or SFP+ transceiver is installed in an SFP or SFP+ module is down | **441**

Diagnosing a Faulty Twisted-Pair Cable (CLI Procedure) | **442**

Troubleshooting Uplink Ports on EX2300 Switches | **446**

Speeds 10-Mbps and 100-Mbps not supported on uplink ports 4 and 5 on EX2300-48MP switches | **447**

6 **Configuration Statements and Operational Commands**

Common Output Fields Description | 450

Junos CLI Reference Overview | 460

About This Guide

Use this guide to configure, monitor, and troubleshoot the various supported Ethernet Interfaces, including aggregated Ethernet Interfaces on Juniper Networks switches.

1
CHAPTER

Configuring Interfaces

IN THIS CHAPTER

- Understanding Interfaces | **2**
- Physical Interface Properties | **32**
- Logical Interface Properties | **49**
- Damping Interfaces | **62**
- Interface Ranges | **77**
- Gigabit Ethernet Interface | **92**
- Optical Transport Network (OTN) Interfaces | **118**
- Energy Efficient Ethernet Interfaces | **129**
- Uplink Failure Detection | **134**
- Targeted Broadcast | **147**
- ARP | **165**
- Use of Resillient Hashing to Minimize Flow Remapping | **173**
- Generic Routing Encapsulation (GRE) | **178**
- Understanding Per-Packet Load Balancing | **186**
- Understanding ECMP Groups | **189**

Understanding Interfaces

IN THIS SECTION

- Interfaces Overview for Switches | 2
- Understanding Interface Naming Conventions | 11
- Understanding Management Interfaces | 30

Junos OS supports different types of interfaces on which the devices function. The following topics provide information of types of interfaces used, the naming conventions and the usage of management interfaces by Juniper Networks.

Interfaces Overview for Switches

IN THIS SECTION

- Network Interfaces for EX Series | 3
- Special Interfaces for EX Series | 4
- Network Interfaces for EX4600, NFX Series, QFX Series, QFabric System | 6
- Special Interfaces for EX4600, NFX Series, QFX Series, QFabric System | 8
- Network Interfaces for OCX Series | 9
- Special Interfaces for OCX Series | 10

Juniper Networks devices have two types of interfaces: network interfaces and special interfaces. This topic provides brief information about these interfaces. For additional information, see the Junos OS Network Interfaces Library for Routing Devices.

Network Interfaces for EX Series

Network interfaces connect to the network and carry network traffic. Table 1 on page 3 lists the types of network interfaces supported on EX Series switches.

Table 1: Network Interfaces Types and Purposes for EX Series

Type	Purpose
Aggregated Ethernet interfaces	All EX Series switches allow you to group Ethernet interfaces at the physical layer to form a single link layer interface. This group is also known as a *link aggregation group (LAG)* or *bundle*. These aggregated Ethernet interfaces help to balance traffic and increase the uplink bandwidth. See "Understanding Aggregated Ethernet Interfaces and LACP for Switches" on page 262.
LAN access interfaces	Use these EX Series switch interfaces to connect the following to the network: • PC • Laptop • File server • Printer When you power on an EX Series switch and use the factory-default configuration, the software automatically configures interfaces in access mode for each of the network ports. The default configuration also enables autonegotiation for both speed and link mode.
Power over Ethernet (PoE) interfaces	EX Series switches provide PoE network ports with various switch models. Use these ports to connect VoIP telephones, wireless access points, video cameras, and point-of-sale devices to safely receive power from the same access ports that are used to connect personal computers to the network. PoE interfaces are enabled by default in the factory configuration. See Understanding PoE on EX Series Switches.
Trunk interfaces	You can connect EX Series access switches to a distribution switch or customer-edge (CE) switches or routers. To use a port for this type of connection, you must explicitly configure the network interface for trunk mode. You must also configure the interfaces from the distribution switch or CE switch to the access switches for trunk mode.

Special Interfaces for EX Series

Table 2 on page 4 lists the types of special interfaces supported on EX Series switches.

Table 2: Special Interfaces Types and Purposes for EX Series

Type	Purpose
Console port	Each EX Series switch has a serial port, labeled **CON** or **CONSOLE**, for connecting tty-type terminals to the switch using standard PC-type tty cables. The console port does not have a physical address or IP address associated with it. However, it is an interface since it provides access to the switch. On an EX3300 *Virtual Chassis*, an EX4200 Virtual Chassis, or an EX4500 Virtual Chassis, you can access the primary device and configure all members of the Virtual Chassis through any member's console port. For more information about the console port in a Virtual Chassis, see Understanding Global Management of a Virtual Chassis.
Loopback	All EX Series switches have this software-only virtual interface that is always up. The loopback interface provides a stable and consistent interface and IP address on the switch.
Management interface	The Juniper Networks Junos operating system (Junos OS) for EX Series switches automatically creates the switch's management Ethernet interface, me0. The management Ethernet interface provides an out-of-band method for connecting to the switch. To use me0 as a management port, you must configure its logical port, me0.0, with a valid IP address. You can connect to the management interface over the network using utilities such as SSH or Telnet. SNMP can use the management interface to gather statistics from the switch. (The management interface me0 is analogous to the fxp0 interfaces on routers running Junos OS.) See "Understanding Management Interfaces" on page 30.
Integrated Routing and Bridging (IRB) Interface or *Routed VLAN Interface* (RVI)	EX Series switches use an integrated routing and bridging (IRB) interface or Routed VLAN Interface (RVI) to route traffic from one broadcast domain to another and to perform other Layer 3 functions such as traffic engineering. These functions are typically performed by a router interface in a traditional network. The IRB interface or RVI functions as a logical router, eliminating the need for having both a switch and a router. Configure these interfaces as part of a broadcast domain or Virtual Private LAN Service (VPLS) routing instance for L3 traffic to be routed from. See Understanding Integrated Routing and Bridging.

Table 2: Special Interfaces Types and Purposes for EX Series *(Continued)*

Type	Purpose
Virtual Chassis port (VCP) interfaces	Virtual Chassis ports (VCPs) are used to interconnect switches in a *Virtual Chassis*: • EX3300 switches—Port 2 and port 3 of the SFP+ uplink ports are preconfigured as VCPs and can be used to interconnect up to six EX3300 switches in an EX3300 Virtual Chassis. See *Setting an Uplink Port on an EX Series or QFX Series Switch as a Virtual Chassis Port*. • EX4100, EX4100-24MP, EX4100-48MP, and EX4100-F switches—Each EX4100, EX4100-24MP, EX4100-48MP, or EX4100-F switch has dedicated VCP ports. You cannot use any other ports on EX4100 switches as VCPs. See EX4100/EX4100-F Switches in a Virtual Chassis. • EX4200 and EX4500 switches—Each EX4200 switch or each EX4500 switch with a Virtual Chassis module installed has two dedicated VCPs on its rear panel. These ports can be used to interconnect up to ten EX4200 switches in an EX4200 Virtual Chassis, up to ten EX4500 switches in an EX4500 Virtual Chassis, and up to ten switches in a mixed EX4200 and EX4500 Virtual Chassis. When you power on switches that are interconnected in this manner, the software automatically configures the VCP interfaces for the dedicated ports that have been interconnected. These VCP interfaces are not configurable or modifiable. See *Understanding the High-Speed Interconnection of the Dedicated Virtual Chassis Ports Connecting EX4200, EX4500, and EX4550 Member Switches*. You can also interconnect EX4200 and EX4500 switches by using uplink module ports. Using uplink ports allows you to connect switches over longer distances than you can by using the dedicated VCPs. To use the uplink ports as VCPs, you must explicitly configure the uplink module ports on the members you want to connect as VCPs. See *Setting an Uplink Port on an EX Series or QFX Series Switch as a Virtual Chassis Port* . • EX4300 switches—All QSFP+ ports are configured as VCPs by default. See *Understanding EX Series Virtual Chassis*. You can also interconnect EX4300 switches into a Virtual Chassis by using SFP+ uplink module ports as VCPs. Using uplink ports as VCPs allows you to connect switches over longer distances than you can by using the QSFP+ ports as VCPs. To use the uplink ports as VCPs, you must explicitly configure the uplink module ports on the members you want to connect as VCPs. See *Setting an Uplink Port on an EX Series or QFX Series Switch as a Virtual Chassis Port*.

Table 2: Special Interfaces Types and Purposes for EX Series *(Continued)*

Type	Purpose
	• EX8200 switches—EX8200 switches can be connected to an XRE200 External Routing Engine to create an EX8200 Virtual Chassis. The XRE200 External Routing Engine has dedicated VCPs that connect to ports on the internal Routing Engines of the EX8200 switches and can connect to another XRE200 External Routing Engine for redundancy. These ports require no configuration. . You can also connect two members of an EX8200 Virtual Chassis so that they can exchange Virtual Chassis Control Protocol (VCCP) traffic. To do so, you explicitly configure network ports on the EX8200 switches as VCPs.
Virtual management Ethernet (VME) interface	EX3300, EX4200, EX4300, and EX4500 switches have a VME interface. This is a *logical interface* that is used for Virtual Chassis configurations and allows you to manage all the members of the Virtual Chassis through the primary device. For more information about the VME interface, see *Understanding Global Management of a Virtual Chassis*. EX8200 switches do not use a VME interface. An EX8200 Virtual Chassis is managed through the management Ethernet (me0) interface on the XRE200 External Routing Engine.

Network Interfaces for EX4600, NFX Series, QFX Series, QFabric System

Network interfaces connect to the network and carry network traffic. Table 3 on page 6 lists the types of network interfaces supported.

Table 3: Network Interfaces Types and Purposes for EX4600, NFX Series, QFX Series, QFabric System

Type	Purpose
Aggregated Ethernet interfaces	Group Ethernet interfaces at the physical layer to form a single link-layer interface, also known as a *link aggregation group (LAG)* or *bundle*. These aggregated Ethernet interfaces help to balance traffic and increase the uplink bandwidth.

Table 3: Network Interfaces Types and Purposes for EX4600, NFX Series, QFX Series, QFabric System *(Continued)*

Type	Purpose
Channelized Interfaces	Depending on the device and software package, 40-Gbps QSFP+ ports can be configured to operate as the following types of interfaces: • 10-Gigabit Ethernet interfaces (*xe*) • 40-Gigabit Ethernet interfaces (*et* and *xle*) • 40-Gigabit data plane uplink interfaces (*fte*) When an *et* port is channelized to four *xe* ports, a colon is used to signify the four separate channels. For example, on a QFX3500 standalone switch with port 2 on PIC 1 configured as four 10-Gigabit Ethernet ports, the interface names are *xe-0/1/2:0*, *xe-0/1/2:1*, *xe-0/1/2:2*, and *xe-0/1/2:3* **NOTE**: You cannot configure channelized interfaces to operate as Virtual Chassis ports.
Ethernet Interfaces	Configure Gigabit Ethernet, 10-Gigabit Ethernet, 40-Gigabit Ethernet interfaces to connect to other servers, storage, and switches. You can configure 40-Gigabit data plane uplink ports to connect a Node device to an Interconnect devices as well as for Virtual Chassis ports (VCPs).
Fibre Channel interfaces	Use Fibre Channel interfaces to connect the switch to a Fibre Channel over Ethernet (FCoE) forwarder or a Fibre Channel switch in a storage area network (SAN). You can configure Fibre Channel interfaces only on ports 0 through 5 and 42 through 47 on QFX3500 devices. Fibre Channel interfaces do not forward Ethernet traffic. See *Overview of Fibre Channel*.
LAN access interfaces	Use these interfaces to connect to other servers, storage, and switches. When you power on a QFX Series product and use the factory-default configuration, the software automatically configures interfaces in access mode for each of the network ports.
Multichassis aggregated Ethernet (MC-AE) interfaces	Group a LAG on one standalone switch with a LAG on another standalone switch to create a MC-AE. The MC-AE provides load balancing and redundancy across the two standalone switches.

Table 3: Network Interfaces Types and Purposes for EX4600, NFX Series, QFX Series, QFabric System *(Continued)*

Type	Purpose
Tagged-access mode interfaces	Use tagged-access interfaces to connect a switch to an access layer device. Tagged-access interfaces can accept VLAN-tagged packets from multiple VLANs.
Trunk interfaces	Use trunk interfaces to connect to other switches or routers. To use a port for this type of connection, you must explicitly configure the port interface for trunk mode. The interfaces from the switches or routers must also be configured for trunk mode. In this mode, the interface can be in multiple VLANs and accept tagged packets from multiple devices. Trunk interfaces typically connect to other switches and to routers on the LAN.
Virtual Chassis ports (VCPs)	You can use Virtual Chassis ports to send and receive Virtual Chassis Control Protocol (VCCP) traffic, and to create, monitor, and maintain the Virtual Chassis. On QFX3500, QFX3600, QFX5100, QFX5110, QFX5200, and EX4600 standalone switches, you can configure 40-Gigabit Ethernet QSFP+ uplink ports (non-channelized) or fixed SFP+ 10-Gigabit Ethernet ports as VCPs by issuing the `request virtual-chassis-vc-port-set` CLI command. QFX5110 switches also support configuring 100-Gigabit QSFP28 ports as VCPs.

Special Interfaces for EX4600, NFX Series, QFX Series, QFabric System

Table 4 on page 8 lists the types of special interfaces supported.

Table 4: Special Interfaces Types and Purposes supported on EX4600, NFX Series, QFX Series, QFabric System

Type	Purpose
Console port	Each device has a serial console port, labeled **CON** or **CONSOLE**, for connecting tty-type terminals to the switch. The console port does not have a physical address or IP address associated with it. However, it is an interface in the sense that it provides access to the switch.
Loopback interface	A software-only virtual interface that is always up. The loopback interface provides a stable and consistent interface and IP address on the switch.

Table 4: Special Interfaces Types and Purposes supported on EX4600, NFX Series, QFX Series, QFabric System *(Continued)*

Type	Purpose
Management interface	The management Ethernet interface provides an out-of-band method for connecting to a standalone switch and QFabric system. **NOTE**: On OCX Series switches, the em0 management interface always has the status up in show command outputs, even if the physical port is empty. The me0 interface is a virtual interface between Junos and the host operating system, therefore its status is independent from the status of the physical port.
Routed VLAN interfaces (*RVI* and IRB interfaces)	Layer 3 routed VLAN interfaces (called RVI in the original CLI, and called IRB in Enhanced Layer 2 Software) route traffic from one broadcast domain to another and perform other Layer 3 functions such as traffic engineering. These functions are typically performed by a router interface in a traditional network. The RVI or IRB functions as a logical router, eliminating the need for having both a switch and a router. The RVI or IRB must be configured as part of a broadcast domain or virtual private LAN service (VPLS) routing instance for Layer 3 traffic to be routed out of it.

Network Interfaces for OCX Series

Network interfaces connect to the network and carry network traffic. Table 5 on page 9 lists the types of network interfaces supported.

Table 5: Network Interfaces Types and Purposes for OCX Series

Type	Purpose
Aggregated Ethernet interfaces	Group Ethernet interfaces at the physical layer to form a single link-layer interface, also known as a *link aggregation group (LAG)* or *bundle*. These aggregated Ethernet interfaces help to balance traffic and increase the uplink bandwidth.
Ethernet Interfaces	Configure Gigabit Ethernet, 10-Gigabit Ethernet, 40-Gigabit Ethernet interfaces to connect to other servers, storage, and switches.

Special Interfaces for OCX Series

Table 6 on page 10 lists the types of special interfaces supported.

Table 6: Special Interfaces Types and Purposes for OCX Series

Type	Purpose
Console port	Each device has a serial console port, labeled **CON** or **CONSOLE**, for connecting tty-type terminals to the switch. The console port does not have a physical address or IP address associated with it. However, it is an interface in the sense that it provides access to the switch.
Loopback interface	A software-only virtual interface that is always up. The loopback interface provides a stable and consistent interface and IP address on the switch.
Management interface	The management Ethernet interface provides an out-of-band method for connecting to a standalone switch and QFabric system. **NOTE**: On OCX Series switches, the em0 management interface always has the status `up` in `show` command outputs, even if the physical port is empty. The me0 interface is a virtual interface between Junos and the host operating system, therefore its status is independent from the status of the physical port.

SEE ALSO

- EX2200 Switches Hardware Overview
- EX3200 System Overview
- EX3300 Switches Hardware Overview
- *EX4200 Switches Hardware Overview*
- *EX4300 Switches Hardware Overview*
- EX4500 Switches Hardware Overview
- EX6210 Switch Hardware Overview
- EX8208 Switch Hardware Overview
- EX8216 Switch Hardware Overview
- Understanding Layer 3 Logical Interfaces
- Understanding Layer 3 Subinterfaces

Understanding Interface Naming Conventions

IN THIS SECTION

- Physical Part of an Interface Name for EX Series | 11
- Logical Part of an Interface Name for EX Series | 14
- Wildcard Characters in Interface Names for EX Series | 14
- Physical Part of an Interface Name for QFX series, NFX Series, EX4600, QFabric System | 14
- Logical Part of an Interface Name on a Switch Running QFabric Software Package for QFX series, NFX Series, EX4600, QFabric System | 28
- Logical Part of a Channelized Interface Name on a Switch Running Enhanced Layer 2 Software for QFX series, NFX Series, EX4600, QFabric System | 29
- Wildcard Characters in Interface Names for QFX series, NFX Series, EX4600, QFabric System | 29
- Physical Part of an Interface Name for OCX1100 | 29
- Wildcard Characters in Interface Names for OCX1100 | 30

The EX Series, QFX Series, NFX Series, OCX1100, QFabric System, and EX4600 devices use a naming convention for defining the interfaces that are similar to that of other platforms running under Juniper Networks Junos OS. This topic provides brief information about the naming conventions used for interfaces on the QFX Series and on EX4600 switches.

For detailed information on interface naming like physical part, logical part, and channel part of the interfaces, see Interface Naming Overview.

This topic describes:

Physical Part of an Interface Name for EX Series

Network interfaces in Junos OS are specified as follows:

type-fpc / pic / port

EX Series switches apply this convention as follows:

- *type*-EX Series interfaces use the following media types:
  - ge—Gigabit Ethernet interface
  - xe—10 Gigabit Ethernet interface

    - et—40 Gigabit Ethernet interface

- *fpc*—Flexible PIC Concentrator. EX Series interfaces use the following convention for the FPC number in interface names:

    - On an EX2200 switch, an EX2300, an EX3200 switch, a standalone EX3300 switch, a standalone EX3400 switch, a standalone EX4200 switch, a standalone EX4300 switch, a standalone EX4500, and a standalone EX4550 switch, FPC refers to the switch itself. The FPC number is **0** by default on these switches.

    - On an EX3300 *Virtual Chassis*, an EX3400 Virtual Chassis, an EX4200 Virtual Chassis, an EX4300 Virtual Chassis, an EX4500 Virtual Chassis, an EX4550 Virtual Chassis, or a mixed Virtual Chassis, the FPC number indicates the member ID of the switch in the Virtual Chassis.

    - On EX4100 and EX4100-F switches, the FPC number ranges from **0** to **9**. On a standalone EX4100 or EX4100-F switch, FPC refers to the switch. The FPC number is **0** by default on the standalone switches.

    - On EX4100 and EX4100-F Virtual Chassis, the FPC number indicates the member ID of the switch in the Virtual Chassis.

    - On an EX6200 switch and a standalone EX8200 switch, the FPC number indicates the slot number of the line card that contains the physical interface. On an EX6200 switch, the FPC number also indicates the slot number of the Switch Fabric and Routing Engine (SRE) module that contains the uplink port.

    - On an EX8200 Virtual Chassis, the FPC number indicates the slot number of the line card on the Virtual Chassis. The line card slots on Virtual Chassis member 0 are numbered 0 through 15; on member 1, they are numbered 16 through 31, and so on.

    - On EX9251 switch, the FPC number is always **0**.

    - The EX9253 switch does not have actual FPCs—the line cards are the FPC equivalents on the switch. In FPC (n), n is a value in the range of 0-1. The value corresponds to the line card slot number in which the line card is installed.

    - On an EX29204 switch, switch does not have actual FPCs—the line cards are the FPC equivalents on the switch. The value ranges from 0-2, and it corresponds to the line card slot number in which the line card is installed.

- *pic*—EX Series interfaces use the following convention for the PIC (*Physical Interface Card*) number in interface names:

    - On EX2200, EX2300, EX3200, EX3300, EX4200, EX4500 switch, and EX4550 switches, the PIC number is **0** for all built-in interfaces (interfaces that are not uplink ports).

- On EX2200, EX2300, EX3200, EX3300, and EX4200 switches, the PIC number is **1** for uplink ports.
- On EX3400 switches, the PIC number is **0** for built-in network ports, **1** for built-in QSFP+ ports (located on the rear panel of the switch), and **2** for uplink module ports.
- On EX4100 and EX4100-F switches, the PIC number ranges from **0** to **2**. The PIC number is **0** for built-in network ports, **1** for SFP28/SFP+ dedicated Virtual Chassis ports, and **2** for SFP/SFP+ uplink ports.
- On EX4300 switches, the PIC number is **0** for built-in network ports, **1** for built-in QSFP+ ports (located on the rear panel of the switch), and **2** for uplink module ports.
- On EX4500 switches, the PIC number is **1** for ports on the left-hand uplink module and **2** for ports on the right-hand uplink module.
- On EX4550 switches, the PIC number is **1** for ports in the expansion module or Virtual Chassis module installed in the module slot on the front panel of the switch and **2** for those in the expansion module or Virtual Chassis module installed in the module slot on the rear panel of the switch.
- On EX6200 and EX8200 switches, the PIC number is always **0**.
- On EX9251 and EX9253 switches, the PIC number is **0** for built-in network ports, **1** for built-in QSFP+ ports (located on the rear panel of the switch).
- On EX9204 switches, the PIC number ranges from 0-3.

- *port*—EX Series interfaces use the following convention for port numbers:
  - On EX2200, EX2300, EX3200, EX3300, EX3400, EX4200, EX4300, EX4500, and EX4550 switches, built-in network ports are numbered from left to right. On models that have two rows of ports, the ports on the top row start with **0** followed by the remaining even-numbered ports, and the ports on the bottom row start with **1** followed by the remaining odd-numbered ports.
  - Uplink ports in EX2200, EX3200, EX3300, EX3400, EX4200, EX4300, EX4500, and EX4550 switches are labeled from left to right, starting with **0**.
  - On EX4100 and EX4100-F switches, the uplink ports are labeled from 0 to 3. The Virtual Chassis ports are also labeled from 0 to 3. The downlink ports are labeled from 0 to 47 (for EX4100-48P, EX4100-48T, EX4100-F-48P, and EX4100-F-48T switches) and from 0 to 23 (for EX4100-24P, EX4100-24T, EX4100-F-24P and EX4100-F-24T switches).
  - On EX6200 and EX8200 switches, the network ports are numbered from left to right on each line card. On line cards that have two rows of ports, the ports on the top row start with **0** followed by the remaining even-numbered ports, and the ports on the bottom row start with **1** followed by the remaining odd-numbered ports.

- Uplink ports on an SRE module in an EX6200 switch are labeled from left to right, starting with **0**.
- EX9251 Switch has eight 10-Gigabit Ethernet ports and four rate-selectable ports that you can configure as 100-Gigabit Ethernet ports or 40-Gigabit Ethernet ports; each rate-selectable port can be configured as four 10-Gigabit Ethernet ports by using a breakout cable. The 10-Gigabit Ethernet ports support SFP+ transceivers and rate-selectable ports support QSFP28 and QSFP+ transceivers.
- EX9253 contains six built-in QSFP+ ports, each of which can house QSFP+ pluggable transceivers and 12 built-in QSFP28 ports, each of which can house QSFP28 pluggable transceivers.

Logical Part of an Interface Name for EX Series

The logical unit part of the interface name corresponds to the logical unit number, which can be a number from 0 through 16384. In the virtual part of the name, a period (.) separates the port and logical unit numbers: *type-fpc/pic/port.logical-unit-number*. For example, if you issue the `show ethernet-switching interfaces` command on a system with a default VLAN, the resulting display shows the logical interfaces associated with the VLAN:

```
Interface   State    VLAN members          Blocking
ge-0/0/0.0  down     remote-analyzer       unblocked
ge-0/0/1.0  down     default               unblocked
ge-0/0/10.0 down     default               unblocked
```

Wildcard Characters in Interface Names for EX Series

In the `show interfaces` and `clear interfaces` commands, you can use wildcard characters in the ***interface-name*** option to specify groups of interface names without having to type each name individually. You must enclose all wildcard characters except the asterisk (*) in square brackets [ ].

Physical Part of an Interface Name for QFX series, NFX Series, EX4600, QFabric System

Interfaces in Junos OS are specified as follows:

device-name:*type*-*fpc*/*pic*/*port*

The convention is as follows (and platform support depends on the Junos OS release in your installation):

- *device-name*—(QFabric systems only) The *device-name* is either the serial number or the alias of the QFabric system component, such as a Node device, Interconnect device, or QFabric infrastructure. The name can contain a maximum of 128 characters and cannot contain any colons.
- *type*—The QFX Series and EX4600 device interfaces use the following media types:
  - **fc**—Fibre Channel interface
  - **ge**—Gigabit Ethernet interface
  - **xe**—10-Gigabit Ethernet interface
  - *sxe*—10-Gigabit Service interface. *sxe* is an internal interface and user must not configure this interface. It supports L2 and L3 configurations like VLANs and IP address.
  - **xle**—40-Gigabit Ethernet interface (QFX3500, QFX3600, and QFX5100 switches running a QFabric software package)
  - **et**—25-Gigabit Ethernet interface (QFX5120 and QFX5200 switches)
  - **et**—40-Gigabit Ethernet interface (QFX3500, QFX3600, QFX5100, QFX5200, QFX10000, and EX4600 switches running Enhanced Layer 2 Software)
  - **et**—100-Gigabit Ethernet interface (QFX5200 and QFX10000 switches running Enhanced Layer 2 Software)
  - **fte**—40-Gigabit data plane uplink interface (QFX3500, QFX3600, and QFX5100 switches running a QFabric software package)
  - **me**—Management interface
  - **em**—Management interface on QFX5100 and EX4600 switches.
- *fpc*—Flexible PIC Concentrator. QFX Series interfaces use the following convention for the FPC number in interface names:
  - On QFX3500, QFX3600, QFX5100 devices running a QFabric software package, and QFX10002 switches, the FPC number is always 0.

    The FPC number indicates the slot number of the line card that contains the physical interface.
  - On QFX3500, QFX3600, QFX5100, QFX5200, EX4600, QFX10002, QFX10008, and QFX10016 switches running Enhanced Layer 2 Software, the member ID of a member in a Virtual Chassis determines the FPC number.

    NOTE: Every member in a Virtual Chassis must have a unique member ID, otherwise the Virtual Chassis will not be created.

- On standalone QFX5100, EX4600, and QFX10002 switches, the FPC number is always 0.

- *pic*—QFX Series and EX4600 device interfaces use the following convention for the PIC (*Physical Interface Card*) number in interface names:

Table 7: Naming Conventions for PICs

Device with Software Package	Convention
QFX3500 switch with QFabric software package	PIC 0 can support 48 ports, PIC 1 can support 16 10-Gigabit Ethernet ports, and PIC 2 can support 4 40-Gigabit Ethernet ports.
QFX3500 switch with Enhanced Layer 2 software	PIC **0** can support 48 ports, and PIC **1** can support 16 10-Gigabit Ethernet ports, and 4 40-Gigabit Ethernet ports.
QFX3500 Node device with a QFabric software package	PIC **0** can support 48 ports and PIC **1** can support four 40-Gigabit data plane uplink ports.
QFX3600 switch with a QFabric software package	PIC **0** can support 64 10-Gigabit Ethernet ports, and PIC **1** can support 16 40-Gigabit Ethernet ports.
QFX3600 switch with Enhanced Layer 2 software	PIC **0** can support 64 10-Gigabit Ethernet ports and can also support 16 40-Gigabit Ethernet ports.
QFX3600 Node device running a QFabric software package	PIC **0** can support 56 10-Gigabit Ethernet ports, and PIC **1** can support 8 40-Gigabit data plane uplink ports, and up to 14 40-Gigabit Ethernet ports.
QFX5100-48S switch with Enhanced Layer 2 software	PIC **0** provides six 40-Gbps QSFP+ ports and 48 10-Gigabit Ethernet interfaces.
EX4600 device with Enhanced Layer 2 software	PIC 0 provides 4 40-Gbps QSFP+ ports and 24 10-Gigabit Ethernet interfaces. There are two expansion bays (PIC 1 and PIC 2), and you can insert QFX-EM-4Q expansion modules and EX4600-EM-8F expansion modules. The QFX-EM-4Q expansion module provide 4 40-Gbps QSFP+ ports. The EX4600-EM-8F expansion module provides 8 10-Gbps SFP+ ports. You can insert any combination of expansion modules. For example, you can insert two EX4600-EM-8F expansion modules, two QFX-EM-4Q expansion modules, or one of each.

Table 7: Naming Conventions for PICs *(Continued)*

Device with Software Package	Convention
QFX5100-48S switch with a QFabric software package	PIC **1** provides six 40-Gbps QSFP+ ports, and PIC **0** provides 48 10-Gigabit Ethernet interfaces.
QFX5100-24Q switch with Enhanced Layer 2 software	PIC **0** provides 24 40-Gbps QSFP+ ports. PIC 1 and PIC 2 can each contain a QFX-EM-4Q expansion module, and each expansion module provides 4 40-Gbps QSFP+ ports
QFX5100-96S switch with Enhanced Layer 2 software	PIC **0** provides 96 10-Gigabit Ethernet interfaces and 8 40-Gbps QSFP+ ports .
QFX5110-48S switch with Enhanced Layer 2 software	PIC **0** can support 48 10-Gigabit Ethernet ports labeled 0 through 47, and 4 QSFP28 ports labeled 48 through 51. Ports 0 through 47 support either 1-Gbps small form-factor pluggable (SFP) or 10-Gbps small form-factor pluggable plus (SFP+) transceivers. You can also use SFP+ DAC cables and 10-Gbps active optical cables (AOC) in any access port. The default 100-Gigabit Ethernet ports can be configured as 40-Gigabit Ethernet, and in this configuration can either operate as dedicated 40-Gigabit Ethernet ports or can be channelized to 4 independent 10-Gigabit Ethernet ports using copper or fiber breakout cables.
QFX5200-32C switch with Enhanced Layer 2 software	PIC **0** provides 32 QSFP28 ports. The 100-Gigabit Ethernet ports can be channelized to two 50-Gigabit Ethernet or four 25-Gigabit Ethernet ports. The default 100-Gigabit Ethernet ports can be configured as 40-Gigabit Ethernet and operate as 40-Gigabit Ethernet or be channelized to four 10-Gigabit Ethernet ports.
QFX10002-36Q switch with Enhanced Layer 2 software	PIC **0** provides 144 10-Gigabit Ethernet interfaces, and 36 40-Gbps QSFP+ ports, and 12 100-Gigabit Ethernet interfaces.
QFX10002-72Q switch with Enhanced Layer 2 software	PIC **0** provides 288 10-Gigabit Ethernet interfaces, and 72 40-Gbps QSFP+ ports, and 24 100-Gigabit Ethernet interfaces.
QFX10008 switch with Enhanced Layer 2 software	PIC **0** provides one-thousand, one-hundred fifty two 10-Gigabit Ethernet interfaces, two-hundred eighty-eight 40-Gbps QSFP+ ports, or two-hundred forty 100-Gigabit Ethernet interfaces.

Table 7: Naming Conventions for PICs *(Continued)*

Device with Software Package	Convention
QFX10016 switch with Enhanced Layer 2 software	PIC **0** provides two-thousand, three-hundred and four 10-Gigabit Ethernet interfaces, five-hundred seventy-six 40-Gbps QSFP+ ports, or four-hundred eighty 100-Gigabit Ethernet interfaces.

- *port*—Interfaces use the following convention for port numbers:

Table 8: Naming Conventions for PORTs

Device with Software Package	Convention
QFX3500 switch with a QFabric software package	There are 48 network access ports (10-Gigabit Ethernet) labeled 0 through 47 on PIC 0 and, 16 network access ports labeled 0 through 15 on PIC 1, and four 40-Gbps QSFP+ ports labeled Q0 through Q3 on PIC 2. You can use the QSFP+ ports to connect the Node device to Interconnect devices. By default, the 40-Gbps QSFP+ ports are configured to operate as 10-Gigabit Ethernet ports. You can use QSFP+ to four SFP+ copper breakout cables to connect the 10-Gigabit Ethernet ports to other servers, storage, and switches. Optionally, you can choose to configure the QSFP+ ports as 40-Gigabit Ethernet ports (see Configuring the QSFP+ Port Type on QFX3500 Standalone Switches).
QFX3500 switch with Enhanced Layer 2 software	There are 48 network access ports labeled 0 through 47 on PIC 0 and 4 40-Gbps QSFP+ ports labeled Q0 through Q3 on PIC 1. See Channelizing Interfaces on QFX3500, QFX3600, QFX5100, QFX10002, QFX10008, QFX10016, and EX4600 Switches for information on how to configure and channelize the 40-Gbps QSFP+ ports.

Table 8: Naming Conventions for PORTs *(Continued)*

Device with Software Package	Convention
QFX3600 switch with a QFabric software package	There are 64 network access ports (10-Gigabit Ethernet) labeled Q0 through Q15 on PIC 0, and there are 16 network access ports (40-Gigabit Ethernet) labeled Q0 through Q15 on PIC **1**. By default, all the QSFP+ ports are configured to operate as 40-Gigabit Ethernet ports. Optionally, you can choose to configure the QSFP+ ports as 10-Gigabit Ethernet ports (see Configuring the Port Type on QFX3600 Standalone Switches) and use QSFP+ to four SFP+ copper breakout cables to connect the 10-Gigabit Ethernet ports to other servers, storage, and switches.
QFX3600 Node device with a QFabric software package	PIC **0** can support up to 56 10-Gigabit Ethernet ports labeled Q2 through Q15, and PIC **1** can support up to 8 40-Gigabit data plane uplink ports labeled Q0 through Q7, and up to 14 40-Gigabit Ethernet ports labeled Q2 through Q15. On a QFX3600 Node device, by default, four 40-Gbps QSFP+ ports (labeled Q0 through Q3) are configured for uplink connections between your Node device and your Interconnect devices, and twelve 40-Gbps QSFP+ ports (labeled Q4 through Q15) use QSFP+ to four SFP+ copper breakout cables to support up to 48 10-Gigabit Ethernet ports for connections to either endpoint systems (such as servers and storage devices) or external networks. Optionally, you can choose to configure the first eight ports (Q0 through Q7) for uplink connections between your Node device and your Interconnect devices, and ports Q2 through Q15 for 10-Gigabit Ethernet or 40-Gigabit Ethernet connections to either endpoint systems or external networks (see *Configuring the Port Type on QFX3600 Node Devices*).
QFX3600 switch with Enhanced Layer 2 software	PIC **0** can support 64 network access ports (10-Gigabit Ethernet ports) labeled Q0 through Q15 and 16 40-Gigabit Ethernet ports labeled Q0 through Q15. See Channelizing Interfaces on QFX3500, QFX3600, QFX5100, QFX10002, QFX10008, QFX10016, and EX4600 Switches for information on how to configure and channelize the 40-Gbps QSFP+ ports.
QFX5100-48S switch with Enhanced Layer 2 software	PIC **0** can support 48 network access ports (10-Gigabit Ethernet ports) labeled 0 through 47 and 6 40-Gbps QSFP+ ports labeled 48 through 53. See Channelizing Interfaces on QFX3500, QFX3600, QFX5100, QFX10002, QFX10008, QFX10016, and EX4600 Switches for information on how to configure and channelize the 40-Gbps QSFP+ ports.

Table 8: Naming Conventions for PORTs *(Continued)*

Device with Software Package	Convention
EX4600 switch with Enhanced Layer 2 software	PIC **0** can support 24 network access ports (10-Gigabit Ethernet ports) labeled 0 through 23 and 4 40-Gbps QSFP+ ports labeled 24 through 27. There are two expansion bays (PIC 1 and PIC 2), and you can insert QFX-EM-4Q expansion modules and EX4600-EM-8F expansion modules. The QFX-EM-4Q expansion module provide 4 40-Gbps QSFP+ ports. The EX4600-EM-8F expansion module provides 8 10-Gbps SFP+ ports. You can insert any combination of expansion modules. For example, you can insert two EX4600-EM-8F expansion modules, two QFX-EM-4Q expansion modules, or one of each. See Channelizing Interfaces on QFX3500, QFX3600, QFX5100, QFX10002, QFX10008, QFX10016, and EX4600 Switches for information on how to configure and channelize the 40-Gbps QSFP+ ports.
QFX5100-48S switch with a QFabric software package	PIC **0** can support 48 network access ports (10-Gigabit Ethernet ports) labeled 0 through 47, and PIC 1 can support 6 40-Gbps QSFP+ ports labeled 0 through 5. See *Configuring the QSFP+ Port Type on QFX5100 Devices* for information on how to configure the port mode of 40-Gbps QSFP+ ports.
QFX5100-24Q switch with Enhanced Layer 2 software	PIC **0** can support 24 40-Gbps QSFP+ ports labeled 0 through 23. PIC 1 and PIC 2 each support 4 40-Gbps QSFP+ port, for a total of eight 40-Gbps QSFP + ports. See Channelizing Interfaces on QFX3500, QFX3600, QFX5100, QFX10002, QFX10008, QFX10016, and EX4600 Switches for information on how to configure and channelize the 40-Gbps QSFP+ ports. **NOTE**: You cannot channelize the 40-Gbps QSFP+ ports provided in the two QFX-EM-4Q expansion modules. Also, even though there is a total of 128 physical ports, only 104 logical ports can be channelized. You can configure different system modes to achieve varying levels of port density on the QFX5100-24Q and QFX5100-96S switches. Depending on the system mode you configure, there are restrictions on which ports you can channelize. If you channelize ports that are restricted, the configuration is ignored. See Configuring the System Mode for information on how to configure the system mode.

Table 8: Naming Conventions for PORTs *(Continued)*

Device with Software Package	Convention
QFX5100-96S switch with Enhanced Layer 2 software	PIC **0** can support 96 10-Gigabit Ethernet ports labeled 0 through 95, and 8 40-Gbps QSFP+ ports labeled 96 through 103. See Channelizing Interfaces on QFX3500, QFX3600, QFX5100, QFX10002, QFX10008, QFX10016, and EX4600 Switches for information on how to configure and channelize the 40-Gbps QSFP+ ports. **NOTE**: You can only channelize the 40-Gbps QSFP+ ports provided in ports 96 and 100, because only 104 logical ports can be channelized. You can configure different system modes to achieve varying levels of port density on the QFX5100-24Q and QFX5100-96S switches. Depending on the system mode you configure, there are restrictions on which ports you can channelize. If you channelize ports that are restricted, the configuration is ignored. See Configuring the System Mode for information on how to configure the system mode.
QFX5110-48S switch with Enhanced Layer 2 software	PIC **0** can support 48 10-Gigabit Ethernet ports labeled 0 through 47, and 4 QSFP28 ports labeled 48 through 51. These data ports (0 through 47) support either 1-Gbps small form-factor pluggable (SFP) or 10-Gbps small form-factor pluggable plus (SFP+) transceivers. You can also use SFP+ DAC cables and 10-Gbps active optical cables (AOC) in any access port.The default 100-Gigabit Ethernet ports can be configured as 40-Gigabit Ethernet, and in this configuration can either operate as dedicated 40-Gigabit Ethernet ports or can be channelized to 4 independent 10-Gigabit Ethernet ports using copper or fiber breakout cables.
QFX5200-32C switch with Enhanced Layer 2 software	There is support for both quad small-form-factor pluggable (QSFP+) and 28-Gbps QSFP+ (QSFP28) transceivers in the 32 QSFP28 sockets. The QSFP28 ports are configured as 100-Gigabit Ethernet ports by default, but can also be configured to speeds of 50, 40, 25, or 10 Gigabit Ethernet. The 100 Gigabit Ethernet ports can be channelized using breakout cables either to 2 independent downstream 50 Gigabit Ethernet or to 4 independent 25 Gigabit Ethernet ports. The default 100 Gigabit Ethernet ports can also be configured as 40 Gigabit Ethernet and in this configuration can either operate as dedicated 40 Gigabit Ethernet ports or can be channelized to 4 independent 10 Gigabit Ethernet ports using breakout cables. See Channelizing Interfaces on QFX5200-32C Switches for information on how to configure and channelize the interfaces.

Table 8: Naming Conventions for PORTs *(Continued)*

Device with Software Package	Convention
QFX10002-36Q switch with Enhanced Layer 2 software	There are 36 quad small-form factor pluggable plus (QSFP+) ports that support 40-Gigabit Ethernet optical transceivers. Out of these 36 ports, 12 ports are QSFP28 capable, which are dual speed 40- or 100-Gigabit Ethernet optical transceivers. Each QSFP28 socket can be configured to support: • 100-Gigabit Ethernet using 28-Gbps QSFP28 optical transceivers. When a QSFP28 transceiver is inserted into the ports marked with a fine black line underneath the socket and the port is configured for 100-Gigabit Ethernet, the two adjacent ports are disabled and the QSFP28 is enabled for 100-Gigabit Ethernet. • 40-Gigabit Ethernet using QSFP+ optical transceivers. • 10-Gigabit Ethernet using breakout cables. When configured for channelization, a breakout cable converts the 40-Gigabit Ethernet port into 4 independent 10-Gigabit Ethernet ports. Any of the 36 ports 0 through 35 can be configured as either uplink or access ports. See Channelizing Interfaces on QFX3500, QFX3600, QFX5100, QFX10002, QFX10008, QFX10016, and EX4600 Switches for information on how to configure and channelize the 40-Gbps QSFP+ ports. Each of the 12 QSFP28 ports support: • 100-Gigabit Ethernet QSFP28 transceivers • 40-Gigabit Ethernet QSFP+ transceivers Each of the 36 QSFP+ ports support: • 40-Gigabit Ethernet QSFP+ transceivers • Access ports

Table 8: Naming Conventions for PORTs *(Continued)*

Device with Software Package	Convention
QFX10002-72Q switch with Enhanced Layer 2 software	There are 72 quad small-form factor pluggable plus (QSFP+) ports that support 40-Gigabit Ethernet optical transceivers. Out of these 72 ports, 24 ports are QSFP28 capable, which are dual speed 40- or 100-Gigabit Ethernet optical transceivers. Each QSFP28 socket can be configured to support: • 100-Gigabit Ethernet using 28-Gbps QSFP28 optical transceivers. When a QSFP28 transceiver is inserted into the ports marked with a fine black line underneath the socket and the port is configured for 100-Gigabit Ethernet, the two adjacent ports are disabled and the QSFP28 is enabled for 100-Gigabit Ethernet. • 40-Gigabit Ethernet using QSFP+ optical transceivers. • 10-Gigabit Ethernet using breakout cables. When configured for channelization, a breakout cable converts the 40-Gigabit Ethernet port into 4 independent 10-Gigabit Ethernet ports. Any of the 72 ports 0 through 71 can be configured as either uplink or access ports. See Channelizing Interfaces on QFX3500, QFX3600, QFX5100, QFX10002, QFX10008, QFX10016, and EX4600 Switches for information on how to configure and channelize the 40-Gbps QSFP+ ports. Each of the 24 QSFP28 ports support: • 100-Gigabit Ethernet QSFP28 transceivers Each of the 72 QSFP+ ports support: • 40-Gigabit Ethernet QSFP+ transceivers Each of the 36 QSFP+ ports support: • 40-Gigabit Ethernet QSFP+ transceivers • Access ports • Uplink ports

Table 8: Naming Conventions for PORTs *(Continued)*

Device with Software Package	Convention
On a QFX10008 switch with Enhanced Layer 2 software, there are two line cards available: QFX10008 with Line Card QFX10000-36Q (ELS)	QFX10000-36Q, a 36-port 40-Gigabit Ethernet quad small form-factor pluggable plus transceiver (QSFP+) or 12-port 100GbE QSFP28 line card The QFX10000-36Q line cards supports Each QSFP28 socket can be configured to support: • 100-Gigabit Ethernet using QSFP28 optical transceivers. When a QSFP28 transceiver is inserted into the ports marked with a fine black line underneath the socket and the port is configured for 100-Gigabit Ethernet, the two adjacent ports are disabled and the QSFP28 socket is enabled for 100-Gigabit Ethernet. • 40-Gigabit Ethernet using QSFP+ optical transceivers. • 10-Gigabit Ethernet using breakout cabling and attached optical transceivers. When configured for channelization, the system converts the 40-Gigabit Ethernet port into 4 independent 10-Gigabit Ethernet ports. Any of the 36 ports 0 through 35 can be configured as either uplink or access ports. See Channelizing Interfaces on QFX3500, QFX3600, QFX5100, QFX10002, QFX10008, QFX10016, and EX4600 Switches for information on how to configure and channelize the 40-Gbps QSFP+ ports. Each of the 12 QSFP28 ports supports: • 100-Gigabit Ethernet QSFP28 transceivers • 40-Gigabit Ethernet QSFP+ transceivers Each of the 12 QSFP28 ports supports: • 100-Gigabit Ethernet QSFP28 transceivers • 40-Gigabit Ethernet QSFP+ transceivers Each of the 36 QSFP+ ports support: • 40-Gigabit Ethernet QSFP+ transceivers

Table 8: Naming Conventions for PORTs *(Continued)*

Device with Software Package	Convention
	• Access ports • Uplink ports
QFX10008 with Line Card QFX10000-30C and QFX10000-30C-M (ELS)	QFX10000-30C and QFX10000-30C-M, a 30-port 100-Gigabit or 40-Gigabit Ethernet QSFP28 line card • The QFX10000-30C and QFX10000-30C-M line cards support: Thirty 28-Gbps QSFP+ Pluggable Solution (QSFP28) cages that support either 40-Gigabit Ethernet or 100-Gigabit Ethernet optical transceivers. The QFX10000-30C and QFX10000-30C-M ports auto detect the type of transceiver installed and set the configuration to the appropriate speed. Each QSFP28 socket can be configured to support: • 100-Gigabit Ethernet using QSFP28 optical transceivers. When a QSFP28 transceiver is inserted into the ports marked with a fine black line underneath the socket and the port is configured for 100-Gigabit Ethernet, the two adjacent ports are disabled and the QSFP28 socket is enabled for 100-Gigabit Ethernet. • 40-Gigabit Ethernet using QSFP+ optical transceivers. See Channelizing Interfaces on QFX3500, QFX3600, QFX5100, QFX10002, QFX10008, QFX10016, and EX4600 Switches for information on how to configure and channelize the 40-Gbps QSFP+ ports. Each of the 30 QSFP28 ports supports: • 100-Gigabit Ethernet QSFP28 transceivers • 40-Gigabit Ethernet QSFP+ transceivers • Access ports • Uplink ports

Table 8: Naming Conventions for PORTs *(Continued)*

Device with Software Package	Convention
On a QFX10016 switch running Enhanced Layer 2 software, there are 16 slots, which you can populate with two types line cards: QFX10016 with Line Card QFX10000-36Q (ELS)	• QFX10000-36Q, a 36-port 40-Gigabit Ethernet quad small form-factor pluggable plus transceiver (QSFP+) or 12-port 100GbE QSFP28 line card The QFX10000-36Q line card consists of 36 quad small form-factor pluggable plus (QSFP+) ports that support 40-Gigabit Ethernet optical transceivers. Out of these 36 ports, 12 ports are QSFP28 capable. The QSFP+ ports are dual speed and can support either 40-Gigabit or 100-Gigabit Ethernet optical transceivers. The line card can support 10-Gigabit Ethernet by channelizing the 40-Gigabit ports. Channelization is supported on fiber break-out cable using standard structured cabling techniques. With 100-Gigabit Ethernet using QSFP28 optical transceivers, when a QSFP28 transceiver is inserted into the ports marked with a fine black line underneath the socket and the port is configured for 100-Gigabit Ethernet, the two adjacent ports are disabled and the QSFP28 socket is enabled for 100-Gigabit Ethernet. You can use 40-Gigabit Ethernet using QSFP+ optical transceivers. With 10-Gigabit Ethernet using breakout cabling and attached optical transceivers, when configured for channelization, the system converts the 40-Gigabit Ethernet port into 4 independent 10-Gigabit Ethernet ports. Any of the 36 ports 0 through 35 can be configured as either uplink or access ports. Each of the 12 QSFP28 ports supports: • 100-Gigabit Ethernet QSFP28 transceivers • 40-Gigabit Ethernet QSFP+ transceivers Each of the 36 QSFP+ ports supports: • 40-Gigabit Ethernet QSFP+ transceivers • Access ports You can use 40-Gigabit Ethernet QSFP+ transceivers in any downstream port.

Table 8: Naming Conventions for PORTs *(Continued)*

Device with Software Package	Convention
	• Uplink ports You can configure all the QSFP+ ports as uplinks. Every second and sixth port in a 6XQSFP cage on a QFX10000-36Q supports 100-Gigabit Ethernet using QSFP28 transceivers. These 100-Gigabit Ethernet ports work either as 100-Gigabit Ethernet or as 40-Gigabit Ethernet, but are recognized as 40-Gigabit Ethernet by default. When a 40-Gigabit Ethernet transceiver is inserted into a 100-Gigabit Ethernet port, the port recognizes the 40-Gigabit Ethernet port speed. When a 100-Gigabit Ethernet transceiver is inserted into the port and enabled in the CLI, the port recognizes the 100-Gigabit Ethernet speed and disables two adjacent 40-Gigabit Ethernet ports. You can also use an 100-Gigabit Ethernet transceiver and run it at 40-Gigabit Ethernet by using the CLI to set the port speed to 40-Gigabit Ethernet. The 40-Gigabit Ethernet ports can operate independently, be channelized into four 10-Gigabit Ethernet ports, or bundled with the next two consecutive ports and channelized into twelve 10-Gigabit Ethernet ports as a port range. Only the first and fourth port in each 6XQSFP cage are available to channelize a port range. The port range must be configured using the set chassis fpc pic port channel-speed command. For example, to channelize the first switch port, use the `set chassis fpc 0 pic 0port 1channel-speed 10g` command.

Table 8: Naming Conventions for PORTs *(Continued)*

Device with Software Package	Convention
QFX10016 with Line Card QFX10000-30C and QFX10000-30C-M (ELS)	QFX10000-30C and QFX10000-30C-M, a 30-port 100-Gigabit or 40-Gigabit Ethernet QSFP28 line card. The QFX10000-30C and QFX10000-30C-M ports auto detect the type of transceiver installed and set the configuration to the appropriate speed. Each QSFP28 socket supports: • 100-Gigabit Ethernet using QSFP28 optical transceivers. When a QSFP28 transceiver is inserted into any of the ports, the QSFP28 socket is enabled for 100-Gigabit Ethernet. • 40-Gigabit Ethernet using QSFP+ optical transceivers. When a QSFP+ transceiver is inserted into any of the ports, the QSFP+ socket is enabled for 40-Gigabit. Any of the 30 ports 0 through 29 can be configured as either uplink or access ports, and of the 30 QSFP28 ports supports: • 100-Gigabit Ethernet QSFP28 transceivers • 40-Gigabit Ethernet QSFP+ transceivers

Logical Part of an Interface Name on a Switch Running QFabric Software Package for QFX series, NFX Series, EX4600, QFabric System

The logical unit part of the interface name corresponds to the logical unit number, which can be a number from 0 through 16384. In the virtual part of the name, a period (.) separates the port and logical unit numbers: *device-name* (QFabric systems only): *type-fpc/pic/port.logical-unit-number*. For example, if you issue the **show ethernet-switching interfaces** command on a system with a default VLAN, the resulting display shows the logical interfaces associated with the VLAN:

```
Interface               State   VLAN members          Blocking
node-device1:xe-0/0/1.0  down     remote-analyzer       unblocked
node-device1:xe-0/0/2.0  down     default               unblocked
node-device1:xe-0/0/3.0  down     default               unblocked
```

When you configure aggregated Ethernet interfaces, you configure a *logical interface*, which is called a or a . Each LAG can include up to eight Ethernet interfaces, depending on the switch model.

Logical Part of a Channelized Interface Name on a Switch Running Enhanced Layer 2 Software for QFX series, NFX Series, EX4600, QFabric System

Channelizing enables you to configure four 10-Gigabit Ethernet interfaces from a 40-Gigabit Ethernet QSFP+ interface. By default, a 40-Gigabit Ethernet QSFP+ interface is named *et-fpc/pic/port*. The resulting 10-Gigabit Ethernet interfaces appear in the following format: *xe-fpc/pic/port:channel*, where channel can be a value of 0 through 3.

For example, if an *et* interface named `et-0/0/3` is channelized to four 10-Gigabit Ethernet interfaces, the resulting 10-Gigabit Ethernet interface names will be `xe-0/0/3:0`, `xe-0/0/3:1`, `xe-0/0/3:2`, and `xe-0/0/3:3`:

```
Interface               Admin Link Proto    Local        Remote
xe-0/0/3:0               up    down
xe-0/0/3:1               up    down
xe-0/0/3:2               up    down
xe-0/0/3:3               up    down
```

Wildcard Characters in Interface Names for QFX series, NFX Series, EX4600, QFabric System

In the **show interfaces** and **clear interfaces** commands, you can use wildcard characters in the ***interface-name*** option to specify groups of interface names without having to type each name individually. You must enclose all wildcard characters except the asterisk (*) in square brackets [ ].

Physical Part of an Interface Name for OCX1100

Interfaces in Junos OS are specified as follows:

type-fpc/pic/port

The convention is as follows:

- *type*—The OCX Series device interfaces use the following media types:
  - **xe**—10-Gigabit Ethernet interface
  - **et**—40-Gigabit Ethernet interface
  - **em**—Management interface
- *fpc*—Flexible PIC Concentrator. OCX Series interfaces use the following convention for the FPC number in interface names:

- On standalone OCX Series switches, the FPC number is always **0**.

  The FPC number indicates the slot number of the line card that contains the physical interface.

- *pic*—The OCX Series interfaces use the following convention for the PIC (*Physical Interface Card*) number in interface names:
  - PIC **0** provides six 40-Gbps QSFP+ ports and 48 10-Gigabit Ethernet interfaces.
- *port*—Interfaces use the following convention for port numbers:
  - PIC **0** can support 48 network access ports (10-Gigabit Ethernet ports) labeled 1 through 48 and 6 40-Gbps QSFP+ ports labeled 49 through 54.

Wildcard Characters in Interface Names for OCX1100

In the **show interfaces** and **clear interfaces** commands, you can use wildcard characters in the ***interface-name*** option to specify groups of interface names without having to type each name individually. You must enclose all wildcard characters except the asterisk (*) in square brackets [ ].

SEE ALSO

Interfaces Overview for Switches | 2

Channelizing Interfaces on QFX3500, QFX3600, QFX5100, QFX10002, QFX10008, QFX10016, and EX4600 Switches

Understanding Management Interfaces | 30

Understanding Port Ranges and System Modes

Configuring the System Mode

Configuring Gigabit Ethernet Interfaces (CLI Procedure)

Configuring Gigabit Ethernet Interfaces for EX Series Switches with ELS support | 111

Junos OS Network Interfaces Library for Routing Devices

Rear Panel of a QFX3500 Device

Front Panel of a QFX3600 Device

Understanding Management Interfaces

You use management interfaces to access devices remotely. Typically, a management interface is not connected to the in-band network, but is connected to a device in the internal network. Through a management interface, you can access the device over the network using utilities such as **ssh** and **telnet**

and configure it from anywhere, regardless of its physical location. As a security feature, users cannot log in as **root** through a management interface. To access the device as **root**, you must use the console port. You can also use **root** to log in using SSH.



NOTE: Before you can use management interfaces, you must configure the logical interfaces with valid IP addresses. Juniper Networks does not support configuring two management interfaces in the same subnet.

Management interface port ranges vary based on device type (and platform support depends on the Junos OS release in your installation):

- QFX3500 devices:

  The valid port range for a management interface (**me**) on a QFX3500 device is between 0 and 6, with a total of seven available ports. On a QFX3500 standalone switch, however, you can only configure me0 and me1 as management interfaces. The management interfaces are labeled **C0** and **C1**, and they correspond to me0 and me1. On a QFX3500 Node device, the RJ-45 management interfaces and SFP management interfaces correspond to **me5** and **me6**

- QFX3600 devices:

  There are two RJ-45 management interfaces (labeled **C0** and **C1**) and two SFP management interfaces (labeled **C0S** and **C1S**). On a QFX3600 standalone switch, the RJ-45 management interfaces and SFP management interfaces correspond to **me0** and **me1**. On a QFX3600 Node device, the RJ-45 management interfaces and SFP management interfaces correspond to **me5** and **me6**. Each pair of management interfaces correspond to one Ethernet interface—for example, both RJ-45 management interfaces (labeled **C0** and **C0s**) can correspond to me0, and both SFP management interfaces (labeled **C1** and **C1S**) can correspond to me1. By default, both RJ-45 management interfaces are active. If you insert an SFP interface into the SFP management port (**C0S**, for example), the SFP interface would become the active management interface, and the corresponding RJ-45 management interface (**C0**) is disabled.

  NOTE: On a QFX3600 device, you can use either the RJ-45 or the SFP management interfaces, but not both at the same time.

- On QFX5100, QFX5200, and EX4600 switches, there is one RJ-45 management interface (labeled **C0** and one SFP management interface (labeled **C1**), and they correspond to em0 and em1. You can use both management interfaces simultaneously.

- On QFX10002 and QFX10008 switches, there is one RJ-45 management interface (labeled **MGMT** and one SFP management interface (labeled **MGMT**), and they correspond to em0 and em1. Although the CLI permits you to configure two management Ethernet interfaces within the same subnet, only one interface is usable and supported.

- On QFX10008 and QFX10016 switches, if you are using em1 for management purpose, then you cannot directly access the backup RE em1 from external network. Indirectly you can access the backup RE from external network, by following these steps:
  - Login to primary RE using SSH/Telnet to its em1.
  - Access backup RE using the following command:

```
user@host>request routing-engine login other-routing-engine
```

- On OCX Series switches:

  There is one RJ-45 management interface (labeled **MGMT**), which corresponds to em0. The em0 interface always has the status `up` in show command outputs, even if the physical port is empty. The me0 interface is a virtual interface between Junos and the host operating system, therefore its status is independent from the status of the physical port.

- QFabric system:

  On a QFabric system, there are management interfaces on the Node devices, Interconnect devices, and Director devices. However, you cannot access the management interfaces on the Node devices or Interconnect devices directly. You can only manage and configure these devices using the Director device. You can connect to the management interface over the network using utilities such as SSH.

  For information on how to use management interfaces on a QFabric system, see *Performing the QFabric System Initial Setup on a QFX3100 Director Group* and *Gaining Access to the QFabric System Through the Default Partition*.

Physical Interface Properties

IN THIS SECTION

- Configure Damping of Shorter Physical Interface Transitions | **33**
- Accounting for Physical Interfaces | **34**
- Enable SNMP Notifications on Physical Interfaces | **38**
- Configuring Ethernet Loopback Capability | **39**
- Configuring Short Reach Mode on QFX5100-48T | **40**

- Configuring Flow Control | **41**
- Setting the Mode on an SFP+ or SFP+ MACSec Uplink Module | **42**
- Setting the Operating Mode on a 2-Port 40-Gigabit Ethernet QSFP+/100-Gigabit Ethernet QSFP28 Uplink Module | **43**
- Configuring the Media Type on Dual-Purpose Uplink Ports | **45**
- Disable a Physical Interface | **46**

The physical interfaces undergo various transitions which is advertised to the Junos OS for proper functioning of the routers and switches. Accounting profiles that specify the characteristics of data about the traffic passing through the routers and switches can also be configured on the physical interfaces. Simple Network Management Protocol (SNMP) notifications can be enabled on the physical interface to provide information about the state of an interface or when a connection changes. The interface offers to configure various modes like short-reach-mode, flow-control and media type on the devices for ease of access.

Configure Damping of Shorter Physical Interface Transitions

By default, when an interface changes from up to down or from down to up, this transition is advertised immediately to the hardware and Junos OS. In some situations, you might want to damp interface transitions.

For example, you may want to configure damping on an interface that is connected to an add/drop multiplexer (ADM) or wavelength-division multiplexer (WDM), or to protect against SONET/SDH framer holes.

Damping the interface means not advertising the interface's transition until a certain period of time has passed, called the *hold-time*. When the interface goes from up to down, the down hold-time timer is triggered. Every interface transition that occurs during the hold time is ignored. If the timer expires and the interface state is still *down*, then the router begins to advertise the interface as being down. Similarly, when an interface goes from down to up, the up hold-time timer is triggered. Every interface transition that occurs during the hold time is ignored. If the timer expires and the interface state is still *up*, then the router begins to advertise the interface as being up.

To configure damping of shorter physical interface transitions in milliseconds:

1. Select the interface to damp, where the interface name is *interface-type-fpc/pic/port*:

```
[edit]
user@host# edit interfaces interface-name
```

2. Configure the hold time for link up and link down.

```
[edit interfaces interface-name]
user@host# set hold-time up milliseconds down milliseconds
```

The hold time can be a value from 0 through 4,294,967,295 milliseconds. The default value is 0, which means that interface transitions are not damped. Junos OS advertises the transition within 100 milliseconds of the time value you specify.

For most Ethernet interfaces, Junos OS implements hold timers using a one-second polling algorithm. For 1-port, 2-port, and 4-port Gigabit Ethernet interfaces with small form-factor pluggable (SFP) transceivers, hold timers are interrupt driven.

NOTE: The `hold-time` option is not available for controller interfaces.

Accounting for Physical Interfaces

IN THIS SECTION

- Overview | 35
- Configure an Accounting Profile for a Physical Interface | 35
- How to Display the Accounting Profile | 37

Devices running Junos OS can collect various kinds of data about traffic passing through the device. You (the systems administrator) can set up one or more *accounting profiles* that specify some common characteristics of this data. These characteristics include the following:

- The fields used in the accounting records

- The number of files that the router or switch retains before discarding, and the number of bytes per file
- The polling period that the system uses to record the data

Overview

There are two types of accounting profiles: filter profiles and interface profiles. Configure the profiles using statements at the `[edit accounting-options]` hierarchy level.

Configure filter profiles by including the `filter-profile` statement at the `[edit accounting-options]` hierarchy level. You apply filter profiles by including the `accounting-profile` statement at the `[edit firewall filter` *`filter-name`*`]` and `[edit firewall family` *`family`* `filter` *`filter-name`*`]` hierarchy levels.

Configure interface profiles by including the `interface-profile` statement at the `[edit accounting-options]` hierarchy level. Read on to learn how to configure interface profiles.

Configure an Accounting Profile for a Physical Interface

Before You Begin

Configure an accounting data log file at the `[edit accounting-options]` hierarchy level. The operating system logs the statistics in the accounting data log file.

For more information about how to configure an accounting data log file, see the *Configuring Accounting-Data Log Files*.

Configuration

Configure an interface profile to collect error and statistic information for input and output packets on a particular physical interface. The interface profile specifies the information that the operating system writes to the log file.

To configure an interface profile:

1. Navigate to the `[edit accounting-options interface-profile]` hierarchy level. Include the *profile-name* to name the interface profile.

```
[edit]
user@host# edit accounting-options interface-profile profile-name
```

2. To configure which statistics should be collected for an interface, include the `fields` statement.

```
[edit accounting-options interface-profile profile-name]
user@host# set fields field-name
```

3. Each accounting profile logs its statistics to a file in the **/var/log** directory. To configure which file to use, use the `file` statement.

```
[edit accounting-options interface-profile profile-name]
user@host# set file filename
```

> **NOTE**: You must specify a `file` statement for the interface profile that has already been configured at the `[edit accounting-options]` hierarchy level.

4. The operating system collects statistics from each interface with an accounting profile enabled. It collects the statistics once per interval time specified for the accounting profile. The operating system schedules statistics collection time evenly over the configured interval. To configure the interval, use the `interval` statement:

```
[edit accounting-options interface-profile profile-name]
user@host# set interval minutes
```

> **NOTE**: The minimum interval allowed is 1 minute. Configuring a low interval in an accounting profile for a large number of interfaces might cause serious performance degradation.

5. Apply the interface profile to a physical interface by including the `accounting-profile` statement at the `[edit interfaces interface-name]` hierarchy level. The operating system performs the accounting on the interfaces that you specify.

```
[edit interfaces]
user@host# set interface-name accounting-profile profile-name
```

SEE ALSO

Configuring Accounting-Data Log Files

How to Display the Accounting Profile

IN THIS SECTION

- Purpose | 37
- Action | 37
- Meaning | 38

Purpose

To display the configured accounting profile of a particular physical interface at the `[edit accounting-options interface-profile profile-name]` hierarchy level that has been configured with the following:

- interface-name—et-1/0/1
- Interface profile —`if_profile`
- File name—`if_stats`
- Interval—15 minutes

Action

- Run the `show` command at the `[edit interfaces et-1/0/1]` hierarchy level.

```
[edit interfaces et-1/0/1]
user@host# show
accounting-profile if_profile;
```

- Run the `show` command at the `[edit accounting-options]` hierarchy level.

```
[edit accounting-options]
user@host# show
interface-profile if_profile {
    interval 15;
    file if_stats {
        fields {
            input-bytes;
            output-bytes;
```

```
            input-packets;
            output-packets;
            input-errors;
            output-errors;
        }
    }
}
```

Meaning

The configured accounting and its associated set options are displayed as expected.

Enable SNMP Notifications on Physical Interfaces

By default, Junos OS sends Simple Network Management Protocol (SNMP) notifications when the state of an interface or a connection changes. You can enable or disable SNMP notifications based on your requirements.

To explicitly enable sending SNMP notifications on the physical interface:

1. In configuration mode, go to the `[edit interfaces interface-name]` hierarchy level:

```
[edit]
user@host# edit interfaces interface-name
```

2. Configure the `traps` option to enable SNMP notifications when the state of the connection changes.

```
[edit interfaces interface-name]
user@host# set traps
```

To disable SNMP notifications on the physical interface:

1. In configuration mode, go to the `[edit interfaces interface-name]` hierarchy level:

```
[edit]
user@host# edit interfaces interface-name
```

2. Configure the `no-traps` option to disable SNMP notifications when the state of the connection changes.

```
[edit interfaces interface-name]
user@host# set no-traps
```

Configuring Ethernet Loopback Capability

To place an interface in loopback mode, include the `loopback` statement:

```
loopback;
```

To return to the default—that is, to disable loopback mode—delete the `loopback` statement from the configuration:

```
[edit]
user@switch# delete interfaces interface-name ether-options loopback
```

To explicitly disable loopback mode, include the `no-loopback` statement:

```
no-loopback;
```

You can include the **loopback** and `no-loopback` statements at the following hierarchy levels:

- `[edit interfaces interface-name aggregated-ether-options]`
- `[edit interfaces interface-name ether-options]`

SEE ALSO

Configuring Gigabit and 10-Gigabit Ethernet Interfaces for EX4600 and QFX Series Switches | 105

Configuring Short Reach Mode on QFX5100-48T

You can enable short-reach mode for individual as well as a range of copper-based 10-Gigabit Ethernet interfaces using short cable lengths (less than 10m) on the QFX5100-48T switch. Short-reach mode reduces power consumption up to 5W on these interfaces.

1. To enable short-reach mode on an individual interface, issue the following command:

```
[edit chassis]
user@switch# set fpc fpc-slot pic pic-slot port port-number short-reach-mode enable
```

   For example, to enable short-reach mode on port 0 on PIC 0, issue the following command:

```
[edit chassis]
user@switch# set fpc 0 pic 0 port 0 short-reach-mode enable
```

2. To enable short-reach mode on a range of interfaces, issue the following command:

```
[edit chassis]
user@switch# set fpc fpc-slot pic pic-slot port-range port-range-low port-range-high short-
reach-mode enable
```

   For example, to enable short-reach mode on a range of interfaces between port 0 and port 47 on PIC 0, issue the following command:

```
[edit chassis]
user@switch# set fpc 0 pic 0 port-range 0 47 short-reach-mode enable
```

3. To disable short-reach mode on an individual interface, issue the following command:

```
[edit chassis]
user@switch# set fpc fpc-slot pic pic-slot port port-number short-reach-mode disable
```

   For example, to disable short-reach mode on port 0 on PIC 0, issue the following command:

```
[edit chassis]
user@switch# set fpc 0 pic 0 port 0 short-reach-mode disable
```

4. To disable short-reach mode on a range of interfaces, issue the following command:

```
[edit chassis]
user@switch# set fpc fpc-slot pic pic-slot port-range port-range-low port-range-high short-
reach-mode disable
```

For example, to disable short-reach mode on a range of interfaces between port 0 and port 47 on PIC 0, issue the following command:

```
[edit chassis]
user@switch# set fpc 0 pic 0 port-range 0 47 short-reach-mode disable
```

SEE ALSO

short-reach-mode

Configuring Flow Control

By default, the router or switch imposes flow control to regulate the amount of traffic sent out on a Fast Ethernet, Tri-Rate Ethernet copper, Gigabit Ethernet, and 10-Gigabit Ethernet interface. Flow control is not supported on the 4-port Fast Ethernet PIC. This is useful if the remote side of the connection is a Fast Ethernet or Gigabit Ethernet switch.

You can disable flow control if you want the router or switch to permit unrestricted traffic. To disable flow control, include the `no-flow-control` statement:

```
no-flow-control;
```

To explicitly reinstate flow control, include the `flow-control` statement:

```
flow-control;
```

You can include these statements at the following hierarchy levels:

- `[edit interfaces interface-name aggregated-ether-options]`
- `[edit interfaces interface-name ether-options]`

- [edit interfaces *interface-name* fastether-options]
- [edit interfaces *interface-name* gigether-options]

> **NOTE**: On the Type 5 FPC, to prioritize control packets in case of ingress oversubscription, you must ensure that the neighboring peers support MAC flow control. If the peers do not support MAC flow control, then you must disable flow control.

SEE ALSO

> *flow-control*
>
> *Ethernet Interfaces Overview*

Setting the Mode on an SFP+ or SFP+ MACSec Uplink Module

SFP+ uplink modules are supported on EX3200 and EX4200 switches, and SFP+ Media Access Control Security (MACSec) uplink modules are supported on EX4200 switches. You can use these uplink modules either for two SFP+ transceivers or four SFP transceivers. You configure the operating mode on the module to match the type of transceiver you want to use—that is, for SFP+ transceivers, you configure the 10-gigabit operating mode, and for SFP transceivers, you configure the 1-gigabit operating mode.

By default, the SFP+ uplink module operates in the 10-gigabit mode and supports only SFP+ transceivers. If you have not changed the module from the default setting and you want to use SFP+ transceivers, you do not need to configure the operating mode.

To set the operating mode of an SFP+ or SFP+ MACSec uplink module:

1. Change the operating mode to the appropriate mode for the transceiver type you want to use by using one of the following commands:

```
[edit]
user@switch# set chassis fpc 0 pic 1 sfpplus pic-mode 1g
```

```
[edit]
user@switch# set chassis fpc 0 pic 1 sfpplus pic-mode 10g
```

2. (SFP+ uplink module only) If the switch is running:

- Junos OS Release 10.1 or later, the changed operating mode takes effect immediately unless a port on the SFP+ uplink module is a Virtual Chassis port (VCP). If any port on the SFP+ uplink module is a VCP, the changed operating mode does not take effect until the next reboot of the switch.

  NOTE: During the operating mode change, the Packet Forwarding Engine is restarted. In a Virtual Chassis configuration, this means that the Flexible PIC Concentrator connection with the primary device is dropped and then reconnected.

- Junos OS Release 10.0 or earlier, reboot the switch.

You can see whether the operating mode has been changed to the new mode you configured by issuing the `show chassis pic fpc-slot slot-number pic-slot 1` command.

SEE ALSO

Uplink Modules in EX3200 Switches

Uplink Modules in EX4200 Switches

Pluggable Transceivers Supported on EX3200 Switches

Pluggable Transceivers Supported on EX4200 Switches

Setting the Operating Mode on a 2-Port 40-Gigabit Ethernet QSFP+/100-Gigabit Ethernet QSFP28 Uplink Module

You can configure the 2-port 4-Gigabit Ethernet QSFP+/100-Gigabit Ethernet QSFP28 uplink module on EX4300-48MP switches to operate either two 40-Gigabit Ethernet ports or two 100-Gigabit Ethernet port. By default, the uplink module operates only the two 40-Gbps ports.

The uplink module on EX4300-48MP switches supports Media Access Control Security (MACsec). See *Understanding Media Access Control Security (MACsec)* for more information.

The uplink module does not support configuring virtual chassis ports.

To set the operating mode on this uplink module:

1. Install the 2-port 40-Gigabit Ethernet QSFP+/100-Gigabit Ethernet QSFP28 uplink module only in PIC slot 2 on the switch. Insert the uplink module in the chassis and check whether it is detected by issuing the `show chassis hardware` command.

2. Change the operating mode to 100-Gigabit Ethernet mode, by issuing the following command on the first port (port 0). The port then recognizes the 100-Gigabit speed and disables the adjacent 40-Gigabit Ethernet port. The adjacent 40-Gigabit Ethernet port is disabled only when port 0 is loaded with 100G optics.

```
[edit]
user@switch# set chassis fpc 0 pic 2 port 0 speed 100G
```

3. You can change the operating mode to 100-Gigabit Ethernet mode on the second (port 1) by using the following command. This command overrides the `set chassis fpc 0 pic 2 port 0 speed 100G` command to change the operating mode to 100-Gigabit Ethernet mode.

```
[edit]
user@switch# run request chassis system-mode mode-2x100G
```

4. Optional: Check whether the operating mode has been changed to the new mode you configured by issuing the `show chassis pic fpc-slot 0 pic-slot 2` command.



NOTE: If you configure both the ports on the uplink module to operate at 100-Gbps speed, the four built-in QSFP+ ports on the switch are disabled.

Starting with Junos OS Release 19.1R1, in the 2-port 40-Gigabit Ethernet QSFP+/1-port 100-Gigabit Ethernet QSFP28 uplink module of EX4300-48MP switches, you can channelize the 100-Gigabit four independent 25-Gigabit Ethernet ports by using breakout cables. You can configure only port 0 of the uplink module as 25-Gigabit Ethernet port. Issue the command `set chassis fpc 0 pic 2 port 0 channel-speed 25g` to channelize the 100-Gigabit Ethernet uplink port to four 25-Gigabit Ethernet uplink ports.

Starting with Junos OS Release 19.3R1, you can configure the 2-port 40-Gigabit Ethernet QSFP+/100-Gigabit Ethernet QSFP28 uplink module on EX4300-48MP switches to operate either two 40-Gigabit Ethernet ports or two 100-Gigabit Ethernet ports.

You can also channelize the 40-Gigabit Ethernet interfaces to four independent 10-Gigabit Ethernet interfaces using breakout cables. To channelize the 100-Gigabit Ethernet interfaces to operate as four independent 25-Gigabit Ethernet, specify the port number and channel speed

1. To configure the 100-Gigabit Ethernet uplink port to operate as a 25-Gigabit Ethernet interface, specify the port number and channel speed by using the following command:

```
[edit chassis fpc 0 pic 2]
user@switch# set port port-number channel-speed speed
```

For example, to configure port 0 to operate as a 25-Gigabit Ethernet port:

```
[edit chassis fpc 0 pic 2]
user@switch# set port 0 channel-speed 25g
```

2. Review your configuration and issue the `commit` command.

```
[edit]
user@switch# commit
commit complete
```

NOTE: If you configure both the ports on the uplink module to operate at 100-Gbps speed, the four QSFP+ ports on the switch are disabled.

SEE ALSO

Uplink Modules in EX4300 Switches

Configuring the Media Type on Dual-Purpose Uplink Ports

EX2200-C switches and ACX1000 routers provide two dual-purpose uplink ports. Each dual uplink port is a single interface that offers a choice of two connections: an RJ-45 connection for a copper Ethernet cable and an SFP connection for a fiber-optic Ethernet cable. You can choose to use either connection, but only one connection can be active at a time.

By default, if you plug a transceiver into the SFP connector, the port becomes a fiber-optic Gigabit Ethernet port, even if a copper Ethernet cable is plugged into the RJ-45 connection as well. If a transceiver is not plugged into the SFP connector, the port defaults to a copper 10/100/1000 Ethernet port.

You can constrain the use of the port to one connection type by configuring the media type for the port to be either copper or fiber. When you configure a media type on the port, the port will no longer accept the alternate connection type. For example, if you configure the uplink port as a fiber port and then plug a copper Ethernet cable into the RJ-45 connector, the interface will not come up.

To configure the media type for an uplink port:

```
user@switch# set interfaces interface-name media-type (Dual-Purpose Uplink Ports) media-type
```

For example, to set the media type for uplink port **ge-0/1/0** to copper:

```
user@switch# set interfaces ge-0/1/0 media-type copper
```

NOTE: When you change the media type setting for a dual-purpose uplink port, it can take up to 6 seconds for the interface to appear in operational commands.

SEE ALSO

EX2200 Switches Hardware Overview

Disable a Physical Interface

IN THIS SECTION

- How to Disable a Physical Interface | 46
- Example: Disable a Physical Interface | 47

You can disable a physical interface, marking it as being down, without removing the interface configuration statements from the configuration.

How to Disable a Physical Interface

CAUTION: Dynamic subscribers and logical interfaces use physical interfaces for connection to the network. You can set the interface to disable and commit the change while dynamic subscribers and logical interfaces are still active. This action results in the loss of all subscriber connections on the interface. Use care when disabling interfaces.

To disable a physical interface:

1. In configuration mode, go to the `[edit interfaces interface-name]` hierarchy level.

```
[edit]
user@host# edit interfaces interface-name
```

2. Include the `disable` statement.

```
[edit interfaces interface-name]
user@device# set disable
```

For example:

```
[edit interfaces et-1/0/7]
user@device# set disable
```

NOTE: When you use the `disable` statement at the `edit interfaces` hierarchy level, depending on the PIC type, the interface might or might not turn off the laser. Older PIC transceivers do not support turning off the laser, but newer Gigabit Ethernet PICs with SFP and XFP transceivers do support it. On a device with newer PICs, the laser turns off when the interface is disabled.

LASER WARNING: Do not stare into the laser beam or view it directly with optical instruments even if the interface has been disabled.

Example: Disable a Physical Interface

Sample interface configuration:

```
[edit interfaces]
user@device# show et-0/3/2 {
    unit 0 {
        description CE2-to-PE1;
        family inet {
            address 20.1.1.6/24;
        }
```

```
    }
}
```

Disable the interface:

```
[edit interfaces et-0/3/2]
user@device# set disable
```

Verify the interface configuration:

```
[edit interfaces et-0/3/2]
user@device# show
disable; # Interface is marked as disabled.
    unit 0 {
    description CE2-to-PE1;
        family inet {
            address 20.1.1.6/24;
        }
}
```

Change History Table

Feature support is determined by the platform and release you are using. Use Feature Explorer to determine if a feature is supported on your platform.

Release	Description
19.3R1	Starting with Junos OS Release 19.3R1, you can configure the 2-port 40-Gigabit Ethernet QSFP+/100-Gigabit Ethernet QSFP28 uplink module on EX4300-48MP switches to operate either two 40-Gigabit Ethernet ports or two 100-Gigabit Ethernet ports.
19.1R1	Starting with Junos OS Release 19.1R1, in the 2-port 40-Gigabit Ethernet QSFP+/1-port 100-Gigabit Ethernet QSFP28 uplink module of EX4300-48MP switches, you can channelize the 100-Gigabit four independent 25-Gigabit Ethernet ports by using breakout cables.

Logical Interface Properties

IN THIS SECTION

- Assign the Interface Address | 49
- Add a Logical Unit Description to the Configuration | 51
- Media MTU Overview | 52
- Configure the Media MTU | 53
- Protocol MTU | 54
- Configure the Interface Bandwidth | 55
- Enable or Disable SNMP Notifications on Logical Interfaces | 56
- Overview of Accounting for the Logical Interface | 57
- Disable a Logical Interface | 60

The logical interfaces can be configured and the description is displayed in the output of the `show` commands. Media maximum transmission unit (MTU) is automatically calculated when configuring an interface and can also be modified. Simple Network Management Protocol (SNMP) notifications can be enabled on the logical interface to provide information about the state of an interface or when a connection changes.

Assign the Interface Address

You assign an address to an interface by specifying the address when configuring the protocol family. For the `inet` or `inet6` family, configure the interface IP address. For the `iso` family, configure one or more addresses for the loopback interface. For the `ccc`, `ethernet-switching`, `tcc`, `mpls`, `tnp`, and `vpls` families, you never configure an address.



NOTE: The Point-to-Point Protocol (PPP) address is taken from the loopback interface address that has the primary attribute. When the loopback interface is configured as an unnumbered interface, it takes the primary address from the donor interface.

To assign an address to an interface, perform the following steps:

1. Configure the interface address at the [edit interfaces *interface-name* unit *logical-unit-number* family *family*] hierarchy level.

   - To configure an IP version 4 (IPv4) address on routers and switches, use the interface *interface-name* unit *number* family inet address *a.b.c.d/nn* statement at the [edit interfaces] hierarchy level.

     You can also assign multiple IPv4 addresses on the same interface.

     ```
     [edit interfaces ]
     user@host# set interface-name unit logical-unit-number family inet address a.b.c.d/nn
     ```

     NOTE:

     - Juniper Networks routers and switches support /31 destination prefixes when used in point-to-point Ethernet configurations; however, they are not supported by many other devices, such as hosts, hubs, routers, or switches. You must determine if the peer system also supports /31 destination prefixes before configuration.
     - You can configure the same IPv4 address on multiple physical interfaces. When you assign the same IPv4 address to multiple physical interfaces, the operational behavior of those interfaces differs, depending on whether they are implicitly or explicitly point-to-point.
     - By default, all interfaces are assumed to be point-to-point (PPP) interfaces. For all interfaces except aggregated Ethernet, Fast Ethernet, and Gigabit Ethernet, you can explicitly configure an interface to be a point-to-point connection.
     - If you configure the same IP address on multiple interfaces in the same routing instance, Junos OS applies the configuration randomly on one of the interfaces. The other interfaces will remain without an IP address.

   - To configure an IP version 6 (IPv6) address on routers and switches, use the interface *interface-name* unit *number* family inet6 address *aaaa:bbbb:...:zzzz/nn* statement at the [edit interfaces] hierarchy level.

     ```
     [edit interfaces ]
     user@host# set interface-name unit logical-unit-number family inet6 address
     aaaa:bbbb:...:zzzz/nn
     ```

NOTE:

- You represent IPv6 addresses in hexadecimal notation using a colon-separated list of 16-bit values. The double colon (::) represents all bits set to 0.
- You must manually configure the router or switch advertisement and advertise the default prefix for autoconfiguration to work on a specific interface.

2. [Optional] Set the broadcast address on the network or subnet.

```
[edit interfaces interface-name unit logical-unit-number family family address address],
user@host# set broadcast address
```

NOTE: The broadcast address must have a host portion of either all ones or all zeros. You cannot specify the addresses `0.0.0.0` or `255.255.255.255`.

3. [Optional] specify the remote address of the connection for the encrypted, PPP-encapsulated, and tunnel interfaces.

```
[edit logical-systems logical-system-name interfaces interface-name unit logical-unit-number
family family address  address]
user@host# set destination address
```

Add a Logical Unit Description to the Configuration

You can include a text description of each logical unit in the configuration file. Any descriptive text that you include displays in the output of the `show interfaces` commands. It is also exposed in the `ifAlias` Management Information Base (MIB) object. It has no impact on the interface's configuration. To add a text description, include the `description` statement:

```
description text;
```

You can include this statement at the following hierarchy levels:

- `[edit interfaces interface-name unit logical-unit-number]`
- `[edit logical-systems logical-system-name interfaces interface-name unit logical-unit-number]`

The description can be a single line of text. If the text contains spaces, enclose it in quotation marks.



NOTE: You can configure the extended DHCP relay to include the interface description in the option 82 Agent Circuit ID suboption. See DHCP Relay Agent Information Option (Option 82).

For information about describing physical interfaces, see Configure the Interface Description.

Media MTU Overview

The media maximum transmission unit (MTU) for an interface is the largest data unit that can be forwarded through that interface without fragmentation.

The default media MTU depends on the encapsulation used on that interface and the default IP MTU. In some cases, the default IP MTU depends on whether the protocol used is IP version 4 (IPv4) or International Organization for Standardization (ISO).

The default media MTU for a physical interface is calculated as follows:

```
Default media MTU = Default IP MTU + encapsulation overhead
```

The actual frames transmitted also contain cyclic redundancy check (CRC) bits, which are not part of the media MTU. For example, the media MTU for a Gigabit Ethernet Version 2 interface is specified as 1514 bytes, but the largest possible frame size is actually 1518 bytes. You need to consider the extra bits when you calculate MTUs for interoperability.

Keep the following in mind when configuring the media MTU:

- The MTU size must be the same on both sides of a point-to-point connection.
- All interfaces in the subnet of point-to-multipoint connections must use the same MTU size.
- The physical MTU for Ethernet interfaces does not include the 4-byte frame check sequence (FCS) field of the Ethernet frame.
- A SONET/SDH interface operating in concatenated mode has a “c” added to the rate descriptor. For example, a concatenated OC48 interface is referred to as OC48c.
- The maximum number of data-link connection identifiers (DLCIs) is determined by the MTU on the interface. If you have keepalives enabled with the MTU set to 5012, the maximum number of DLCIs is 1000.

Because tunnel services interfaces are considered logical interfaces, you cannot configure the MTU setting for the associated physical interface. This means that you cannot configure the MTU size for the following interface types:

- Generic routing encapsulation (gr-)
- IP-IP (ip-)
- Loopback (lo-)
- Link services (ls-)
- Multilink services (ml-)
- Multicast (pe-, pd-)



NOTE: Starting in Junos OS Evolved release 24.4R1-S2, for PTX10008 routers, the minimum MTU value for IPv4 traffic will be 576 bytes.

Configure the Media MTU

If you change the size of the media MTU, you must ensure that the size is equal to or greater than the sum of the protocol MTU and the encapsulation overhead. In other words:

```
Minimum media MTU = protocol MTU + encapsulation overhead
```

The maximum media MTU size that you can configure depends on your device and the type of interface.



NOTE: Changing the media MTU or protocol MTU causes an interface to be deleted and added again. This causes the link to flap.

To configure the media MTU:

1. In configuration mode, go to the `[edit interfaces interface-name]` hierarchy level.

```
[edit]
user@host# edit interfaces interface-name
```

2. Include the `mtu` statement.

```
[edit interfaces interface-name]
user@host# set mtu bytes
```

Protocol MTU

IN THIS SECTION

- Overview | 54
- Platform-Specific Protocol MTU Behavior | 55

Overview

The default protocol MTU depends on your device and the interface type. When you initially configure an interface, the protocol MTU is calculated automatically. If you subsequently change the media MTU, the protocol MTU on existing address families automatically changes.

Use Feature Explorer to confirm platform and release support for specific features.

Review the "Platform-Specific Protocol MTU Behavior" on page 55 section for notes related to your platform.

If you reduce the media MTU size but one or more address families are already configured and active on the interface, you must also reduce the protocol MTU size. If you increase the size of the protocol MTU, you must ensure that the size of the media MTU is equal to or greater than the sum of the protocol MTU and the encapsulation overhead.

If you do not configure an MPLS MTU, Junos OS derives the MPLS MTU from the physical interface MTU. From this value, the software subtracts the encapsulation-specific overhead and space for the maximum number of labels that might be pushed in the Packet Forwarding Engine. The software provides for three labels of four bytes each, for a total of 12 bytes.

In other words, the formula used to determine the MPLS MTU is as follows:

```
MPLS MTU = physical interface MTU - encapsulation overhead - 12
```

You can configure the protocol MTU on all tunnel interfaces except virtual tunnel (VT) interfaces. Junos OS sets the MTU size for VT interfaces to unlimited by default.

Platform-Specific Protocol MTU Behavior

Platform	Difference
ACX Series	• ACX Series routers that support protocol MTU need to explicitly configure MTU at the family level for IPv4 and IPv6 make MTU exception work in egress. Follow the guidelines below while configuring MTUs. • If you configure MTUs for both inet and inet6 families, inet MTU gets precedence. • If you configure MTU only at inet level, the same value applies to inet6 as well. • If you configure MTU only for inet6 level, the same value applies to inet as well.

Configure the Interface Bandwidth

By default, the operating system uses the physical interface speed for the MIB-II object, `ifSpeed`. You can configure the logical unit to populate the `ifSpeed` variable by configuring a bandwidth value for the logical interface. The `bandwidth` statement sets an informational-only parameter; you cannot adjust the actual bandwidth of an interface with this statement.



NOTE: We recommend that you be careful when setting this value. Any interface bandwidth value that you configure using the `bandwidth` statement affects how the interface cost calculation for a dynamic routing protocol, such as OSPF. By default, the interface cost for a dynamic routing protocol is the following formula:

```
cost = reference-bandwidth/bandwidth,
```

In the formula, bandwidth is the physical interface speed. However, if you specify a value for bandwidth using the `bandwidth` statement, that value is used to calculate the interface cost rather than the actual physical interface bandwidth.

To configure the bandwidth value for a logical interface, include the `bandwidth` statement:

```
bandwidth rate;
```

You can include this statement at the following hierarchy levels:

- `[edit interfaces interface-name unit logical-unit-number]`
- `[edit logical-systems logical-system-name interfaces interface-name unit logical-unit-number]`

rate is the peak rate, in bits per second (bps) or cells per second (cps). You can specify a value in bps either as a complete decimal number or as a decimal number followed by the abbreviation `k` (1000), `m` (1,000,000), or `g` (1,000,000,000). You can also specify a value in cps by entering a decimal number followed by the abbreviation `c`. Values expressed in cps are converted to bps using the formula 1 cps = 384 bps. The value can be any positive integer. The `bandwidth` statement is valid for all logical interfaces except multilink interfaces.

Enable or Disable SNMP Notifications on Logical Interfaces

By default, Simple Network Management Protocol (SNMP) notifications are sent when the state of an interface or a connection changes.

To explicitly enable these notifications on the logical interface, include the `traps` statement:

```
(traps);
```

To explicitly disable these notifications on the logical interface, include the `no-traps` statement:

```
(no-traps);
```

You can include these statements at the following hierarchy levels:

- `[edit interfaces interface-name unit logical-unit-number]`
- `[edit logical-systems logical-system-name interfaces interface-name unit logical-unit-number]`

Overview of Accounting for the Logical Interface

IN THIS SECTION

- Accounting Profiles Overview | 57
- Configure Accounting for the Logical Interface | 57
- Introduction to Displaying the Accounting Profile for the Logical Interface | 59

This section discusses on how to configure accounting on logical interfaces.

Accounting Profiles Overview

Juniper Networks routers and switches can collect various kinds of data about traffic passing through the router and switch. You can set up one or more *accounting profiles* that specify some common characteristics of this data, including the following:

- The fields used in the accounting records
- The number of files that the router or switch retains before discarding, and the number of bytes per file
- The polling period that the system uses to record the data

You configure the profiles and define a unique name for each profile using statements at the `[edit accounting-options]` hierarchy level. There are two types of accounting profiles: interface profiles and filter profiles. You configure interface profiles by including the `interface-profile` statement at the `[edit accounting-options]` hierarchy level. You configure filter profiles by including the `filter-profile` statement at the `[edit accounting-options]` hierarchy level. For more information, see the Junos OS Network Management Administration Guide for Routing Devices.

You apply filter profiles by including the `accounting-profile` statement at the `[edit firewall filter` *`filter-name`*`]` and `[edit firewall family` *`family`* `filter` *`filter-name`*`]` hierarchy levels. For more information, see the Routing Policies, Firewall Filters, and Traffic Policers User Guide.

Configure Accounting for the Logical Interface

Before you begin

You must configure a profile to collect error and statistic information for input and output packets on a particular logical interface. An accounting profile specifies which statistics are collected and written to a

log file. For more information about how to configure an accounting-data log file, see the *Configuring Accounting-Data Log Files*.

An interface profile specifies the information collected and written to a log file. You can configure a profile to collect error and statistic information for input and output packets on a particular logical interface.

1. To configure which statistics are collected for an interface, include the `fields` statement at the `[edit accounting-options interface-profile profile-name]` hierarchy level.

```
[edit accounting-options interface-profile profile-name]
user@host# set fields field-name
```

2. Each accounting profile logs its statistics to a file in the **/var/log** directory. To configure which file to use, include the `file` statement at the `[edit accounting-options interface-profile profile-name]` hierarchy level.

```
[edit accounting-options interface-profile profile-name]
user@host# set file filename
```



NOTE: You must specify a `file` statement for the interface profile that has already been configured at the `[edit accounting-options]` hierarchy level. For more information, see Configuring Accounting-Data Log Files.

3. Each interface with an accounting profile enabled has statistics collected once per interval time specified for the accounting profile. Statistics collection time is scheduled evenly over the configured interval. To configure the interval, include the interval statement at the `[edit accounting-options interface-profile profile-name]` hierarchy level.

```
[edit accounting-options interface-profile profile-name]
user@host# set interval minutes
```



NOTE: The minimum interval allowed is 1 minute. Configuring a low interval in an accounting profile for a large number of interfaces might cause serious performance degradation.

4. To configure the interfaces on which the accounting needs to be performed, apply the interface profile to a logical interface by including the `accounting-profile` statement at the `[edit interfaces interface-name unit logical-unit-number]` hierarchy level.

```
[edit interfaces]
user@host# set interface-name unit logical-unit-number  accounting-profile profile-name
```

SEE ALSO

Accounting Options Overview

Configuring Accounting-Data Log Files

Introduction to Displaying the Accounting Profile for the Logical Interface

IN THIS SECTION

- Purpose | 59
- Action | 60
- Meaning | 60

Purpose

Displaying the configured accounting profile of a particular logical interface at the `[edit accounting-options interface-profile profile-name]` hierarchy level requires that you specify certain parameters:

- interface-name—ge-1/0/1
- Logical unit number—1
- Interface profile —`if_profile`
- File name—`if_stats`
- Interval—15 minutes

Action

- Run the `show` command at the `[edit interfaces ge-1/0/1 unit 1]` hierarchy level.

```
[edit interfaces ge-1/0/1 unit 1]
accounting-profile if_profile;
```

- Run the `show` command at the `[edit accounting-options]` hierarchy level.

```
interface-profile if_profile {
    interval 15;
    file if_stats {
        fields {
            input-bytes;
            output-bytes;
            input-packets;
            output-packets;
            input-errors;
            output-errors;
        }
    }
}
```

Meaning

The configured accounting and its associated set options are displayed as expected.

Disable a Logical Interface

You can unconfigure a logical interface, effectively disabling that interface, without removing the logical interface configuration statements from the configuration. To unconfigure a logical interface, include the `disable` statement:

```
disable;
```

You can include this statement at the following hierarchy levels:

- `[edit interfaces` *`interface-name`* `unit` *`logical-unit-number`*`]`

- [edit logical-systems *logical-system-name* interfaces *interface-name* unit *logical-unit-number*]

When an interface is disabled, a route (pointing to the reserved target "REJECT") with the IP address of the interface and a 32–bit subnet mask is installed in the routing table. See *Routing Protocols*.

Example: Disable a Logical Interface

Sample interface configuration:

```
[edit interfaces]
user@host# show
et-2/1/1  {
    vlan-tagging;
    encapsulation flexible-ethernet-services;
    unit 0 {
        vlan-id 1000;
        family inet {
            address 11.0.0.20/24;
        }
    }
}
```

Disabling the interface:

```
[edit interfaces et-2/1/1 unit 0]
user@host# set disable
```

Verifying the interface configuration:

```
[edit interfaces et-2/1/1]
user@host# show
disable; # Interface is marked as disabled.
    unit 0 {
    vlan-id 1000;
        family inet {
            address 11.0.0.20/24;
        }
}
```

Damping Interfaces

SUMMARY

You (the network administrator) can configure damping to reduce the advertisement of physical interface transitions between up and down states.

IN THIS SECTION

- Physical Interface Damping Overview | **62**
- Configure Damping of Shorter Physical Interface Transitions | **70**
- Configure Damping of Aggregated Ethernet Interface Transitions | **71**
- Configure Damping of Longer Physical Interface Transitions | **72**
- Example: Configure Physical Interface Damping | **74**

Physical Interface Damping Overview

IN THIS SECTION

- Damping Overview for Shorter Physical Interface Transitions | **63**
- Damping Overview for Longer Physical Interface Transitions | **64**

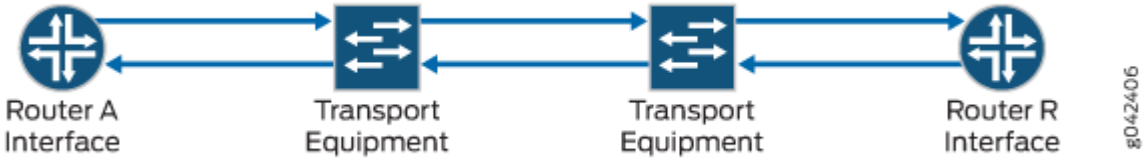
Physical interface damping limits the advertisement of the up-and-down transitions (flapping) on an interface. Each time a transition occurs, the interface state is changed, which generates an advertisement to the upper-level routing protocols. Damping helps reduce the number of these advertisements.

From the viewpoint of network deployment, physical interface flaps fall into the following categories:

- Nearly instantaneous multiple flaps of short duration (milliseconds)
- Periodic flaps of long duration (seconds)

Figure 1 on page 63 is used to describe these types of interface flaps and the damping configuration that you can use in each case.

Figure 1: Two Router Interfaces Connected Through Transport Equipment



NOTE: We recommend that you use similar damping configurations on both ends of the physical interface. Configuring interface damping on one end and not configuring interface damping on the other end can result in undesired behavior.

The types of interface damping depend upon the transition time length.

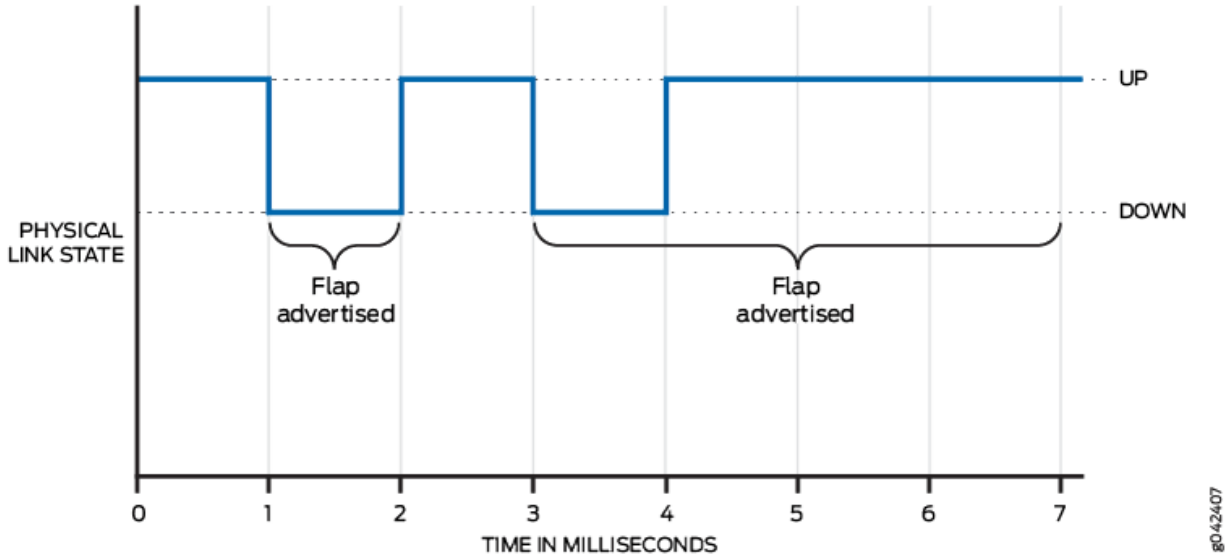
Damping Overview for Shorter Physical Interface Transitions

Figure 1 on page 63 shows two routers with two transport devices between them. If a redundant link between the two transport devices fails, Junos OS performs link switching. Link switching takes a number of milliseconds. As shown in Figure 2 on page 64, during switching, both device interfaces might encounter multiple flaps with an up-and-down duration of several milliseconds. These multiple flaps, if advertised to the upper-level routing protocols, might result in undesired route updates. This is why you might want to damp these interface flaps.

NOTE: Damping is suitable only with routing protocols.

For shorter physical interface transitions, you configure interface damping with the `hold-time` statement on the interface. The hold timer enables interface damping by not advertising interface transitions until the hold timer duration has passed. When a hold-down timer is configured and the interface goes from up to down, the down hold-time timer is triggered. Every interface transition that occurs during the hold time is ignored. When the timer expires and the interface state is still *down*, then the router begins to advertise the interface as being down. Similarly, when a hold-up timer is configured and an interface goes from down to up, the up hold-time timer is triggered. Every interface transition that occurs during the hold time is ignored. When the timer expires and the interface state is still *up*, then the router begins to advertise the interface as being up.

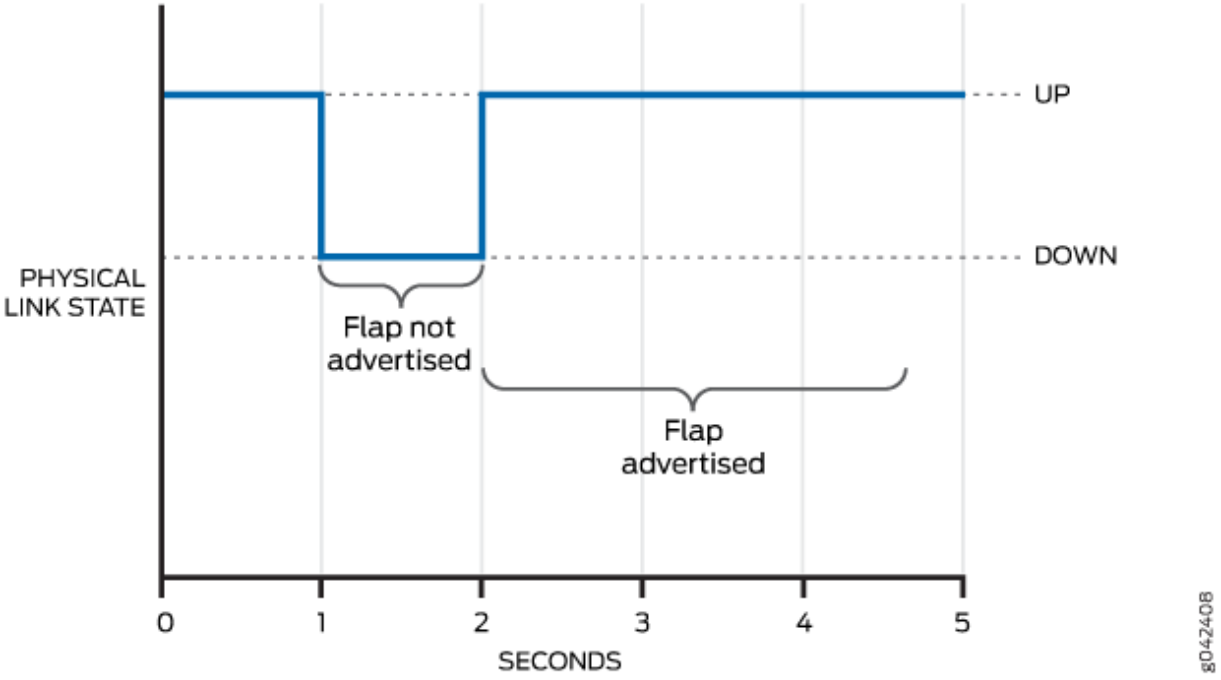
Figure 2: Multiple Flaps of Short Duration (Milliseconds)



Damping Overview for Longer Physical Interface Transitions

When the link between a router interface and the transport devices is not stable, this can lead to periodic flapping, as shown in Figure 3 on page 65. Flaps occur in the order of seconds or more, with an up-and-down flap duration in the order of a second or more. In this case, using the hold timer feature might not produce optimal results because it cannot suppress the relatively longer and repeated interface flaps. Increasing the hold-time duration to seconds still allows the system to send route updates on the flapping interface. Increasing the duration therefore fails to suppress periodically flapping interfaces on the system.

Figure 3: Periodic Flaps of Long Duration (Seconds)



For longer periodic interface flaps, configure interface damping with the `damping` statement on the interface. This damping method uses an exponential back-off algorithm to suppress interface up-and-down event reporting to the upper-level protocols. Every time an interface goes down, Junos OS adds a penalty to the interface penalty counter. If at some point the accumulated penalty exceeds the suppress level, Junos OS places the interface in the suppress state. In this case, Junos OS does not report further interface link up-and-down events to the upper-level protocols.

The penalty added on every interface flap is 1000. At all times, the interface penalty counter follows an exponential decay process. Figure 4 on page 67 and Figure 5 on page 69 show the decay process as it applies to recovery when the physical level link is down or up. As soon as the accumulated penalty reaches the lower boundary of the reuse level, the interface is marked as unsuppressed, and further changes in the interface link state are again reported to the upper-level protocols. You use the `max-suppress` option to configure the maximum time for restricting the accumulation of the penalty beyond the value of the maximum penalty. The value of the maximum penalty is calculated by the software. The maximum penalty corresponds to the time it would take max-suppress to decay and reach the reuse level. The penalty continues to decay after crossing the reuse level.

Figure 4 on page 67 and Figure 5 on page 69 show the accumulated penalty and the decay over time as a curve. Whenever the penalty is below the reuse level and the physical level link changes state, state changes are advertised to the system and cause SNMP state changes.

Figure 4 on page 67 shows the penalty dropping below the reuse level when the physical link is down. The system is notified of a state change only after the physical level link transitions to up.

Figure 4: Physical-Level Link Is Down When the Penalty Falls Below the Reuse Level

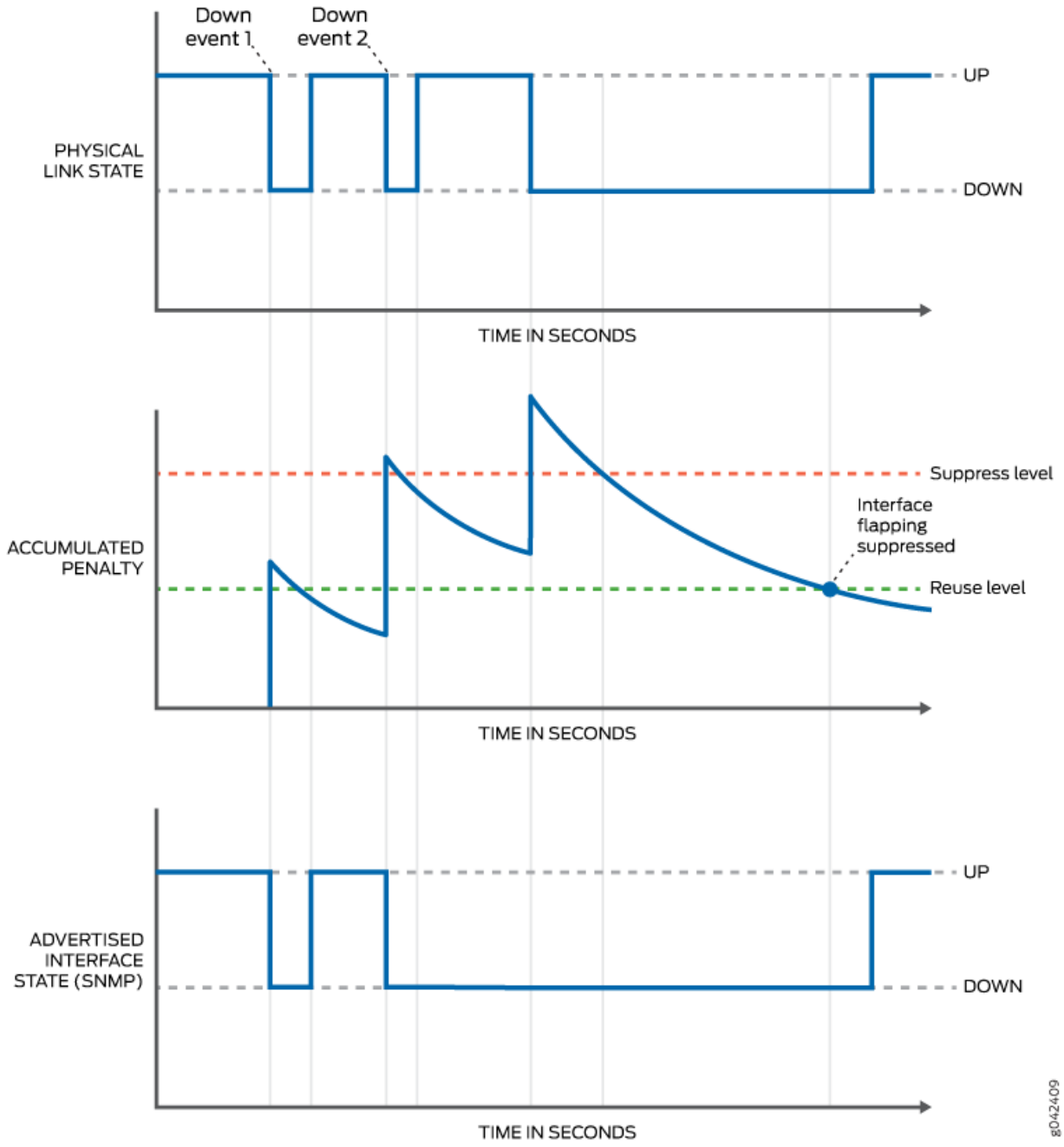
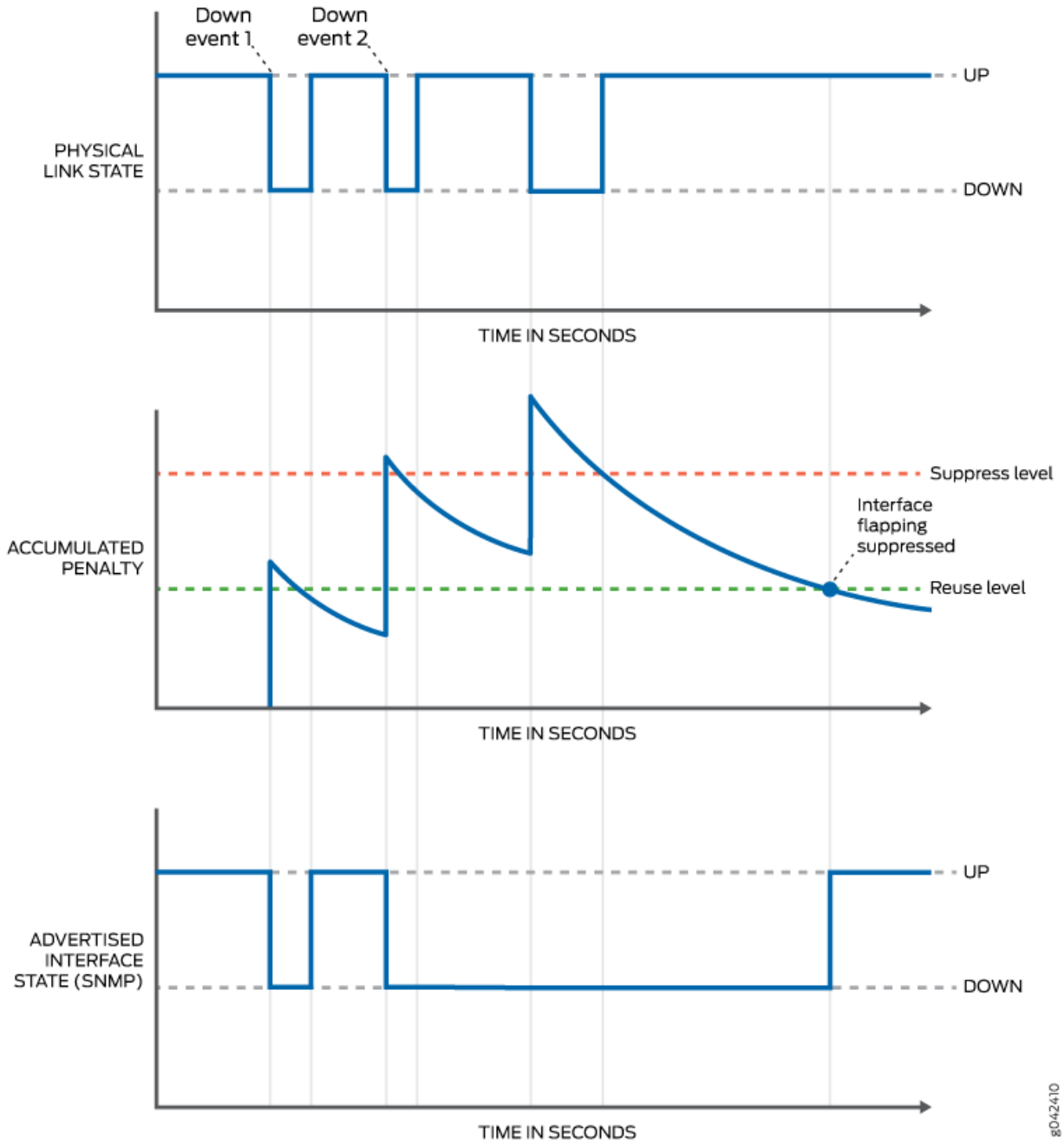


Figure 5 on page 69 shows the penalty dropping below the reuse level when the physical link is up. The system is notified of a state change immediately.

Figure 5: Physical-Level Link Is Up When the Penalty Falls Below the Reuse Level





NOTE: The QFX10002-72Q switches do not support `hold-time down` time of less than 1 second on 100G interfaces. The recommended `hold-time down` time is 3 seconds.

Configure Damping of Shorter Physical Interface Transitions

By default, when an interface changes from up to down or from down to up, this transition is advertised immediately to the hardware and Junos OS. In some situations, you might want to damp interface transitions.

For example, you may want to configure damping on an interface that is connected to an add/drop multiplexer (ADM) or wavelength-division multiplexer (WDM), or to protect against SONET/SDH framer holes.

Damping the interface means not advertising the interface's transition until a certain period of time has passed, called the *hold-time*. When the interface goes from up to down, the down hold-time timer is triggered. Every interface transition that occurs during the hold time is ignored. If the timer expires and the interface state is still *down*, then the router begins to advertise the interface as being down. Similarly, when an interface goes from down to up, the up hold-time timer is triggered. Every interface transition that occurs during the hold time is ignored. If the timer expires and the interface state is still *up*, then the router begins to advertise the interface as being up.

To configure damping of shorter physical interface transitions in milliseconds:

1. Select the interface to damp, where the interface name is *interface-type-fpc/pic/port*:

```
[edit]
user@host# edit interfaces interface-name
```

2. Configure the hold time for link up and link down.

```
[edit interfaces interface-name]
user@host# set hold-time up milliseconds down milliseconds
```

The hold time can be a value from 0 through 4,294,967,295 milliseconds. The default value is 0, which means that interface transitions are not damped. Junos OS advertises the transition within 100 milliseconds of the time value you specify.

For most Ethernet interfaces, Junos OS implements hold timers using a one-second polling algorithm. For 1-port, 2-port, and 4-port Gigabit Ethernet interfaces with small form-factor pluggable (SFP) transceivers, hold timers are interrupt driven.

NOTE: The `hold-time` option is not available for controller interfaces.

Configure Damping of Aggregated Ethernet Interface Transitions

By default, when an interface changes from up to down or from down to up, this transition is advertised immediately to the hardware and Junos OS. In some situations, you might want to damp interface transitions.

For example, you may want to configure damping on an interface that is connected to an add/drop multiplexer (ADM) or wavelength-division multiplexer (WDM), or to protect against SONET/SDH framer holes.

Damping the interface means not advertising the interface's transition until a certain period of time has passed, called the *hold-time*. When the interface goes from up to down, the down hold-time timer is triggered. Every interface transition that occurs during the hold time is ignored. If the timer expires and the interface state is still *down*, then the router begins to advertise the interface as being down. Similarly, when an interface goes from down to up, the up hold-time timer is triggered. Every interface transition that occurs during the hold time is ignored. If the timer expires and the interface state is still *up*, then the router begins to advertise the interface as being up.

To configure damping of aggregated ethernet interface transitions in ms:

1. Select the interface to damp, where the interface name is *interface-type-fpc/pic/port*:

```
[edit]
user@host# edit interfaces aex
```

2. Configure the hold time for link up and link down.

```
[edit interfaces aex]
user@host# set hold-time up milliseconds down milliseconds
```

The hold time can be a value from 0 through 4,294,967,295 milliseconds. The default value is 0, which means that interface transitions are not damped. Junos OS advertises the transition within 100 milliseconds of the time value you specify.

For most Ethernet interfaces, Junos OS implements hold timers using a one-second polling algorithm. For 1-port, 2-port, and 4-port Gigabit Ethernet interfaces with small form-factor pluggable (SFP) transceivers, hold timers are interrupt driven.

Starting from Junos OS release 21.4R1, you can specify the hold-time value on aggregated ethernet interfaces (aex). When you configure hold timer for aggregated Ethernet interfaces, we recommend not to configure the hold-time for member links.



NOTE: The `hold-time` option is not available for controller interfaces.

Configure Damping of Longer Physical Interface Transitions

Physical interface damping limits the advertisement of the up-and-down transitions (flapping) on an interface. An unstable link between a router Interface and the transport devices can lead to periodic flapping. Longer flaps occur with a period of about five seconds or more, with an up-and-down duration of one second.

For these longer periodic interface flaps, configure interface damping with the `damping` statement on the interface. This damping method uses an exponential back-off algorithm to suppress interface up-and-down event reporting to the upper-level protocols. Every time an interface goes down, a penalty is added to the interface penalty counter. If at some point the accumulated penalty exceeds the suppress level `max-suppress`, the interface is placed in the suppress state, and further interface state up-and-down transitions are not reported to the upper-level protocols.

You can view the damping parameters with the `show interfaces extensive` command.



NOTE: Only PTX Series routers, T Series routers, MX2010 routers, MX2020 routers, MX960 routers, MX480 routers, MX240 routers, MX80 routers, and M10i routers support interface damping for longer periodic interface flaps.

To configure damping of longer physical interface transitions:

1. Select the interface to damp, where the interface name is *interface-type-fpc/pic/port* or an interface range:

```
[edit]
user@host# edit interfaces interface-name damping
```

2. Enable longer interface transition damping on a physical interface:

```
[edit interfaces interface-name damping]
user@host# set enable
```

3. (Optional) Set the maximum time in seconds that an interface can be suppressed no matter how unstable the interface has been.



NOTE: Configure `max-suppress` to a value that is greater than the value of `half-life`; otherwise, the configuration is rejected.

```
[edit interfaces interface-name damping]
user@host# set max-suppress maximum-seconds
```

4. (Optional) Set the decay half-life in seconds, which is the interval after which the accumulated interface penalty counter is reduced by half if the interface remains stable.



NOTE: Configure `half-life` to a value that is less than the value of `max-suppress`; otherwise, the configuration is rejected.

```
[edit interfaces interface-name damping]
user@host# set half-life seconds
```

5. (Optional) Set the reuse threshold (no units). When the accumulated interface penalty counter falls below this value, the interface is no longer suppressed.

```
[edit interfaces interface-name damping]
user@host# set reuse number
```

6. (Optional) Set the suppression threshold (no units). When the accumulated interface penalty counter exceeds this value, the interface is suppressed.

```
[edit interfaces interface-name damping]
user@host# set suppress number
```

NOTE: The system does not indicate whether an interface is down because of suppression or because that is the actual state of the physical interface. Therefore, neither SNMP link traps nor Operation, Administration, and Maintenance (OAM) protocols can differentiate the damped version of the link state from the real version. Therefore, traps and protocols might not work as expected.

You can verify suppression by viewing the information in the `Damping` field of the `show interface extensive` command output.

Example: Configure Physical Interface Damping

IN THIS SECTION

- Requirements | 74
- Overview | 74
- Configuration | 75
- Verification | 76

This example shows how to configure damping for a physical interface on a PTX Series Packet Transport Router.

Requirements

This example uses the following hardware and software components:

- One PTX Series Packet Transport Router
- One or more routers that provide input packets and receive output packets
- Junos OS Release 14.1 or later

Overview

Physical interface damping provides a smoothing of the up-and-down transitions (flapping) on an interface. Each time a transition occurs, the interface state is changed, which generates an advertisement to the upper-level routing protocols. Damping helps reduce the number of these advertisements.

From the viewpoint of network deployment, physical interface flaps fall into these categories:

- Nearly instantaneous multiple flaps of short duration (milliseconds). For shorter physical interface transitions, you configure interface damping with the `hold-time` statement on the interface. The hold timer enables interface damping by not advertising interface transitions until the hold-timer duration has passed. When a hold-down timer is configured and the interface goes from up to down, the interface is not advertised to the rest of the system as being down until it has remained down for the hold-down timer period. Similarly, when a hold-up timer is configured and an interface goes from down to up, it is not advertised as being up until it has remained up for the hold-up timer period.
- Periodic flaps of long duration (seconds). For longer periodic interface flaps, you configure interface damping with the `damping` statement on the interface. This damping method uses an exponential back-off algorithm to suppress interface up-and-down event reporting to the upper-level protocols. Every time an interface goes down, a penalty is added to the interface penalty counter. If at some point the accumulated penalty exceeds the suppress level, the interface is placed in the suppress state, and further interface state up transitions are not reported to the upper-level protocols.

Configuration

IN THIS SECTION

- CLI Quick Configuration | 75
- Procedure | 75
- Results | 76

CLI Quick Configuration

To quickly configure this example, copy the following commands, paste them into a text file, remove any line breaks, change any details necessary to match your network configuration, and then copy and paste the commands into the CLI at the `[edit]` hierarchy level.

```
set interfaces xe-6/0/0 damping half-life 11 max-suppress 2222 reuse 3333 suppress 4444 enable
```

Procedure

Step-by-Step Procedure

To configure damping on the PTX Series Packet Transport Router:

1. Set the half-life interval, maximum suppression, reuse, suppress values, and enable:

```
[edit interface]
user@router# set xe-6/0/0 damping half-life 11 max-suppress 2222 reuse 3333 suppress 4444
enable
```

2. Commit the configuration:

```
[edit]
user@router# commit
```

Results

From configuration mode, confirm your configuration by entering the `show interfaces` command. If the output does not display the intended configuration, repeat the instructions in this example to correct the configuration.

```
user@router# show interfaces
xe-6/0/0 {
    damping {
        half-life 11;
        max-suppress 2222;
        reuse 3333;
        suppress 4444;
        enable;
    }
```

Verification

IN THIS SECTION

- Verify Interface Damping on xe-6/0/0 | **77**

To confirm that the configuration is working properly, perform this task:

Verify Interface Damping on xe-6/0/0

Purpose

Verify that damping is enabled on the interface and that the damping parameter values are correctly set.

Action

From operational mode, run the `show interfaces extensive` command.

```
user@router# run show interfaces xe-6/0/0 extensive
Physical interface: xe-6/0/0, Enabled, Physical link is Up
  Interface index: 158, SNMP ifIndex: 535, Generation: 161
  Link-level type: Ethernet, MTU: 1514, LAN-PHY mode, Speed: 10Gbps, BPDU Error: None, Loopback:
None,
  Source filtering: Disabled, Flow control: Enabled
  Device flags   : Present Running
  Interface flags: SNMP-Traps Internal: 0x4000
  Link flags     : None
  CoS queues     : 8 supported, 8 maximum usable queues
  Hold-times     : Up 0 ms, Down 0 ms
  Damping        : half-life: 11 sec, max-suppress: 2222 sec, reuse: 3333, suppress: 4444,
state: unsuppressed
```

Meaning

Damping is enabled and configured successfully on the xe-6/0/0 interface.

Interface Ranges

IN THIS SECTION

- Understanding Interface Ranges for Switches | **78**
- Configuring Interface Ranges for EX Series Switches with ELS | **81**

Interface ranges represent similar type of interfaces with common configurations that are grouped together. The ranges containa a name, a range and the configuration statements which is common to all the similar interfaces.

Understanding Interface Ranges for Switches

You can use the interface ranges to group interfaces of the same type that share a common configuration profile. This helps reduce the time and effort in configuring interfaces on Juniper Networks EX Series Ethernet switches. The configurations common to all the interfaces can be included in the interface range definition.

The interface range definition contains the name of the interface range defined, the names of the individual member interfaces that do not fall in a series of interfaces, a range of interfaces defined in the member range, and the configuration statements common to all the interfaces. An interface range defined with member ranges and individual members but without any common configurations, is also a valid definition.



NOTE: The interface range definition is supported only for Gigabit, 10-Gigabit, and Fast Ethernet interfaces. OCX Series switches do not support Fibre Channel interfaces.

Starting in Junos OS Release 14.1X53-D15 and later, the common configurations defined in the interface range will not be overridden but appended to the local configuration. In Junos OS Releases prior to 14.1X53-D15, the common configurations defined in the interface range will be overridden by the local configuration.

The defined interface ranges can be used at places where the interface node is used in the following configuration hierarchies:

Table 9 on page 79 lists the configuration hierarchies for the EX Series, NFX, OCX, QFX Series, and QFabric Series.

Table 9: Configuration hierarchies for EX Series

Configuration Hierarchies for EX Series	Configuration Hierarchies for EX4600, NFX, QFX Series, and QFabric Systems	Configuration Hierarchies for EX Series with ELS
• ethernet-switching-options analyzer *name* input egress interface • ethernet-switching-options analyzer *name* input ingress interface • ethernet-switching-options analyzer output interface • ethernet-switching-options bpdu-block interface • ethernet-switching-options interfaces • ethernet-switching-options redundant-trunk-group group-name interface • ethernet-switching-options secure-access-port interface • ethernet-switching-options voip interface • poe interface • protocols dot1x authentication interface • protocols gvrp interface • protocols igmp interface • protocols igmp-snooping vlan *vlan-name* interface • protocols isis interface	• protocols isis interface • protocols sflow interfaces **NOTE**: These statements are not supported on OCX Series switches.	• forwarding-options analyzer *name* input egress interface • forwarding-options analyzer *name* input ingress interface • poe interface • protocols dot1x authenticator interface • protocols igmp interface • protocols isis interface • protocols layer2-control bpdu-block interface • protocols link-management peer *name* lmp-control-channel • protocols link-management te-link *name* interface • protocols lldp interface • protocols lldp-med interface • protocols mstp interface • protocols oam ethernet link-fault-management interface • protocols ospf area *area-id* interface • protocols pim interface • protocols router-advertisement interface

Table 9: Configuration hierarchies for EX Series *(Continued)*

Configuration Hierarchies for EX Series	Configuration Hierarchies for EX4600, NFX, QFX Series, and QFabric Systems	Configuration Hierarchies for EX Series with ELS
• protocols link-management peer lmp-control-channel interface • protocols link-management te-link *name* interface • protocols lldp interface • protocols lldp-med interface • protocols mpls interface • protocols mstp interface • protocols mstp msti-*id* interface • protocols mstp msti-*id* vlan *vlan-id* interface • protocols oam ethernet link-fault-management interface • protocols ospf area • protocols pim interface • protocols rip group *group-name* neighbor • protocols ripng group *group-name* neighbor • protocols router-advertisement interface • protocols router-discovery interface • protocols rsvp interface		• protocols router-discovery interface • protocols rsvp interface • protocols sflow interfaces • protocols vstp vlan *vlan-id* interface • switch-options redundant-trunk-group *group-name* interface • switch-options voip interface For ELS details, see Using the Enhanced Layer 2 Software CLI.

Table 9: Configuration hierarchies for EX Series *(Continued)*

Configuration Hierarchies for EX Series	Configuration Hierarchies for EX4600, NFX, QFX Series, and QFabric Systems	Configuration Hierarchies for EX Series with ELS
• protocols sflow interfaces • protocols stp interface • protocols vstp vlan *vlan-id* interface • vlans *vlan-name* interface		

SEE ALSO

- Configuring Interface Ranges
- Configuring Gigabit Ethernet Interfaces (CLI Procedure)
- Configuring Aggregated Ethernet Links (CLI Procedure)
- Configuring a Layer 3 Subinterface (CLI Procedure)
- interface-range
- Configuring Link Aggregation | **283**
- Configuring a Layer 3 Logical Interface
- Junos OS Network Interfaces Library for Routing Devices

Configuring Interface Ranges for EX Series Switches with ELS

IN THIS SECTION

- Configuring Interface Ranges on Switches | **82**
- Expanded Interface Range Statements | **85**
- Configuration Inheritance for Member Interfaces | **86**

- Configuration Group Inheritance | 88
- Common Configuration Inheritance | 90
- Configuration Inheritance Priority | 90
- Configuration Expansion Where Interface Range Is Used | 91

NOTE: This task uses Junos OS for EX Series switches with support for the Enhanced Layer 2 Software (ELS) configuration style. If your switch runs software that does not support ELS, see Configuring Interface Ranges. For ELS details, see Using the Enhanced Layer 2 Software CLI.

Junos OS allows you to group a range of identical interfaces into an *interface range*. You first specify the group of identical interfaces in the interface range. Then you can apply a common configuration to the specified interface range, reducing the number of configuration statements required and saving time while producing a compact configuration.

Configuring Interface Ranges on Switches

To configure an interface range, include the `interface-range` statement at the `[edit interfaces]` hierarchy level.

The `interface-range` statement accepts only physical networking interface names in its definition.

Interfaces can be grouped either as a range of interfaces or using a number range under the `interface-range` statement definition.

Interfaces in an **interface-range** definition can be added as part of a member range or as individual members or multiple members using a number range.

To specify a member range, use the `member-range` statement at the `[edit interfaces interface-range` *`name`*`]` hierarchy level.

To specify interfaces in lexical order, use the `member-range` *`start-range`* `to` *`end-range`* statement.

A range for a member statement must contain the following:

- ***—All, specifies sequential interfaces from 0 through 47.

CAUTION: The wildcard * in a member statement does not take into account the interface numbers supported by a specific interface type. Irrespective of the interface

type, * includes interface numbers ranging from 0 through 47 to the interface group. Therefore, use * in a member statement with caution.

- **num**—Number; specifies one specific interface by its number.
- `[low-high]`—Numbers between low to high; specifies a range of sequential interfaces.
- `[num1, num2, num3]`—Numbers **num1**, **num2**, and **num3** specify multiple specific interfaces.

Example: Specifying an Interface Range Member Range

```
member-range ge-0/0/0 to ge-4/0/40;
```

To specify one or multiple members, use the `member` statement at the `[edit interfaces interface-range` *`name`*`]` hierarchy level.

To specify the list of interface range members individually or for multiple interfaces using regex, use the `member` *`list of interface names`* statement.

Example: Specifying an Interface Range Member

```
member ge-0/0/0;
member ge-0/*/*
member ge-0/[1-10]/0;
member ge-0/[1,2,3]/3;
```

Regex or wildcards are not supported for interface-type prefixes. For example, prefixes **ge**, **fe**, and **xe** must be mentioned explicitly.

An **interface-range** definition can contain both **member** and `member-range` statements within it. There is no maximum limit on the number of **member** or `member-range` statements within an interface-range. However, at least one **member** or `member-range` statement must exist within an **interface-range** definition.

Example: Interface Range Common Configuration

Configuration common to an interface range can be added as a part of the **interface-range** definition, as follows:

```
[edit]
interfaces {
    +    interface-range foo {
    +        member-range ge-1/0/0 to ge-4/0/40;
    +        member ge-0/1/1;
```

```
    +        member ge-5/[1-10]/*;

        /*Common configuration is added as part of interface-range definition*/
        mtu 256;
        hold-time up 10;
        ether-options {
            flow-control;
            speed {
                100m;
            }
            802.3ad primary;
        }

    }
}
```

An **interface-range** definition having just **member** or `member-range` statements and no common configurations statements is valid.

These defined interface ranges can be used in other configuration hierarchies, in places where an **interface** node exists.

Example: Interface-Range foo Used Under the Protocols Hierarchy

```
protocols {
    dot1x {
        authenticator {
            interface foo{
                retries 1;
            }
        }
    }
}
```

foo should be an **interface-range** defined at the `[interfaces]` hierarchy level. In the above example, the **interface** node can accept both individual interfaces and interface ranges.



TIP: To view an interface range in expanded configuration, use the (`show | display inheritance`) command. For more information, see the Junos OS CLI User Guide.

The defined interface ranges can be used at places where the interface node is used. To view the configuration hierarchies, see "Understanding Interface Ranges for Switches" on page 78.

Expanded Interface Range Statements

The operating system expands all `member` and `member-range` statements in an interface range definition to generate the final list of interface names for the specified interface range.

An example configuration looks like this before it is expanded:

```
[edit]
interfaces {
    interface-range range1 {
        member-range et-0/0/0 to et-4/0/20;
        member et-10/1/1;
        member et-5/[0-5]/*;

        /*Common configuration is added as part of the interface-range definition*/
        mtu 256;
        hold-time up 10;
        ether-options {
            flow-control;
            speed {
                100m;
            }
            802.3ad primary;
        }

    }
}
```

For the `member-range` statement, all possible interfaces between `start-range` and `end-range` are considered in expanding the members. For example, the following `member-range` statement:

```
member-range et-0/0/0 to et-4/0/20
```

expands to:

```
      [et-0/0/0, et-0/0/1 ... et-0/0/max_ports
       et-0/1/0  et-0/1/1 ... et-0/1/max_ports
       et-0/2/0  et-0/2/1 ... et-0/2/max_ports
                           .
                           .
       et-0/MAX_PICS/0 ... et-0/max_pics/max_ports
```

```
        et-1/0/0  et-1/0/1 ... et-1/0/max_ports
                           .
        et-1/MAX_PICS/0 ... et-1/max_pics/max_ports
                           .
                           .
        et-4/0/0 et-4/0/1  ... et-4/0/max_ports]
```

The following `member` statement:

```
et-5/[0-5]/*
```

expands to:

```
      et-5/0/0 ... et-5/0/max_ports
      et-5/1/0 ... et-5/0/max_ports
          .
           .
      et-5/5/0 ... et-5/5/max_ports
```

The following `member` statement:

```
et-5/1/[2,3,6,10]
```

expands to:

```
      et-5/1/2
      et-5/1/3
      et-5/1/6
      et-5/1/10
```

Configuration Inheritance for Member Interfaces

When Junos OS expands the `member` and `member-range` statements present in an `interface-range`, it creates *interface objects* if they are not explicitly defined in the configuration. The operating system copies the common configuration to all the interface range's member interfaces.

Foreground interface configuration takes priority over configuration that the interface inherits from the interface range configuration.

In this example, interface `et-1/0/1` has an MTU value of 1024 because that is its foreground configuration:

```
interfaces {
    interface-range range1 {
        member-range et-1/0/0 to et-7/0/47;
        mtu 500;
        }

    et-1/0/1 {
        mtu 1024;
    }
}
```

You can verify this in the output of the `show interfaces | display inheritance` command:

```
user@host: show interfaces | display inheritance
##
## 'et-1/0/0' was expanded from interface-range 'range1'
##
et-1/0/0 {
    ##
    ## '500' was expanded from interface-range 'range1'
    ##
    mtu 500;
}
et-1/0/1 {
    mtu 1024;
}
##
## 'et-1/0/2' was expanded from interface-range 'range1'
##
et-1/0/2 {
    ##
    ## '500' was expanded from interface-range 'range1'
    ##
    mtu 500;
}
      .........
      .........
##
```

```
## 'et-10/0/47' was expanded from interface-range 'range1'
##
et-10/0/47 {
    ##
    ## '500' was expanded from interface-range 'range1'
    ##
    mtu 500;
}
```

Configuration Group Inheritance

Interface range member interfaces inherit configurations from configuration groups like any other foreground configuration. The only difference is that the `interface-range` goes through a member interfaces expansion before the operating system reads this configuration.

In this example, Junos OS applies the `hold-time` configuration to all members of the interface range `range1`:

```
groups {
     global {
         interfaces {
             <*> {
                  hold-time up 10;
             }
         }
     }
}
apply-groups [global];
interfaces {
     interface-range range1 {
         member-range et-1/0/0 to et-7/0/47;
         mtu 500;
     }
}
```

You can verify this with `show interfaces | display inheritance`, as follows:

```
user@host# show interfaces | display inheritance
[...]
##
## 'et-1/0/0' was expanded from interface-range 'range1'
##
```

```
et-1/0/0 {
    ##
    ## '500' was expanded from interface-range 'range1'
    ##
    mtu 500;
    ##
    ## 'hold-time' was inherited from group 'global'
    ## '10' was inherited from group 'global'
    ##
    hold-time up 10;
}
##
## 'et-1/0/1' was expanded from interface-range 'range1'
##
et-1/0/1 {
    ##
    ## '500' was expanded from interface-range 'range1'
    ##
    mtu 500;
    ##
    ## 'hold-time' was inherited from group 'global'
    ## '10' was inherited from group 'global'
    ##
    hold-time up 10;
}
##
## 'et-7/0/47' was expanded from interface-range 'range1'
##
et-7/0/47 {
    ##
    ## '500' was expanded from interface-range 'range1'
    ##
    mtu 500;
    ##
    ## 'hold-time' was inherited from group 'global'
    ## '10' was inherited from group 'global'
    ##
    hold-time up 10;
}
```

SEE ALSO

> *Using Wildcards with Configuration Groups*

Common Configuration Inheritance

If an interface is a member of multiple interface ranges, that interface will inherit the common configuration from all of those interface ranges.

For example:

```
[edit]
interfaces {
    interface-range int-grp-one {
        member-range et-0/0/0 to et-4/0/40;

        mtu 256;
    }
    interface-range int-grp-two {
        member-range et-4/0/0 to et-4/0/40;

        hold-time up 10;
    }
}
```

In this example, interfaces `et-4/0/0` through `et-4/0/40` have both `hold-time` and `mtu` configured.

Configuration Inheritance Priority

The interface ranges are defined in the order of inheritance priority. The first interface range configuration data takes priority over subsequent interface ranges.

In this example, interface `et-1/1/1` exists in both interface range `int-grp-one` and interface range `int-grp-two`:

```
[edit]
interfaces {
    interface-range int-grp-one {
        member-range et-0/0/0 to et-4/0/47;
        member et-1/1/1;

        /*Common config is added part of the interface-range definition*/
        mtu 500;
        hold-time up 10;
    }
```

```
    interface-range int-grp-two {
        member-range et-5/0/0 to et-7/0/47;
        member et-1/1/1;

        mtu 1024;
    }
}
```

Interface et-1/1/1 inherits mtu *500* from interface range int-grp-one because it was defined first.

Configuration Expansion Where Interface Range Is Used

In this example, interface-range *range1* is used under the protocols hierarchy:

```
[edit]
interfaces {
    interface-range range1 {
        member et-7/1/1;
        member et-5/0/1;

        mtu 500;
        hold-time up 10;
        ether-options {
            flow-control;
            speed {
                100m;
            }
            802.3ad primary;
        }
    }
    protocols {
        dot1x {
            authenticator {
                interface range1 {
                    retries 1;
                }
            }
        }
    }
}
```

The `interface` node present under `authenticator` expands into member interfaces of the interface range `range1` as follows:

```
protocols {
    dot1x {
        authenticator {
            interface et-7/1/1 {
                retries 1;
            }
            interface et-5/0/1 {
                retries 1;
            }
        }
    }
}
```

The `interface` *`range-1`* statement is expanded into two interfaces, et-7/1/1 and et-5/0/1, and the operating system copies the configuration `retries` *`1`* under those two interfaces.

You can verify this configuration using the `show protocols dot1x | display inheritance` command.

RELATED DOCUMENTATION

- Physical Interfaces

Gigabit Ethernet Interface

SUMMARY

You can configure Gigabit Ethernet Interface with various modes like speed options, autonegotiation options, VLAN options, IP options, interface modes, link settings on the switches. The configuration uses Junos OS for EX Series switches with support for the Enhanced Layer 2 Software (ELS) configuration style.

IN THIS SECTION

- Speed and Autonegotiation | **93**
- Configuring Gigabit and 10-Gigabit Ethernet Interfaces for EX4600 and QFX Series Switches | **105**

- Configuring Gigabit Ethernet Interfaces for EX Series Switches with ELS support | **111**
- Configuring Gigabit and 10-Gigabit Ethernet Interfaces for OCX Series Switches | **116**

Speed and Autonegotiation

IN THIS SECTION

- Configure Interface Speed on Switches | **93**
- Configure Speed on EX2300-48MP and EX2300-24MP Switches | **94**
- Configure Speed and Autonegotiation on QFX5100-48T Switches | **94**
- Autonegotiation Support for EX4300-48MP Switches | **96**
- Autonegotiation Support for EX4400 Switches | **97**
- Autonegotiation Support for EX4100 Switches | **98**
- Autonegotiation Support for EX4600-40F, QFX5110-48S and QFX5100-48S with JNP-SFPP-10GE-T Transceiver | **99**
- Autonegotiation Support for QFX5120-48Y with JNP-SFPP-10GE-T Transceiver | **100**
- Operating Speed of Interfaces on EX Switches | **101**

Configure Interface Speed on Switches

On 1/10GbE capable SFP interfaces, the duplex is always full and the speed matches that of the inserted optic. These interfaces support either 1 GbE or 10 GbE SFP optics.

- For EX Series Switches and QFX Series Switches, the CLI configuration needs to match the optic speed. For 1 GbE based optics, the configuration needs to start with ge-. For 10GbE inserted optics, the configuration needs to start with xe-. By default, both ge and xe choices are in the default configuration. User must match the CLI syntax to the optic speed.

NOTE: Only 10 Gbps and 40 Gbps interfaces are supported on OCX Series switches.

Following are the steps to configure the speeds on EX Series Switches and QFX Series Switches:

1. In configuration mode, go to the `[edit interfaces interface-name]` hierarchy level.

```
[edit ]
user@host# edit interfaces interface-name
```

2. To configure the speed, include the `speed` statement at the `[edit interfaces interface-name]` hierarchy level.

```
[edit interfaces interface-name]
user@host# set speed speed;
```

Configure Speed on EX2300-48MP and EX2300-24MP Switches

Follow these guidelines when you configure the speed on EX2300-48 MP and EX2300-24MP switches:

- The mge interface is a rate-selectable (Multi-Rate) GbE interface that supports speeds of 10-Gbps, 5-Gbps, and 2.5-Gbps over CAT5e/CAT6/CAT6a cables. In the EX2300-48 MP and EX2300-24MP switches, the mge interface supports 100-Mbps, 1-Gbps, and 2.5-Gbps speeds. Note that 10 Mbps speed is supported only on `ge` interfaces of EX2300 switch.
- You can configure both multi-rate gigabit ethernet interface (mge) and gigabit ethernet (ge) interface using the `speed` configuration statement.
- The multi-rate gigabit ethernet (mge) interface on EX2300-24MP and EX2300-48MP switches flaps (becomes unavailable, and then available again) while performing timeout detection and recovery (TDR) test.
- If both Energy Efficient Ethernet (EEE) and 100-Mbps speed are configured on a rate-selectable or mge port, the port operates only at 100-Mbps speed but EEE is not enabled on that port. EEE is supported only on mge interfaces that operate at 1-Gbps and 2.5-Gbps speeds.

Configure Speed and Autonegotiation on QFX5100-48T Switches

For information about speed support, see *speed*.

Table 10 on page 94 provides QFX5100-48T details and description.

Table 10: QFX5100-48T Details and Description

Detail	Description
Duplex Mode	Full duplex

Table 10: QFX5100-48T Details and Description *(Continued)*

Detail	Description
Autonegotiation	The autonegotiation option is to negotiate the speeds. 10 Gbps, 1 Gbps, 100 Mbps-By default, autonegotiation is enabled.

Following are guidelines for configuring speed and autonegotiation on QFX5100-48T switch:

- If the speed on the switch is set to 10-Gbps or auto, the switch advertises all the speeds.
- If the speed on the switch is set to 1-Gbps, the switch advertises 1-Gbps and 100-Mbps.
- If you've configured the speed to 100-Mbps on the switch, then you must also configure the speed as 100-Mbps and duplex to full duplex on the link partner.
- If the link partner is set to autonegotiate at 100-Mbps, then you must configure the speed to 10-Gbps, 1-Gbps or auto on QFX5100-48T switch.
- The `no-auto-negotiation` statement does no action. Hence, we recommend not to use the `no-auto-negotiation` statement.

Table 11 on page 95 provides the configuration steps to configure speed and autonegotiation.

Table 11: Configure Speed and Autonegotiation

Configure Speed/ Autonegotiation	Use Configuration

To configure a particular speed, mention the speed.	For a port to only advertise a specific speed, start with a specific speed. It is mandatory that both the autonegotiation option must be set (enabled). The interface must also be configured with a specific supported speed.
	
	`set interfaces xe-0/0/0 ether-options auto-negotiation`
	`set interfaces xe-0/0/0 speed` *speed*
	
	For example to configure 1-Gbps speed, execute the following command:
	
	`set interfaces xe-0/0/0 ether-options auto-negotiation`
	`set interfaces xe-0/0/0 speed 1g`

Table 11: Configure Speed and Autonegotiation *(Continued)*

Configure Speed/ Autonegotiation	Use Configuration

To enable auto-negotiation and advertise all speeds.	With or without below, QFX5100-48T interface support auto-negotiation to one of either 10 GbE and 1 GbE.
	
	`set interfaces xe-0/0/0 ether-options auto-negotiation`
	`set interfaces xe-0/0/0 speed auto`
	
	This configuration does not change any functionality. If the speed is set to fixed at 10 GbE, the interface still operates as auto, and advertises 10 GbE/1 GbE/100 Mbps.
	
	When you configure a port using the `speed auto` option, the port deletes the last configured speed, comes up again and advertises all the possible speeds.

Autonegotiation Support for EX4300-48MP Switches

The 4-port 1-Gigabit Ethernet/10-Gigabit Ethernet uplink module (EX-UM-4SFPP-MR) on EX4300-48MP switches supports 1-Gbps speed. You do not need to explicitly configure 1-Gbps speed on the uplink module as it automatically identifies the installed 1-gigabit SFP transceivers and creates the interface accordingly.

If both Energy Efficient Ethernet (EEE) and 100-Mbps speed are configured on a rate-selectable (mge) port, the port operates only at 100-Mbps speed but EEE is not enabled on that port. Note that EEE is supported only on mge interfaces that operate at 1-Gbps, 2.5-Gbps, 5-Gbps, and 10-Gbps speeds.

On EX4300-48MP, the status LED of 1-Gigabit Ethernet uplink module port is solid green (instead of blinking green) because of a device limitation. However, the limitation does not impact on the device functionality.

Table 12 on page 97 summarizes the autonegotiation and half duplex support on EX4300-48MP switches.

Table 12: Autonegotiation and Half-Duplex Support for EX4300-48MP Switches

Port Numbers	Interface Name	Autonegotiation Supported (YES/NO)?	Half Duplex Supported (YES/NO)?
PIC 0 PORTS 24-47	mge	Yes. Speed supported (10 GbE/5 GbE/2.5 GbE/ 1 GbE/100 Mbps)	No
PIC 2 PORT 0 – 3 (Uplink ports)	xe	No	No
PIC 0 PORTS 0 – 23	GbE	1 GbE/100 Mbps/ 10 Mbps	Yes, but Half duplex cannot be configured on EX4300-48MP switches. If the link partner is half duplex and capable of autonegotiating half duplex, then these ports can work a half duplex.
PIC 2 PORT 0 - 3	Uplink 4x10G	No. Based on inserted transceiver, port is ge for 1 GbE SFP and xe for 10 GbE SFP.	No
PIC 2 PORT 0, 1	Uplink 2x40G	No	No

Autonegotiation Support for EX4400 Switches

Table 13: Autonegotiation and Half-Duplex Support for EX4400 Switches

Interface Name	Autonegotiation Supported (YES/ NO)?	Half Duplex Supported (YES/NO)?
mge	Yes. Speed supported (10 GbE/ 5 GbE/2.5 GbE/1 GbE/100 Mbps)	No

Table 13: Autonegotiation and Half-Duplex Support for EX4400 Switches *(Continued)*

Interface Name	Autonegotiation Supported (YES/NO)?	Half Duplex Supported (YES/NO)?
xe	No	No
GbE	1 GbE/100 Mbps/10 Mbps	Yes, but half duplex cannot be configured on EX4400 switches. If the link partner is half duplex and capable of autonegotiating half duplex, then these ports can work a half duplex.
Uplink 1x100G, 4x10G, and 4x25G	No. Based on inserted transceiver, port is ge for 1 GbE SFP and xe for 10 GbE SFP.	No
2x40G, 2x100G (PIC 1)	No	No

Autonegotiation Support for EX4100 Switches

Table 14: Autonegotiation and Half-Duplex Support for EX4100 Switches

Interface Name	Autonegotiation Supported (YES/NO)?	Half Duplex Supported (YES/NO)?
mge	Yes. Speed supported (10 GbE/ 5 GbE/2.5 GbE/1 GbE/100 Mbps)	No
xe	No	No
ge	1 GbE/100 Mbps/10 Mbps	Yes, half duplex can be configured on EX4100 switches.

Autonegotiation Support for EX4600-40F, QFX5110-48S and QFX5100-48S with JNP-SFPP-10GE-T Transceiver

The interfaces on which the JNP-SFPP-10GE-T transceivers are present come up based on the speed (100 Mbps, 1 Gbps, or 10-Gbps) configured using the `set interfaces` *`interface-name`* `speed` *`speed`* command at the remote end.

For information about platforms support, see Hardware Compatibility Tool.

Table 15 on page 99 discusses EX4600-40F, QFX5110-48S and QFX5100-48S switches with JNP-SFPP-10GE-T transceiver details.

Table 15: QFX5110-48S and QFX5100-48S Switches with JNP-SFPP-10GE-T Transceiver Details and Description

Details	Description
EX4600-40F, QFX5110-48S and QFX5100-48S switches with JNP-SFPP-10GE-T transceiver	Interface created - Gigabit Ethernet (ge) interface. On EX4600-40F, QFX5110-48S and QFX5100-48S, the ge interface supports 100-Mbps1-Gbps, and 10-Gbps speeds, which can be configured using the `speed` configuration statement. Use the `set interfaces ge-0/0/0 speed (100M|1G|10G)` command to configure the speed and autonegotiation.
EX4600-40F, QFX5110-48S and QFX5100-48S switches with other transceivers	Interface created - `ge` or the `xe` interface.
Duplex Mode	Full duplex
Viewing Media Specific Information	You can execute the `show interfaces media` command to view the media-specific information. In the output of `show interfaces` *`name`* `media`, the output field `speed` displays the speed that is configured for the mge interface (with a default of 10GbE). The configured speed signifies the highest speed that the JNP-SFPP-10GE-T is capable of working at. You should enable auto-negotiation for the JNP-SFPP-10GE-T, unless it works in 100-Mbps speed where it can use the parallel detect capability using which it can detect when the link partner is in forced 100BASE-Tx mode and bring the link up. The speed displayed under the `Link partner` denotes the actual speed at which the link is working. The `Link partner` speed is dynamic and displays the highest speed that both ends have negotiated and can work at.

When the interface is configured with a particular speed, it means that the transceiver can support connection to a peer at rates lesser than or equal to the configured speed, as shown in Table 16 on page 100:

Table 16: Interface Speed Based on Configured Speed and Remote End Speed

Configured Speed	The interface will be up with remote end speed of
10 GbE	10G, 1G, and 100M
1 GbE	1G and 100M
100 Mbps	100M

Autonegotiation Support for QFX5120-48Y with JNP-SFPP-10GE-T Transceiver

For information about platforms support, see Hardware Compatibility Tool.

Table 17 on page 100 discusses QFX5120-48Y with JNP-SFPP-10GE-T transceiver details.

Table 17: QFX5120-48Y with JNP-SFPP-10GE-T Transceiver Details and Description

Details	Description
Supported speeds	10 Gbps and 1 Gbps Default speed: 10 Gbps (with or without JNP-SFPP-10GE-T transceiver connected) If the peer does not support 10-Gbps speed, then the link will be down.
Duplex Mode	Full duplex

Table 18 on page 101 configure 1-Gbps and 10-Gbps speeds on QFX5120-48Y with JNP-SFPP-10GE-T transceiver.

Table 18: Configure and Delete 1-Gbps Speed

Configure Speed.	Description
1 Gbps	Use the `set chassis fpc 0 pic 0 port` *`port-number`* `speed` *`1G`* command. Due to hardware limitations, you can configure the *`port-number`* value only in multiples of four, starting from port 0. You must also configure sets of four consecutive ports (for example, 0-3, 4-7, and so on) to operate at the common speed. On QFX5120 switch, mge interfaces are not supported due to hardware limitations.
To revert to 10-Gbps speed after setting 1-Gbps speed.	Delete the `1G` speed configuration.

Operating Speed of Interfaces on EX Switches

SUMMARY

Describes the operating speed of copper and fiber interfaces in both autonegotiation-enabled and autonegotiation-disabled scenarios.

Operating Speed of Copper Interfaces

Execute the `show interfaces` command to retrieve the operating speed of ge/xe/mge copper interfaces from the autonegotiation information.

An example where autonegotiation is enabled:

```
user@host# run show interfaces mge-0/0/4 media

Physical interface: mge-0/0/4, Enabled, Physical link is Up

  Interface index: 654, SNMP ifIndex: 520

  Link-level type: Ethernet, MTU: 1514, LAN-PHY mode, Link-mode: Full-duplex,

  Speed: 2500mbps, BPDU Error: None, Loop Detect PDU Error: None, <== max speed that the port is
```

```
capable of

  Ethernet-Switching Error: None, Remote Bounce: None, MAC-REWRITE Error: None,

  Loopback: Disabled, Source filtering: Disabled, Flow control: Enabled,

  Auto-negotiation: Enabled, Remote fault: Online, Media type: Copper,

  IEEE 802.3az Energy Efficient Ethernet: Disabled, Auto-MDIX: Enabled

  ::

::

  Autonegotiation information:

    Negotiation status: Complete

    Link partner:

        Link mode: Full-duplex, Flow control: Symmetric, Remote fault: OK,

        Link partner Speed: 1000 Mbps

    Local resolution:

        Flow control: Symmetric, Flow control tx: Enabled,

        Flow control rx: Enabled, Remote fault: Link OK,

        Local link Speed: 1000 Mbps, Link mode: Full-duplex <== operating speed
```

If the autonegotiation is disabled for the interface, obtain the operating speed of the linked-up interface from the first stanza of the output.

An example where autonegotiation is disabled:

```
user@host# run show interfaces mge-0/0/0 media

Physical interface: mge-0/0/0, Enabled, Physical link is Up
```

```
Interface index: 650, SNMP ifIndex: 516

Link-level type: Ethernet, MTU: 1514, LAN-PHY mode, Link-mode: Full-duplex,

Speed: 100mbps, BPDU Error: None, Loop Detect PDU Error: None, <==== Operating Speed

Ethernet-Switching Error: None, Remote Bounce: None, MAC-REWRITE Error: None,

Loopback: Disabled, Source filtering: Disabled, Flow control: Enabled,

Auto-negotiation: Disabled, Remote fault: Online, Media type: Copper,

IEEE 802.3az Energy Efficient Ethernet: Disabled, Auto-MDIX: Enabled

Device flags   : Present Running

Interface Specific flags: Internal: 0x80000

Interface flags: SNMP-Traps Internal: 0x4000

Link flags     : None

CoS queues     : 12 supported, 12 maximum usable queues

Current address: 48:5a:0d:ef:31:03, Hardware address: 48:5a:0d:ef:31:03

Last flapped   : 2025-02-13 18:51:43 IST (00:00:09 ago)

Input rate     : 1136 bps (0 pps)

Output rate    : 0 bps (0 pps)

Active alarms  : None

Active defects : None

PCS statistics                      Seconds

  Bit errors                             0

  Errored blocks                         0

MAC statistics:
```

```
    Input bytes: 3544369, Input packets: 41491, Output bytes: 1084277,
    Output packets: 3015
  PRBS Mode : Disabled
  Autonegotiation information:
    Negotiation status: No-autonegotiation
  Interface transmit statistics: Disabled
 Interface transmit statistics: Disabled
```

Operating Speed of Fiber Interfaces

The ge, xe, and et fiber interfaces do not autonegotiate to other speeds. You can obtain the operating speed of the interface from the `show interfaces` command output.

```
user@host> show interfaces et-0/1/0 extensive
Physical interface: et-0/1/0, Enabled, Physical link is Up
  Interface index: 651, SNMP ifIndex: 516, Generation: 142
  Link-level type: Ethernet, MTU: 1514, LAN-PHY mode, Speed: 100Gbps, <=== Operating speed
  BPDU Error: None, Loop Detect PDU Error: None, Ethernet-Switching Error: None,
  Remote Bounce: None, MAC-REWRITE Error: None, Loopback: Disabled,
  Source filtering: Disabled, Flow control: Enabled, Media type: Fiber
```

RELATED DOCUMENTATION

- *speed*
- *auto-negotiation (Switches)*

Configuring Gigabit and 10-Gigabit Ethernet Interfaces for EX4600 and QFX Series Switches

IN THIS SECTION

- Configuring Port Mode on QFX5100-48S, QFX5100-48T, QFX5100-24Q, and EX4600 Switches | **106**
- Configuring the Link Settings for Gigabit Ethernet Interfaces on QFX5100-48S, QFX5100-96S, and EX4600 Switches | **106**
- Configuring Gigabit Ethernet Interfaces on QFX5100-48T Switches | **107**
- Configuring the Link Settings for 10-Gigabit Ethernet Interfaces on QFX5100-48S, QFX5100-24Q, QFX5100-96S, and EX4600 Switches | **108**
- Configuring the Link Settings for 10-Gigabit Ethernet Interfaces on QFX5100-48T Switches | **108**
- Configuring the Link Settings for 10-Gigabit Ethernet Interfaces on QFX5120-48T Switches | **109**
- Configuring the IP Options on QFX5100-48S, QFX5100-48T, QFX5100-24Q, and EX4600 Switches | **109**
- Configuring the Link Settings for 1-Gigabit Ethernet Interfaces on EX4100 EX4100-F Multigigabit Switches | **110**

Devices include a factory default configuration that:

- Enables all 10-Gigabit Ethernet network interfaces on the switch
- Sets a default port mode (access)
- Sets default link settings
- Specifies a logical unit (**unit 0**) and assigns it to **family ethernet-switching**
- Configures Storm Control on all 10-Gigabit Ethernet network interfaces
- Provides basic Rapid Spanning Tree Protocol (RSTP) and Link Layer Discovery Protocol (LLDP) configuration

The `ether-options` statement enables you to modify the following options:

- **802.3ad**—Specify an aggregated Ethernet bundle for both Gigabit Ethernet and 10-Gigabit Ethernet interfaces.
- **autonegotiation**—Enable or disable autonegotation of flow control, link mode, and speed for interfaces.
- **link-mode**—Specify **full-duplex**, **half-duplex**, or **automatic** for Gigabit Ethernet interfaces.
- **loopback**—Enable or disable a loopback interface for both Gigabit Ethernet and 10-Gigabit Ethernet interfaces.

To set **ether-options** for both Gigabit Ethernet and 10-Gigabit Ethernet interfaces:

```
[edit]
user@switch# set interfaces interface-name ether-options
```

This topic describes:

Configuring Port Mode on QFX5100-48S, QFX5100-48T, QFX5100-24Q, and EX4600 Switches

If you are connecting a switch to other switches and to routers on the LAN, you need to assign the interface to a logical port and you need to configure the logical port as a trunk port.

To configure a Gigabit Ethernet or 10-Gigabit interface for trunk port mode on the Enhanced Layer 2 software (ELS):

```
[edit]
user@switch# set interfaces interface-name unit logical-unit-number family ethernet-switching interface-
mode trunk
```

SEE ALSO

Monitoring Interface Status and Traffic | 431

Configuring the Link Settings for Gigabit Ethernet Interfaces on QFX5100-48S, QFX5100-96S, and EX4600 Switches

Devices include a factory default configuration that enables Gigabit Ethernet interfaces with applicable link settings.

The following default configurations are available on Gigabit Ethernet interfaces:

- You cannot set the speed on these interfaces.

  On QFX5100-48S and QFX5100-96S devices using 1-Gigabit Ethernet SFP interfaces, the speed is set to 1 Gbps by default and cannot be configured to operate in a different speed.

- On QFX5100 devices, the interface naming for Gigabit Ethernet interfaces changes automatically to xe-0/0/0, ge-0/0/0, or et-0/0/0 when the appropriate SFP is inserted.

- Gigabit Ethernet interfaces operate in full-duplex mode.

- Autonegotiation is supported by default. Autonegotiation is enabled by default, and will autonegotiate the speed with the link partner. We recommend that you keep autonegotiation enabled for interfaces operating at 100M and 1G. By default, autonegotiation is disabled on 10-Gigabit fiber ports.

  If for some reason you have disabled autonegotiation, you can enable it by issuing the `set interfaces` *`name`* `ether-options auto-negotiate` command.

  To disable autonegotiation, issue the `delete interfaces` *`name`* `ether-options auto-negotiate` command.

  NOTE: Do not use the `set interfaces` *`name`* `ether-options no-auto-negoatiate` command to remove the autonegotiation configuration.

  Issue the `show interfaces` *`name`* `extensive` command to see if autonegotiation is enabled or disabled and the negotiated speed of the interface.

Configuring Gigabit Ethernet Interfaces on QFX5100-48T Switches

Devices include a factory default configuration that enables Gigabit Ethernet interfaces with applicable link settings.

The following default configurations are available on Gigabit Ethernet interfaces:

- Gigabit Ethernet interfaces operate in full-duplex mode.

- Gigabit Ethernet interfaces must be configured as xe-*fpc*/*pic*/*port*, and not ge-*fpc*/*pic*/*port*.

- Autonegotiation is enabled by default, and will autonegotiate the speed with the link partner. You can not disable auto-negotiation.

  Issue the `show interfaces` *`name`* `extensive` command to see the negotiated speed of the interface.

- For a port to start with a specific speed, it is mandatory that both the auto-negotiation must be enabled and interface must be configured with a particular speed. Otherwise, the switch will remain with the last negotiated speed.

Configuring the Link Settings for 10-Gigabit Ethernet Interfaces on QFX5100-48S, QFX5100-24Q, QFX5100-96S, and EX4600 Switches

The following default configurations are available on 10-Gigabit Ethernet interfaces:

- All the 10-Gigabit Ethernet interfaces are set to **auto-negotiation**.
- Flow control for 10-Gigabit Ethernet interfaces is set to **enabled** by default. You can disable flow control by specifying the **no-flow-control** option.
- The speed cannot be configured.

  On QFX5100-48S, QFX5100-96S, and QFX5100-24Q devices using 10-Gigabit Ethernet SFP interfaces, the speed is set to 10 Gbps by default and cannot be configured to operate in a different speed.
- On QFX5100 devices, the interface naming for Gigabit Ethernet interfaces changes automatically to xe-0/0/0, ge-0/0/0, or et-0/0/0 when the appropriate SFP is inserted.
- 10-Gigabit Ethernet interfaces operate in full-duplex mode by default.
- Autonegotiation is enabled by default, and will autonegotiate the speed with the link partner. We recommend that you keep autonegotiation enabled for interfaces operating at 100M and 1G. By default, autonegotiation is disabled on 10-Gigabit fiber ports.

  If for some reason you have disabled autonegotiation, you can enable it by issuing the `set interfaces` *`name`* `ether-options auto-negotiate` command.

  To disable autonegotiation, issue the `delete interfaces` *`name`* `ether-options auto-negotiate` command.

  NOTE: Do not use the `set interfaces` *`name`* `ether-options no-auto-negoatiate` command to remove the autonegotiation configuration.

  Issue the `show interfaces` *`name`* `extensive` command to see if autonegotiation is enabled or disabled and the negotiated speed of the interface.

Configuring the Link Settings for 10-Gigabit Ethernet Interfaces on QFX5100-48T Switches

The following default configurations are available on 10-Gigabit Ethernet interfaces:

- All the 10-Gigabit Ethernet interfaces are set to **auto-negotiation**.
- Flow control for 10-Gigabit Ethernet interfaces is set to **enabled** by default. You can disable flow control by specifying the **no-flow-control** option.

- 10-Gigabit Ethernet interfaces operate in full-duplex mode by default.
- Autonegotiation is enabled by default, and will autonegotiate the speed with the link partner. You can't disable autonegotiation.

  NOTE:
  If you've configured a switch with 100-Mbps speed, then you must also configure the link partner with 100-Mbps speed and duplex to full duplex. If you want to connect to the link partner with 100-Mbps speed that supports autonegotiation at 100-Mbps, we recommend that you configure the speed to 10-Gbps, 1-Gbps, or auto on QFX5100-48T switch.

  Issue the `show interfaces` *`name`* `extensive` command to see if autonegotiation is enabled or disabled and the negotiated speed of the interface.

Configuring the Link Settings for 10-Gigabit Ethernet Interfaces on QFX5120-48T Switches

The following default configurations are available on 10-Gigabit Ethernet interfaces:

- The 10-Gigabit Ethernet interfaces are set to auto-negotiation by default.
- The 10-Gigabit Ethernet interfaces operate in full-duplex mode by default.
- Flow control for the 10-Gigabit Ethernet interfaces is set to enabled by default. You can disable flow control by specifying the no-flow-control option.
- Six 40GbE/100GbE QSFP28 ports support both manual and auto channelization. Auto channelization is enabled by default.

Issue the `show interfaces name extensive` command to see if autonegotiation is enabled or disabled and the negotiated speed of the interface.

Configuring the IP Options on QFX5100-48S, QFX5100-48T, QFX5100-24Q, and EX4600 Switches

To specify an IP address for the logical unit:

```
[edit]
user@switch# set interfaces interface-name unit logical-unit-number family inet address ip-address
```

Configuring the Link Settings for 1-Gigabit Ethernet Interfaces on EX4100 EX4100-F Multigigabit Switches

The following default configurations are available on 1-Gigabit Ethernet interfaces:

- Autonegotiation is enabled by default, and autonegotiates the speed with the link partner.
- 1-Gigabit Ethernet interfaces operate in full-duplex mode by default.
- Flow control for 1-Gigabit Ethernet interfaces is enabled by default. You can disable flow control by specifying the **no-flow-control** option.

  If you configure the half-duplex mode without specifying the **no-flow-control** option,

  the system displays an error message as given in the following example:

```
root@ex4100-device# set interfaces ge-0/0/3 speed 100m

      {master:0}[edit]
      root@ex4100-device# set interfaces ge-0/0/3 link-mode half-duplex

      {master:0}[edit]
      root@ex4100-device# commit
      [edit interfaces]
        'ge-0/0/3'
          Half duplex and flow control enable is an invalid configuration.
      error: configuration check-out failed

      {master:0}[edit]
```

  Add the **no-flow-control** option and the configuration is successful:

```
root@ex4100-device# ... speed 100m ether-options no-flow-control

      {master:0}[edit]
      root@ex4100-device# commit
      configuration check succeeds
      commit complete

      {master:0}[edit]
```

  Issue the `show interfaces` *name* `extensive` command to see if autonegotiation is enabled or disabled and the negotiated speed of the interface.

RELATED DOCUMENTATION

- Monitoring Interface Status and Traffic | 431
- *show interfaces xe*
- *show interfaces ge*
- *speed*

Configuring Gigabit Ethernet Interfaces for EX Series Switches with ELS support

IN THIS SECTION

- Configuring VLAN Options and Interface Mode | 112
- Configuring the Link Settings | 112
- Configuring the IP Options | 115

NOTE: This task uses Junos OS for EX Series switches with support for the Enhanced Layer 2 Software (ELS) configuration style. If your switch runs software that does not support ELS, see Configuring Gigabit Ethernet Interfaces (CLI Procedure). For ELS details, see Using the Enhanced Layer 2 Software CLI.

An Ethernet interface must be configured for optimal performance in a high-traffic network. EX Series switches include a factory default configuration that:

- Enables all the network interfaces on the switch
- Sets a default interface mode (access)
- Sets default link settings
- Specifies a logical unit (`unit 0`) and assigns it to `family ethernet-switching` (except on EX8200 switches and Virtual Chassis)
- Specifies Rapid Spanning Tree Protocol (RSTP) and Link Layer Discovery Protocol (LLDP)

This topic describes:

Configuring VLAN Options and Interface Mode

By default, when you boot a switch and use the factory default configuration, or when you boot the switch and do not explicitly configure a port mode, all interfaces on the switch are in access mode and accept only untagged packets from the VLAN named `default`. You can optionally configure another VLAN and use that instead of `default`. You can also configure a port to accept untagged packets from the user-configured VLAN. For details on this concept (native VLAN), see Understanding Bridging and VLANs on Switches.

If you are connecting either a desktop phone, wireless access point or a security camera to a Power over Ethernet (PoE) port, you can configure some parameters for the PoE interface. PoE interfaces are enabled by default. For detailed information about PoE settings, see *Configuring PoE Interfaces on EX Series Switches*.

If you are connecting a device to other switches and to routers on the LAN, you need to assign the interface to a logical port and configure the logical port as a trunk port. See Port Role Configuration with the J-Web Interface (with CLI References) for more information about port configuration.

If you are connecting to a server that contains virtual machines and a VEPA for packet aggregation from those virtual machines, configure the port as a tagged-access port. See Understanding Bridging and VLANs on Switches for more information about tagged access.

To configure a 1-Gigabit, 10-Gigabit, or 40-Gigabit Ethernet interface for trunk port mode:

```
[edit]
user@switch# set interfaces interface-name unit logical-unit-number family ethernet-switching interface-
mode trunk
```

SEE ALSO

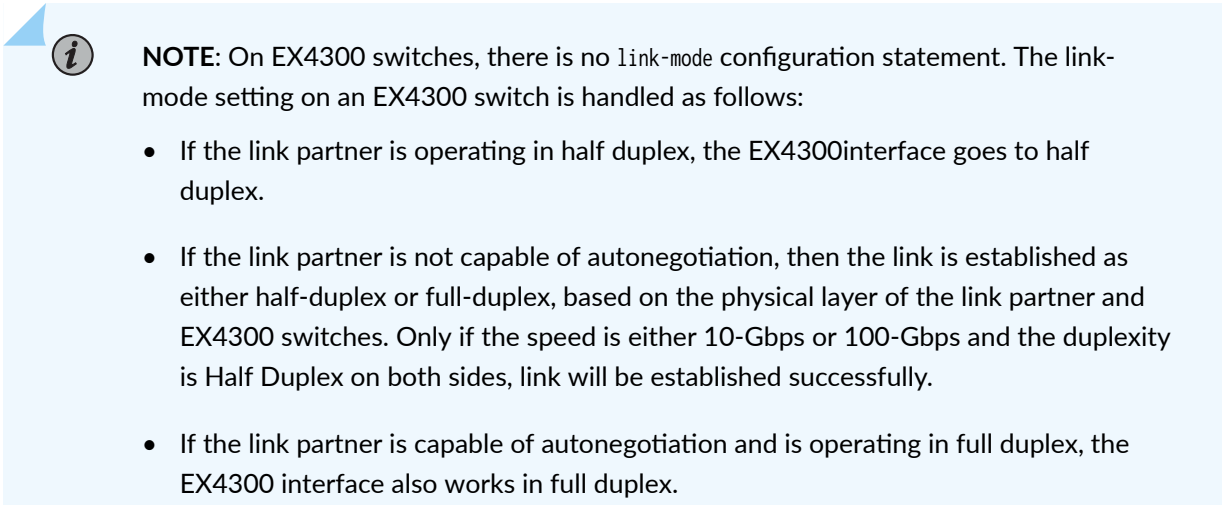
- Monitoring Interface Status and Traffic

Configuring the Link Settings

EX Series switches include a factory default configuration that enables interfaces with the link settings provided in Table 19 on page 113.

Table 19: Factory Default Configuration Link Settings for EX Series Switches

Ethernet Interface	Autonegotiation	Flow Control	Link Mode	Link Speed
1 gigabit	Enabled	Enabled	Autonegotiation (full duplex or half duplex) For information about EX4300, see the Note below this table.	Autonegotiation (10 Mbps, 100 Mbps, or 1 Gbps)
10 gigabit (using a DAC cable)	Enabled	Enabled	Full duplex	10 Gbps
10 gigabit (using a fiber-optic cable)	Disabled	Enabled	Full duplex	10 Gbps
40 gigabit (using a DAC cable)	Enabled	Enabled	Full duplex	40 Gbps
40 gigabit (using a fiber-optic cable)	Disabled	Enabled	Full duplex	40 Gbps



NOTE: On EX4300 switches, there is no `link-mode` configuration statement. The link-mode setting on an EX4300 switch is handled as follows:

- If the link partner is operating in half duplex, the EX4300interface goes to half duplex.
- If the link partner is not capable of autonegotiation, then the link is established as either half-duplex or full-duplex, based on the physical layer of the link partner and EX4300 switches. Only if the speed is either 10-Gbps or 100-Gbps and the duplexity is Half Duplex on both sides, link will be established successfully.
- If the link partner is capable of autonegotiation and is operating in full duplex, the EX4300 interface also works in full duplex.

- To force an EX4300 interface to stay in full-duplex mode, configure the interface's speed as 10 Mbps or 100 Mbps and also configure the interface with the `no-autonegotiation` statement.

To configure the link mode and speed settings for a 1-Gigabit, 10-Gigabit, or 40-Gigabit Ethernet interface:

NOTE: On EX4300 switches, there is no `link-mode` configuration statement. See information earlier in this document regarding how the link mode is set on EX4300 switches.

```
[edit]
user@switch# set interfaces interface-name
```

To configure additional link settings for a 1-Gigabit, 10-Gigabit, or 40-Gigabit Ethernet interface:

```
[edit]
user@switch# set interfaces interface-name ether-options
```

For detailed information about the FPC, PIC, and port numbers used for EX Series switches, see "Understanding Interface Naming Conventions" on page 11.

Configurable link settings include:

- `802.3ad`—Specify an aggregated Ethernet bundle. See Configuring Aggregated Ethernet Links (CLI Procedure).
- `auto-negotiation`—Enable or disable autonegotiation of flow control, link mode, and speed.

  NOTE: Starting with Junos OS Releases 14.1X53-D40, 15.1R4, and 17.1R1, half-duplex communication is supported on all built-in network copper ports on EX4300 switches. *Half-duplex* is bidirectional communication; however, signals can flow in only one direction at a time. *Full-duplex* communication means that both ends of the communication can send and receive signals at the same time.

  Half-duplex is configured by default on EX4300switches. If the link partner is set to autonegotiate the link, then the link is autonegotiated to full duplex or half duplex. If the link is not set to autonegotiation, then the EX4300link defaults to half duplex unless the interface is explicitly configured for full duplex.

To explicitly configure full duplex:

```
[edit]
user@switch# set interfaces interface-name speed 10m-or-100m
[edit]
user@switch# set interfaces interface-name ether-options no-auto-negotiation
```

To verify a half-duplex (or a full-duplex) setting:

```
user@switch> show interfaces interface-name  extensive
```

- `flow-control`—Enable or disable flow control.
- `link-mode`—Specify full duplex, half duplex, or autonegotiation.



NOTE: On EX4300 switches, there is no `link-mode` configuration statement. See information earlier in this document regarding how the link mode is set on EX4300 switches.

- `loopback`—Enable or disable loopback mode.
- `speed`—Specify 10 Mbps, 100 Mbps, 1 Gbps, or autonegotiation.

Configuring the IP Options

To specify an IP address for the logical unit using IPv4:

```
[edit]
user@switch# set interfaces interface-name unit logical-unit-number family inet address ip-address
```

To specify an IP address for the logical unit using IPv6:

```
[edit]
user@switch# set interfaces interface-name unit logical-unit-number family inet6 address ip-address
```



NOTE: Access interfaces on EX4300 switches are set to `family ethernet-switching` by default. You might have to delete this or any other user-configured family setting before changing the setting to `family inet` or `family inet6`.

RELATED DOCUMENTATION

- Configuring Gigabit Ethernet Interfaces (J-Web Procedure)
- Monitoring Interface Status and Traffic
- *show interfaces ge*
- *show interfaces xe*
- Understanding Interface Naming Conventions | **11**

Configuring Gigabit and 10-Gigabit Ethernet Interfaces for OCX Series Switches

IN THIS SECTION

- Configuring the Link Settings for Gigabit Ethernet and 10-Gigabit Ethernet Interfaces | **116**
- Configuring the IP Options | **117**

Devices include a factory default configuration that:

- Enables all 10-Gigabit Ethernet network interfaces on the switch
- Sets default link settings
- Specifies a logical unit (**unit 0**) and assigns it to **family ethernet-switching**
- Configures Storm Control on all 10-Gigabit Ethernet network interfaces

This topic describes:

Configuring the Link Settings for Gigabit Ethernet and 10-Gigabit Ethernet Interfaces

Devices include a factory default configuration that enables 10-Gigabit Ethernet and interfaces with applicable link settings.

The following default configurations are available on 10-Gigabit Ethernet interfaces:

- The speed for 10-Gigabit Ethernet interfaces is set to 10 Gbps by default. The speed cannot be configured.
- 10-Gigabit Ethernet interfaces operate in full-duplex mode by default.

- Autonegotiation is not supported.

The `ether-options` statement enables you to modify the following options:

- **802.3ad**—Specify an aggregated Ethernet bundle for 10-Gigabit Ethernet interfaces.
- **loopback**—Enable or disable a loopback interface for 10-Gigabit Ethernet interfaces.

To set **ether-options** for 10-Gigabit Ethernet interfaces:

```
[edit]
user@switch# set interfaces interface-name ether-options
```

Configuring the IP Options

To specify an IP address for the logical unit:

```
[edit]
user@switch# set interfaces interface-name unit logical-unit-number family inet address ip-address
```

RELATED DOCUMENTATION

Monitoring Interface Status and Traffic

show interfaces xe

show interfaces ge

speed

Change History Table

Feature support is determined by the platform and release you are using. Use Feature Explorer to determine if a feature is supported on your platform.

Release	Description
19.1R1	The 4-port 1-Gigabit Ethernet/10-Gigabit Ethernet uplink module (EX-UM-4SFPP-MR) on EX4300-48MP switches supports 1-Gbps speed.
18.2R1	The mge interface is a rate-selectable (multirate gigabit ethernet) interface that support speeds of 10-Gbps, 5-Gbps, and 2.5-Gbps over CAT5e/CAT6/CAT6a cables. In the EX2300, the mge interface supports 100-Mbps, 1-Gbps, and 2.5-Gbps speeds. Note that 10 Mbps speed is supported only on `ge` interfaces of EX2300 switch.

14.1X53-40	Starting with Junos OS Releases 14.1X53-D40, 15.1R4, and 17.1R1, half-duplex communication is supported on all built-in network copper ports on EX4300 switches.

Optical Transport Network (OTN) Interfaces

IN THIS SECTION

- Understanding the QFX10K-12C-DWDM Line Card | **118**
- Configuring OTN Interface Options on QFX10K-12C-DWDM | **121**
- Support for 400G-ZR Optics on QFX5220-32CD and QFX5130 | **127**

The QFX10K-12C-DWDM line card supports the optical transport interfaces (OTN) which is used for high end packet forwarding by cloud providers, service providers and enterprises. There are various optic-specific options that can be configured on the QFX10K-12C-DWDM line card including the forward error correction (FEC) mode and enabling the threshold crossing alarms.

Understanding the QFX10K-12C-DWDM Line Card

IN THIS SECTION

- Software Features | **119**
- OTN Alarms and Defects | **120**

The QFX10000-12C-DWDM line card provides up to 1.2 Tbps packet forwarding for cloud providers, service providers, and enterprises that need coherent dense wavelength-division multiplexing (DWDM) with MACsec security features.

The QFX10K-12C-DWDM line card is supported on Junos OS Release 17.2R1 and later.

The following sections explain the features of the QFX10K-12C-DWDM line card in detail:

Software Features

The following interface features are supported on the QFX10000-12C-DWDM:

- Compliant with ITU G.709 and G.798
- Performance monitoring features such as alarms, threshold-crossing alarms, OTU/ODU error seconds, and FEC and bit error rate (BER) statistics.
- SNMP management of the MIC based on RFC 3591, Managed Objects for the Optical Interface Type, including the following:
  - Black Link MIB–jnx-bl.mib
  - IFOTN MIB–jnx-ifotn.mib
  - Optics MIB–jnx-optics.mib
  - FRU MIB–jnx-fru.mib
- User-configurable optics options:
  - Modulation format: 16QAM, 8QAM, QPSK
  - FEC mode (15% SDFEC or 25% SDFEC)
  - Differential and non-differential encoding modes
  - Transmit (TX) laser enable and disable
  - TX output power
  - Wavelength
  - Threshold crossing alarms (TCAs)
- IEEE 802.1ag OAM
- IEEE 802.3ah OAM
- IFINFO/IFMON
- IEEE 802.3ad link aggregation
- Flexible Ethernet services encapsulation
- Flexible VLAN tagging
- Source address MAC accounting per logical interface

- Source address MAC filter per port
- Source address MAC filter per logical interface
- Destination address MAC filter per port
- Up to 8000 logical interfaces shared across all ports on a single PFE

OTN Alarms and Defects

The following OTN alarms and defects are supported on the QFX10K-12C-DWDM line card:

Optical Channel(OC) Alarms and Defects

- OC-LOS—Loss Of Signal
- OC-LOF—Loss Of Frame
- OC-LOM—Loss Of Multiframe
- OC-Wavelength-Lock—Wavelength Lock

Optical Channel Data Unit (ODU) Defects

- ODU-AIS—ODU Alarm Indication Signal
- ODU-BDI—ODU Backward Defect Indication
- ODU-IAE—ODU Incoming Alignment Error
- ODU-LCK—ODU Locked
- ODU-LTC—ODU Loss of Tandem Connection
- ODU-OCI—ODU Open Connection Error
- ODU-SSF—ODU Server Signal Failure
- ODU-TSF—ODU Trail Signal Failure
- ODU-TTIM—ODU Trail Trace Identifier Mismatch

Optical Channel Transport Unit (OTU) Defects

- OTU-AIS—OTU Alarm Indication Signal
- OTU-BDI—OTU Backward Defect Indication
- OTU-BIAE—OTU Backward Incoming Alignment Error

- OTU-FEC-DEG—OTU Forward Error Correction Degrade
- OTU-FEC-EXCESS-FEC—OTU Forward Error Correction Excessive FEC Errors
- OTU-IAE—OTU Incoming Alignment Error
- OTU-SSF—OTU Server Signal Failure
- OTU-TSF—OTU Trail Signal Failure
- OTU-TTIM—OTU Trail Trace Identifier Mismatch

Threshold-Crossing Alarms

Threshold-crossing alarms (TCA) are alarms that are activated when a certain configurable threshold — near-end measurement threshold or far-end measurement threshold—is crossed and remains so until the end of the 15 minutes interval for parameters such as OTU and ODU. The following alarms are supported:

- Background block error threshold (BBE)
- Errored seconds threshold (ES)
- Severely errored seconds threshold (SES)
- Unavailable seconds threshold (UES)

Configuring OTN Interface Options on QFX10K-12C-DWDM

The QFX10000-12C-DWDM line card provides up to 1.2 Tbps packet forwarding for cloud providers, service providers, and enterprises that need coherent dense wavelength-division multiplexing (DWDM) with MACsec security features. The QFX10K-12C-DWDM line card is supported on Junos OS Release 17.2R1 and later.

Each QFX10K-12C-DWDM has 6 physical interfaces (ot-x/x/x) that connect to one of three built-in flexible rate optical transponders. Each transponder connects four 100-Gigabit Ethernet logical interfaces (et-x/x/x) to one of three forwarding ASICs.

To configure the optics-specific options on the interface:

1. Specify the modulation format at the [`edit interface` *`interface-name`* `optics-options`] hierarchy level.

```
[edit interfaces interface-name optics-options]
user@host# set modulation-format (qpsk|8qam|16qam)
```

2. Specify encoding.

```
[edit interfaces interface-name optics-options]
user@host# set encoding (differential|non-differential)
```

3. Specify the optical transmit laser output power in dBm. The default transmit laser output value is 0 dBm.

```
[edit interfaces interface-name optics-options]
user@host# set tx-power value
```

4. Specify the wavelength of the optics in nanometers. For a list of wavelengths supported, see *wavelength*.

```
[edit interfaces interface-name optics-options]
user@host# set wavelength nm
```

To configure the OTN-specific options on the interface:

1. At the [edit interfaces *interface-name* otn-options] enable the laser on the OTN interface. The laser is disabled by default for all OTN interfaces.

```
[edit interfaces interface-name otn-options]
user@host# set laser-enable
```

2. Set an trail trace identifier for the source access point and for the destination access point for ODU and OTU on the OTN interface.

```
[edit interfaces interface-name otn-options]
user@host# set tti (odu-dapi | odu-expected-receive-dapi | odu-expected-receive-sapi | odu-
sapi | otu-dapi | otu-expected-receive-dapi | otu-expected-receive-sapi | otu-sapi)
```

3. By default, triggers are ignored. Specify defect triggers and the set the trigger hold time for the trigger. Possible values for the trigger hold time are as follows: down—Delay before marking interface down when defect occurs (1..65534 milliseconds) and up—Delay before marking interface up when defect is absent (1..65534 milliseconds).



NOTE: The hold time value only impacts the alarm reporting time and does not mark an interface down when the defect occurs. To mark the interface up or down, you must also configure the physical interface hold time at the [`edit interfaces` *interface-name*] hierarchy level.

```
[edit interfaces interface-name otn-options]
user@host# set trigger (oc-lof | oc-lom | oc-los | oc-tsf | odu-ais | odu-bdi | odu-bei |
odu-iae | odu-lck | odu-oci | odu-sd | odu-ttim |opu-ptim | otu-ais | otu-bdi | otu-fec-deg
| otu-fec-exe | otu-iae | otu-sd | otu-ttim) (hold-time (down value | up value) | ignore)
```

4. Enable the threshold crossing alarms for the OTN interface along with the trigger for the defect.

```
[edit interfaces interface-name otn-options]
user@host# set tca (odu-tca-bbe | odu-tca-es | odu-tca-ses | odu-tca-uas | otu-tca-bbe |
otu-tca-es | otu-tca-ses | otu-tca-uas ) (enable-tca | no-enable-tca | threshold)
```

5. Set the OTN header bytes as a transmit payload type from 0 bytes through 255 bytes for the packets that are transmitted on the OTN interface.

```
[edit interfaces interface-name otn-options]
user@host# set bytes transmit-payload-type value
```

6. Configure the forward error correction (FEC) mode for the OTN interface. Possible values are: Generic Forward Error Correction (GFEC), or High Gain Forward Error Correction (HGFEC) or Soft Decision Forward Error Correction (SDFEC). The default forward error correction mode is SDFEC.

```
[edit interfaces interface-name otn-options]
user@host# set fec (gfec | hgfec | sdfec)
```

7. Enable line loopback or local host loopback for the OTN interface.

```
[edit interfaces interface-name otn-options]
user@host# set line-loopback
user@host# set local-loopback
```

8. Enable an ODU locked maintenance signal on the OTN interface to send the signal pattern 01010101.

```
[edit interfaces interface-name otn-options]
user@host# set insert-odu-lck
```

9. Enable an ODU open connection indication signal on the OTN interface to send to send the signal pattern 01100110.

```
[edit interfaces interface-name otn-options]
user@host# set insert-odu-oci
```

10. Enable a consequent action as listed in the ITU-T G.798 standard for ODU trail trace identifier mismatch (TTIM) on the OTN interface.

```
[edit interfaces interface-name otn-options]
user@host# set odu-ttim-action-enable
```

11. Enable a consequent action as listed in the ITU-T G.798 standard for OTU trail trace identifier mismatch (TTIM) on the OTN interface.

```
[edit interfaces interface-name otn-options]
user@host# set out-ttim-action-enable
```

12. Configure the OTN payload pseudorandom binary sequence (PRBS) on the OTN interface.

```
[edit interfaces interface-name otn-options]
user@host# set prbs
```

13. Configure the line rate or speed of the OTN signal to OTU4 (100Gbps) for the OTN interface.



NOTE: If you specify a value other than OTU4, the value is ignored. To verify the line rate, use the `show interfaces` *interface-name* `extensive` command.

```
[edit interfaces interface-name otn-options]
user@host# set rate otu4
```

14. Configure the threshold value for signal degradation when an alarm needs to be raised. Configure the threshold value after signal degradation when the alarm needs to be cleared. When you configure the interval along with the `ber-threshold-signal-degrade` *value* statement, the bit error rate (BER) must stay above the signal degradation threshold for the configured interval after which the alarm is raised. When the interval is configured along with the `ber-threshold-clear` *value* statement, then BER must stay below the clear threshold for the configured interval after which the alarm is cleared.

```
[edit interfaces interface-name otn-options signal-degrade]
user@host# set ber-threshold-signal-degrade value
user@host# set ber-threshold-clear value
user@host# set interval value
```

15. Enable the following actions for the preemptive-fast-reroute statement:

    - Backward FRR—Insert the local pre-FEC status into the transmitted OTN frames and monitor the received OTN frames for the pre-FEC status.

      ```
      [edit interfaces interface-name otn-options preemptive-fast-reroute]
      user@host# set backward-frr-enable
      ```

    - ODU backward FRR—Insert the ODU status into the transmitted OTN frames and monitor the received OTN frames for the ODU BER status.

      ```
      [edit interfaces interface-name otn-options preemptive-fast-reroute]
      user@host# set odu-backward-frr-enable
      ```

- Monitoring of signal degradation of pre-FEC OTN frames.

```
[edit interfaces interface-name otn-options preemptive-fast-reroute]
user@host# set signal-degrade-monitor-enable
```

- Monitoring of signal degradation of ODU BER in the received OTN frames.

```
[edit interfaces interface-name otn-options preemptive-fast-reroute]
user@host# set odu-signal-degrade-monitor-enable
```

16. Configure the following options for ODU BER signal degradation on the OTN interface:

- Configure the threshold for signal degradation for ODU BER when an alarm needs to be raised.

```
[edit interfaces interface-name otn-options odu-signal-degrade]
user@host# set ber-threshold-signal-degrade value
```

- Configure the threshold for ODU BER after signal degradation when the alarm needs to be cleared.

```
[edit interfaces interface-name otn-options odu-signal-degrade]
user@host# set ber-threshold-clear value
```

- When you configure the interval along with the `ber-threshold-signal-degrade value` statement, the ODU bit error rate (BER) must stay above the signal degradation threshold for the configured interval after which the alarm is raised. When the interval is configured along with the `ber-threshold-clear value` statement, then ODU BER must stay below the clear threshold for the configured interval after which the alarm is cleared.

```
[edit interfaces interface-name otn-options odu-signal-degrade]
user@host# set interval value
```

SEE ALSO

optics-options

otn-options

signal-degrade

preemptive-fast-reroute

Support for 400G-ZR Optics on QFX5220-32CD and QFX5130

IN THIS SECTION

- Benefit of 400G-ZR Optics Support | **127**

400-ZR is a standard for transporting 400Gb Ethernet. The standard aims at a minimum distance of 80 kilometers and implemented on small, pluggable form factor modules such as QSFP-DD.

Some of the applications that use 400-ZR optics fiber are the following:

- Data Center Interconnectivity (DCI) links
- Campus DWDM
- Metro DWDM

Benefit of 400G-ZR Optics Support

- Low latency and high speed

The following are the guidelines when you configure the 400-ZR optics:

The number of ports supporting the 400G-ZR optics is restricted based on the power budget on the QFX5220-32CD and QFX5130-32CD devices. For better thermal handling and power consumption, 16 ports (0, 3, 4, 7, 8, 11, 12, 15, 16, 19, 20, 23, 24, 27, 28, 31) in zigzag pattern support the 400G-ZR optics. Each port supporting the 400G-ZR optic is mapped to another port. You must configure the mapped port to “unused”. You must also configure the supported ports with a high-power mode to power on the optics module.

For example:

Use the following command to set the corresponding port (port 1) to unused, if the 400G-ZR optic module is connected to port 0:

- For QFX5220: `set chassis fpc 0 pic 0 port 1 unused`

- For QFX5130:`set interfaces et-0/0/1 unused`

These commands power on the port 0.

Use the following commands if the 400G-ZR optics module is connected to port 0:

- For QFX5220: `set interfaces et-0/0/0 optics-options high-power-mode`
- For QFX5130: `set interfaces et-0/0/0 optics-options high-power-mode`

The following table shows the supported ports and corresponding unused ports:

Ports supporting 400G-ZR optic	Corresponding ports to be set unused
0	1
3	2
4	5
7	6
8	9
11	10
12	13
15	14
16	17
19	18
20	21
23	22
24	25
27	26
28	29

(Continued)

Ports supporting 400G-ZR optic	Corresponding ports to be set unused
31	30

- If the 400G-ZR optics is used in channelized mode (4x100G), the high-power mode configuration needs to be present on channel 0 (for both QFX5130-32CD and QFX5220-32CD).

  ```
  set interfaces et-0/0/0:0 optics-options high-power-mode
  ```

- If the 400G-ZR optics module is inserted in an unsupported port, the module is not powered on.

  The following alarm is raised on the port:

  ```
  High power optics can not be supported on the port
  ```

- The following alarm is raised if the 400G-ZR optics module is plugged into the supported port, but high-power mode configuration is not configured.

  ```
  optics-options high-power-mode config needed to support high power optics on the port
  ```

- If none of the ports have a 400-ZR optics module, high-power mode and unused port settings are not required.

Energy Efficient Ethernet Interfaces

IN THIS SECTION

- Reduce Power Consumption on Interfaces using Energy Efficient Ethernet | **130**
- Configure Energy Efficient Ethernet on Interfaces | **130**
- Verify EEE-Enabled Ports | **131**

The energy efficient ethernet (EEE) helps in reducing the power consumption on physical layer devices. Configuring these EEE on interfaces includes enabling EEE on Base-T copper ethernet port based on the power utilization and also verifying if EEE is saving energy on the configured ports.

Reduce Power Consumption on Interfaces using Energy Efficient Ethernet

Energy Efficient Ethernet (EEE), an Institute of Electrical and Electronics Engineers (IEEE) 802.3az standard, reduces the power consumption of physical layer devices (PHYs) during periods of low link utilization. EEE saves energy by switching part of the transmission circuit into low power mode when the link is idle.

An Ethernet link consumes power even when a link is idle. EEE provides a method to utilize power in such a way that Ethernet links use power only during data transmission. EEE uses a signaling protocol, Low Power Idle (LPI) for achieving the power saving when an Ethernet link is idle. EEE allows PHYs to exchange LPI indications to signal the transition to low power mode when there is no traffic. LPI indicates when a link can go idle and when the link needs to resume after a predefined delay without impacting data transmission.

The following copper PHYs are standardized by IEEE 802.3az:

- 100BASE-T
- 1000BASE-T
- 10GBASE-T

Configure Energy Efficient Ethernet on Interfaces

IN THIS SECTION

- Enable EEE on an EEE-Capable Base-T Copper Ethernet Port | **131**
- Disable EEE on a Base-T Copper Ethernet Port | **131**

Configure EEE only on EEE-capable Base-T copper Ethernet ports. If you configure EEE on unsupported ports, the console displays the message: **“EEE not supported”**.

This topic describes:

Enable EEE on an EEE-Capable Base-T Copper Ethernet Port

To enable EEE on an EEE-capable Base-T copper Ethernet interface:

```
[edit]
 user@switch# set interfaces interface-name  ether-options ieee-802-3az-eee
```

You can view the EEE status by using the `show interfaces interface-name detail` command.

Disable EEE on a Base-T Copper Ethernet Port

To disable EEE on a Base-T copper Ethernet interface:

```
[edit]
 user@switch# delete interfaces interface-name  ether-options ieee-802-3az-eee
```

By default, EEE is disabled on EEE-capable ports.

Verify EEE-Enabled Ports

IN THIS SECTION

- Purpose | 131
- Action | 131
- Meaning | 134

Purpose

Verify that enabling EEE saves energy on Base-T Copper Ethernet ports.

Action

You can see the amount of energy that is saved by EEE on an EX Series Switches using the `show chassis power-budget-statistics` command.

1. View the power budget of an EX Series Switches before enabling EEE.

- On an EX6210 switch:

```
user@switch>show chassis power-budget-statistics
     PSU  2     (EX6200-PWR-AC2500)                :    2500 W   Online
            PSU  3                 )                :       0 W  Offline
            Total Power supplied by all Online PSUs  :    2500 W
            Power Redundancy Configuration           :     N+1
            Power Reserved for the Chassis           :     500 W
    Fan Tray Statistics                        Base power   Power Used
    FTC  0                                    :     300 W        nan W
    FPC Statistics                             Base power   Power Used   PoE power
Priority
    FPC  3   (EX6200-48T)                     :     150 W      61.54 W          0 W       9
    FPC  4   (EX6200-SRE64-4XS)               :     100 W      48.25 W          0 W       0
    FPC  5   (EX6200-SRE64-4XS)               :     100 W      48.00 W          0 W       0
    FPC  7   (EX6200-48T)                     :     150 W      63.11 W          0 W       9
    FPC  8   (EX6200-48T)                     :     150 W      12.17 W          0 W       9

    Total (non-PoE) Power allocated           :     950 W
            Total Power allocated for PoE            :       0 W
            Power Available (Redundant case)         :       0 W
            Total Power Available                    :    1550 W
```

- On an EX4300 switch:

```
user@switch>show chassis power-budget-statistics fpc 1
     PSU  1    (JPSU-1100-AC-AFO-A)           :    1100 W   Online
            Power redundancy configuration           :    N+0
            Total power supplied by all online PSUs  :    1100 W
            Base power reserved                      :    175 W
            Non-PoE power being consumed             :    95 W
            Total Power allocated for PoE            :    925 W
            Total PoE power consumed                 :    0 W
            Total PoE power remaining                :    925 W
```

2. Enable EEE on Base-T Copper Ethernet ports and save the configuration.

3. View the power budget of the switch after enabling EEE.

- On an EX6210 switch:

```
user@switch> show chassis power-budget-statistics
     PSU  2     (EX6200-PWR-AC2500)                 :    2500 W   Online
            PSU  3                 )                 :       0 W  Offline
            Total Power supplied by all Online PSUs  :    2500 W
            Power Redundancy Configuration           :     N+1
            Power Reserved for the Chassis           :     500 W
    Fan Tray Statistics                        Base power   Power Used
    FTC  0                                    :     300 W        nan W
    FPC Statistics                             Base power   Power Used   PoE power
Priority
    FPC  3   (EX6200-48T)                     :     150 W      50.36 W          0 W       9
    FPC  4   (EX6200-SRE64-4XS)               :     100 W      48.60 W          0 W       0
    FPC  5   (EX6200-SRE64-4XS)               :     100 W      48.09 W          0 W       0
    FPC  7   (EX6200-48T)                     :     150 W      51.38 W          0 W       9
    FPC  8   (EX6200-48T)                     :     150 W      12.17 W          0 W       9

    Total (non-PoE) Power allocated           :     950 W
            Total Power allocated for PoE              :       0 W
            Power Available (Redundant case)           :       0 W
            Total Power Available                      :    1550 W
```

- On an EX4300 switch:

```
user@switch> show chassis power-budget-statistics fpc 1
     PSU  1    (JPSU-1100-AC-AFO-A)           :    1100 W   Online
            Power redundancy configuration           :    N+0
            Total power supplied by all online PSUs  :    1100 W
            Base power reserved                      :    175 W
            Non-PoE power being consumed             :    86 W
            Total Power allocated for PoE            :    925 W
            Total PoE power consumed                 :    0 W
            Total PoE power remaining                :    925 W
```

4. See show interfaces extensive for EEE-enabled status of the interface for Low Power Idle (LPI).

Meaning

On an EX6210 switch, the `Power Used` field in the output shows the actual power being consumed by the line card or SRE module, including PoE power. If you compare the values in the `Power Used` field before and after enabling EEE for FPC 3 and FPC 7, you notice the saved power when you enable EEE.



NOTE: Only for EX6210 switches, the output displays the `Power Used` field.

On an EX4300 switch, if you compare the values in the `Non-PoE power being consumed` field before and after enabling EEE, you notice the saved power when you enable EEE.

Uplink Failure Detection

IN THIS SECTION

- Overview of Uplink Failure Detection | 134
- Configuring Interfaces for Uplink Failure Detection | 137
- Example: Configuring Interfaces for Uplink Failure Detection | 139
- Verifying That Uplink Failure Detection Is Working Correctly | 145

Uplink failure detection detects the failure on uplink interfaces and advertises this information to the downlink interfaces so that the switch over of interfaces is possible to avoid loss of traffic. The topics below discuss the functions of uplink failure detections and the steps to configure and verify the working of it.

Overview of Uplink Failure Detection

IN THIS SECTION

- Uplink Failure Detection Configuration | 135

Failure Detection Pair | 136
Debounce Interval | 137

Uplink failure detection allows a switch to detect link failure on uplink interfaces and to propagate this information to the downlink interfaces, so that servers connected to those downlinks can switch over to secondary interfaces.

Uplink failure detection supports network adapter teaming and provides network redundancy. In network adapter teaming, all of the network interface cards (NICs) on a server are configured in a primary or secondary relationship and share the same IP address. When the primary link goes down, the server transparently shifts the connection to the secondary link. With uplink failure detection, the switch monitors uplink interfaces for link failures. When it detects a failure, it disables the downlink interfaces. When the server detects disabled downlink interfaces, it switches over to the secondary link to help ensure that the traffic of the failed link is not dropped.

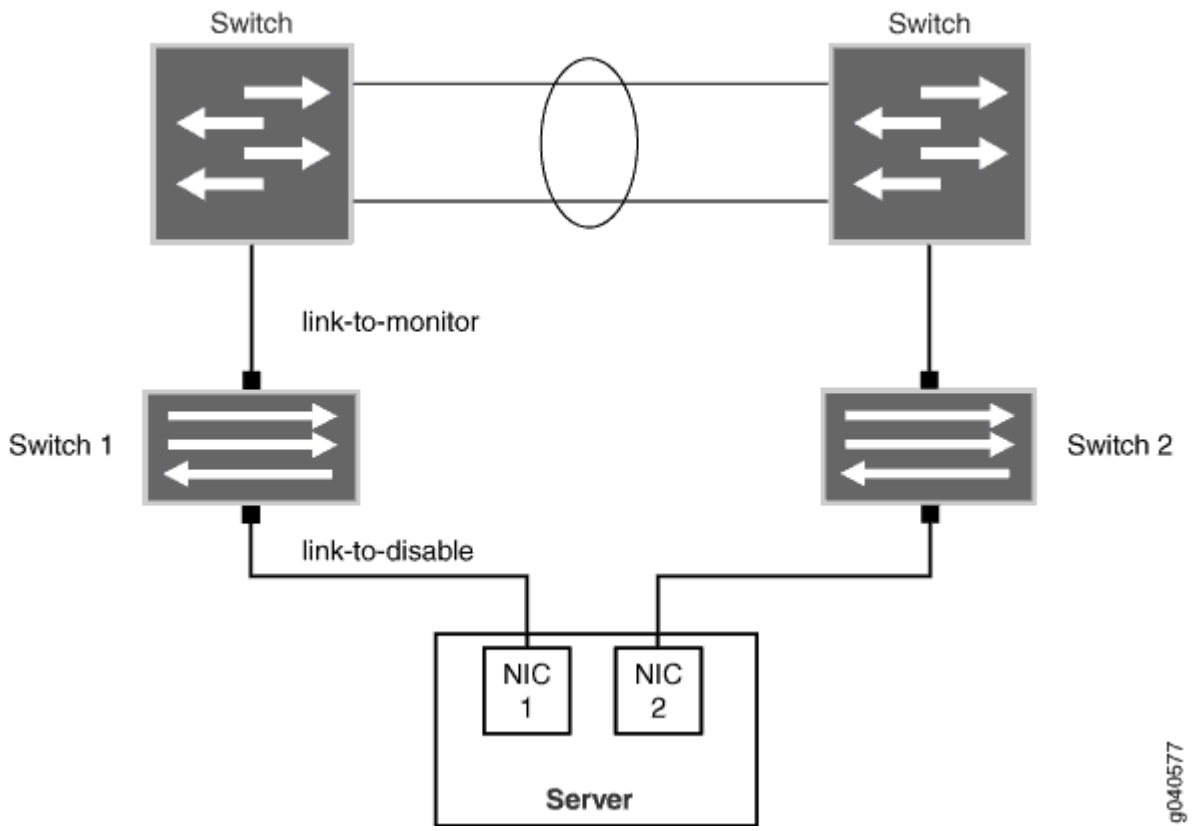
This topic describes:

Uplink Failure Detection Configuration

Uplink failure detection allows switches to monitor uplink interfaces to spot link failures. When a switch detects a link failure, it automatically disables the downlink interfaces bound to the uplink interface. A server that is connected to the disabled downlink interface triggers a network adapter failover to a secondary link to avoid any traffic loss.

Figure 6 on page 136 illustrates a typical setup for uplink failure detection.

Figure 6: Uplink Failure Detection Configuration on Switches



For uplink failure detection, you specify a group of uplink interfaces to be monitored and downlink interfaces to be brought down when an uplink fails. The downlink interfaces are bound to the uplink interfaces within the group. If all uplink interfaces in a group go down, then the switch brings down all downlink interfaces within that group. If any uplink interface returns to service, then the switch brings all downlink interfaces in that group back to service.

The switch can monitor both physical interface links and *logical interface* links for uplink failures, but you must put the two types of interfaces into separate groups.

NOTE: For logical interfaces, the server must send keepalives between the switch and the server to detect failure of logical links.

Failure Detection Pair

Uplink failure detection requires that you create pairs of uplink and downlink interfaces in a group. Each pair includes one of each of the following:

- A link-to-monitor interface—The link-to-monitor interfaces specify the uplinks the switch monitors. You can configure a maximum of 48 uplink interfaces as link-to-monitor interfaces for a group.
- A link-to-disable interface—The link-to-disable interfaces specify the downlinks the switch disables when the switch detects an uplink failure. You can configure a maximum of 48 downlinks to disable in the group.

The link-to-disable interfaces are bound to the link-to-monitor interfaces within the group. When a link-to-monitor interface returns to service, the switch automatically enables all link-to-disable interfaces in the group.

Debounce Interval

The debounce interval is the amount of time, in seconds, that elapses before the downlink interfaces are brought up after corresponding state changes of the uplink interfaces. You can configure the debounce interval for the uplink failure detection group. In absence of the debounce interval configuration, the downlink interfaces are brought up immediately after a state change of the uplink interfaces, which might introduce unnecessary state changes of the downlink interfaces, as well as unnecessary failovers on the servers connected to these ports.

In the event that the uplink interface goes down during the debounce interval, the debounce timer will start when the uplink interface comes back up. If the uplink interface goes down before the debounce interval expires, the debounce timer restarts when the uplink interface comes back up.

Any change you make to the debounce interval takes effect immediately. If you make a change to the debounce interval while the debounce timer is in effect, the change will take place if the new expiry time is in the future. If not, the timer stops immediately.

If uplink failure detection restarts during the debounce interval, the debounce timer resets, and the time that elapsed before uplink failure detection restarted is lost. The link-to-disable interface comes up without waiting for the debounce interval to elapse.

If the link-to-disable interface does not come up after the debounce timer expires, there might be latency between the time the debounce timer expires and the time when the link-to-disable interface actually comes up.

Configuring Interfaces for Uplink Failure Detection

You can configure uplink failure detection to help ensure balanced traffic flow. Using this feature, switches can monitor and detect link failure on uplink interfaces and can propagate the failure information to downlink interfaces, so that servers connected to those downlinks can switch over to secondary interfaces.

Follow these configuration guidelines:

- Configure an interface in only one group.
- Configure a maximum of 48groups for each switch.
- Configure a maximum of 48 uplinks to monitor and a maximum of 48 downlinks to disable in each group.
- Configure physical links and logical links in separate groups.

To configure uplink failure detection on a switch:

1. Specify a name for an uplink failure detection group:

```
[edit protocols]
user@switch# set uplink-failure-detection group group-name
```

2. Add an uplink interface to the group:

```
[edit protocols]
user@switch# set uplink-failure-detection group group-name link-to-monitor interface-name
```

3. Configure the debounce interval for the group:

```
[edit protocols]
user@switch# set uplink-failure-detection group group1 debounce-interval seconds
```

4. Repeat Step 2 for each uplink interface you add to the group.
5. Add a downlink interface to the group:

```
[edit protocols]
user@switch# set uplink-failure-detection group group-name link-to-disable interface-name
```

6. Repeat Step 4 for each downlink interface you add to the group.



NOTE: After you have configured an uplink failure detection group, use the **show uplink-failure-detection group (Uplink Failure Detection)** ***group-name*** command to verify that all interfaces in the group are up. If the interfaces are down, uplink failure detection does not work.

Example: Configuring Interfaces for Uplink Failure Detection

IN THIS SECTION

- Requirements | 139
- Overview and Topology | 139
- Configuring Uplink Failure Detection on Both Switches | 141
- Verification | 144

Uplink failure detection allows a switch to detect link failure on uplink interfaces and to propagate the failure information to the downlink interfaces. All of the network interface cards (NICs) on a server are configured as being either the primary link or the secondary link and share the same IP address. When the primary link goes down, the server transparently shifts the connection to the secondary link to ensure that the traffic on the failed link is not dropped.

This example describes:

Requirements

This example uses the following software and hardware components:

- Junos OS Release 19.2R1 or later for the QFX Series
- Two QFX5100, QFX5110, QFX5120, QFX5200, or QFX5210 switches
- Two aggregation switches
- One dual-homed server

Overview and Topology

IN THIS SECTION

- Topology | 141

The topology in this example illustrates how to configure uplink failure detection on Switch 1 and Switch B. Switch 1 and Switch 2 are both configured with a link-to-monitor interface (the uplink interface to the

aggregation switch) and a link-to-disable interface (the downlink interface to the server). For simplicity, only one group of link-to-monitor interfaces and link-to-disable interfaces is configured for each switch. The server is dual-homed to both Switch 1 and Switch 2. In this scenario, if the link-to-monitor interface to Switch 1 is disabled, the server uses the link-to-monitor interface to Switch 2 instead.

NOTE: This example does not describe how to configure the dual-homed server or the aggregation switches. Please refer to the documentation for each of these devices for more information.

Figure 7 on page 140 illustrates a typical setup for uplink failure detection.

Figure 7: Uplink Failure Detection Configuration on Switches

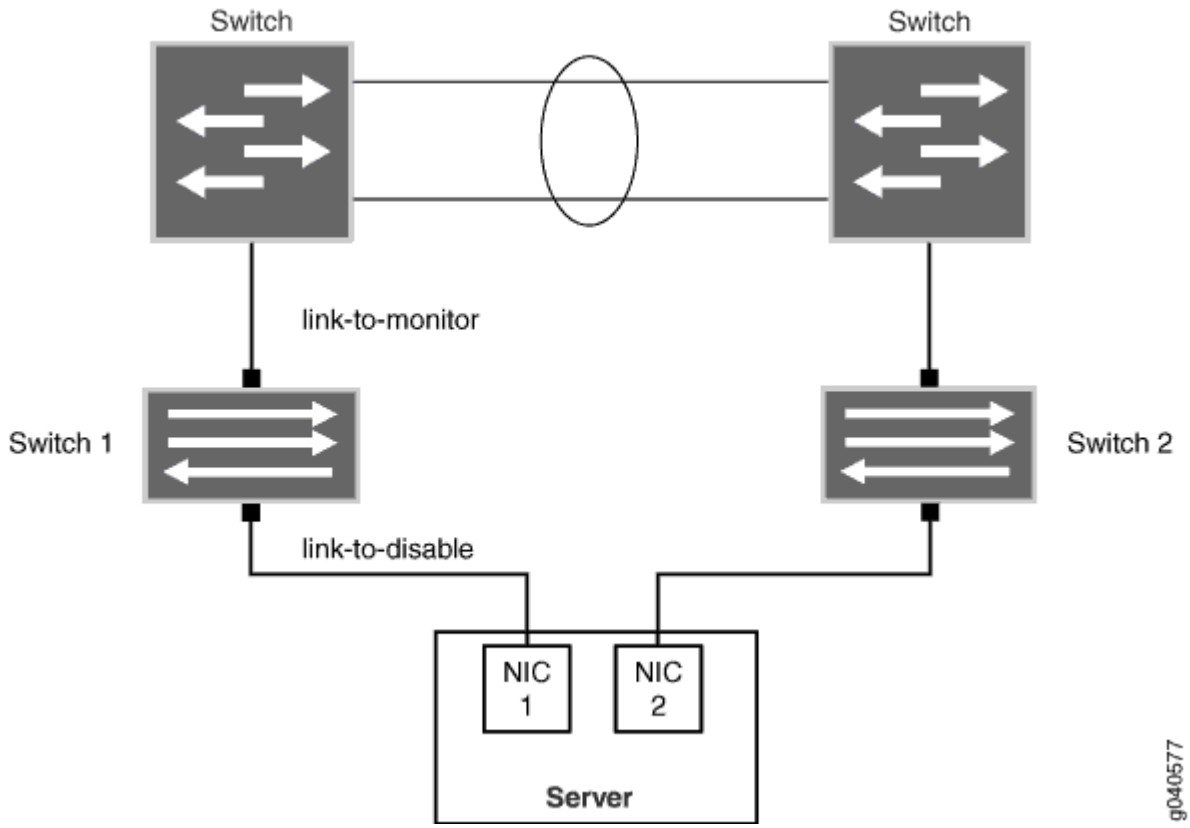


Table 20 on page 141 lists uplink failure settings for each QFX3500 switch.

Topology

Table 20: Settings for Uplink Failure Protection Example

Switch 1	Switch 2
• Group name: Group1 • Link-to-monitor interface: **xe-0/0/0** • Link-to-disable interface: **xe-0/0/1** • Debounce interval: **20**	• Group name: Group2 • Link-to-monitor interface: **xe-0/0/0** • Link-to-disable interface: **xe-0/0/1** • Debounce interval: **20**

Configuring Uplink Failure Detection on Both Switches

IN THIS SECTION

• Procedure | 141

To configure uplink failure detection on both switches, perform these tasks.

Procedure

CLI Quick Configuration

To quickly configure uplink failure protection on Switch 1 and Switch 2, copy the following commands and paste them into the switch terminal window:

```
[edit protocols]
set uplink-failure-detection group group1
set uplink-failure-detection group group2
set uplink-failure-detection group group1 link-to-monitor xe-0/0/0
set uplink-failure-detection group group1 debounce-interval 20
set uplink-failure-detection group group2 link-to-monitor xe-0/0/0
set uplink-failure-detection group group2 debounce-interval 20
```

```
set uplink-failure-detection group group1 link-to-disable xe-0/0/1
set uplink-failure-detection group group2 link-to-disable xe-0/0/1
```

Step-by-Step Procedure

To configure uplink failure protection on both switches:

1. Specify a name for the uplink failure detection group on Switch 1:

```
[edit protocols]
user@switch# set uplink-failure-detection group group1
```

2. Add an uplink interface to the group on Switch 1:

```
[edit protocols]
user@switch# set uplink-failure-detection group group1 link-to-monitor xe-0/0/0
```

3. Add a downlink interface to the group on Switch 1:

```
[edit protocols]
user@switch# set uplink-failure-detection group group1 link-to-disable xe-0/0/1
```

4. Configure the debounce interval for group1 on Switch 1:

```
[edit protocols]
user@switch# set uplink-failure-detection group group1 debounce-interval 20
```

5. Specify a name for the uplink failure detection group on Switch 2:

```
[edit protocols]
user@switch# set uplink-failure-detection group group2
```

6. Add an uplink interface to the group on Switch 2:

```
[edit protocols]
user@switch# set uplink-failure-detection group group2 link-to-monitor xe-0/0/0
```

7. Configure the debounce interval for group2 on Switch 1:

```
[edit protocols]
user@switch# set uplink-failure-detection group group2 debounce-interval 20
```

8. Add a downlink interface to the group on Switch 2:

```
[edit protocols]
user@switch# set uplink-failure-detection group group2 link-to-disable xe-0/0/1
```

Results

Display the results of the configuration:

```
uplink-failure-detection {
    group {
        group1 {
            debounce-interval 20;
                link-to-monitor {
                xe-0/0/0;
            }
            link-to-disable {
                xe-0/0/1;
            }
        }
        group2 {
            debounce-interval 20;
                link-to-monitor {
                xe-0/0/0;
            }
            link-to-disable {
                xe-0/0/1;
            }
        }
    }
}
```

Verification

IN THIS SECTION

- Verifying That Uplink Failure Detection is Working Correctly | 144

To verify that uplink failure detection is working correctly, perform the following tasks on Switch 1 and Switch 2:

Verifying That Uplink Failure Detection is Working Correctly

Purpose

Verify that the switch disables the downlink interface when it detects an uplink failure.

Action

1. View the current uplink failure detection status:

```
user@switch> show uplink-failure-detection
Group                     : group1
Uplink                    :  xe-0/0/0*
Downlink                  :  xe-0/0/1*
Failure Action            :  Inactive
Debounce Interval :  20
```



NOTE: The asterisk (*) indicates that the link is up.

2. Disable the uplink interface:

```
[edit]
user@switch# set interface xe-0/0/0 disable
```

3. Save the configuration on the switch.

4. View the current uplink failure detection status:

```
user@switch> show uplink-failure-detection
Group                       : group1
Uplink                      :  xe-0/0/0
Downlink                    :  xe-0/0/1
Failure Action              :  Active
Debounce Interval :  20
```

Meaning

The output in Step 1 shows that the uplink interface is up, and hence that the downlink interface is also up, and that the status of **Failure Action** is **Inactive**.

The output in Step 4 shows that both the uplink and downlink interfaces are down (there are no asterisks after the interface name) and that the status of **Failure Action** is changed to **Active**. This output shows that uplink failure detection is working.

Verifying That Uplink Failure Detection Is Working Correctly

IN THIS SECTION

- Purpose | **145**
- Action | **145**
- Meaning | **146**

Purpose

Verify that the switch disables the downlink interface when it detects an uplink failure.

Action

1. View the current uplink failure detection status:

```
user@switch> show uplink-failure-detection
Group                       : group1
Uplink                      :  xe-0/0/0*
Downlink                    :  xe-0/0/1*
Failure Action              :  Inactive
Debounce Interval :  20
```



NOTE: The asterisk (*) indicates that the link is up.

2. Disable the uplink interface:

```
[edit]
user@switch# set interface xe-0/0/0 disable
```

3. Save the configuration on the switch.

4. View the current uplink failure detection status:

```
user@switch> show uplink-failure-detection
Group                       : group1
Uplink                      :  xe-0/0/0
Downlink                    :  xe-0/0/1
Failure Action              :  Active
Debounce Interval :  20
```

Meaning

The output in Step 1 shows that the uplink interface is up, and hence that the downlink interface is also up, and that the status of **Failure Action** is **Inactive**.

The output in Step 4 shows that both the uplink and downlink interfaces are down (there are no asterisks after the interface name) and that the status of **Failure Action** is changed to **Active**. This output shows that uplink failure detection is working.

Targeted Broadcast

IN THIS SECTION

- Understanding Targeted Broadcast | 147
- Understanding IP Directed Broadcast | 148
- Configure Targeted Broadcast | 150
- Configuring IP Directed Broadcast (CLI Procedure) | 153
- Example: Configuring IP Directed Broadcast on a Switch | 155
- Verifying IP Directed Broadcast Status | 164

Targeted broadcast helps in remote administration tasks such as backups and wake-on LAN (WOL) on a LAN interface, and supports virtual routing and forwarding (VRF) instances. The below topic discuss the process and functioning of targeted broadcast, its configuration details, and the status of the broadcast on various platforms.

Understanding Targeted Broadcast

Targeted broadcast is a process of flooding a target subnet with Layer 3 broadcast IP packets originating from a different subnet. The intent of targeted broadcast is to flood the target subnet with the broadcast packets on a LAN interface without broadcasting to the entire network. Targeted broadcast is configured with various options on the egress interface of the router or switch, and the IP packets are broadcast only on the LAN (egress) interface. Targeted broadcast helps you implement remote administration tasks, such as backups and wake-on LAN (WOL) on a LAN interface, and supports virtual routing and forwarding (VRF) instances.

Regular Layer 3 broadcast IP packets originating from a subnet are broadcast within the same subnet. When these IP packets reach a different subnet, they are forwarded to the Routing Engine (to be forwarded to other applications). Because of this, remote administration tasks such as backups cannot be performed on a particular subnet through another subnet. As a workaround, you can enable targeted broadcast to forward broadcast packets that originate from a different subnet.

Layer 3 broadcast IP packets have a destination IP address that is a valid broadcast address for the target subnet. These IP packets traverse the network in the same way as unicast IP packets until they reach the destination subnet, as follows:

1. In the destination subnet, if the receiving router has targeted broadcast enabled on the egress interface, the IP packets are forwarded to an egress interface and the Routing Engine or to an egress interface only.
2. The IP packets are then translated into broadcast IP packets, which flood the target subnet only through the LAN interface, and all hosts on the target subnet receive the IP packets. The packets are discarded If no LAN interface exists.
3. The final step in the sequence depends on targeted broadcast:
   - If targeted broadcast is not enabled on the receiving router, the IP packets are treated as regular Layer 3 broadcast IP packets and are forwarded to the Routing Engine.
   - If targeted broadcast is enabled without any options, the IP packets are forwarded to the Routing Engine.

You can configure targeted broadcast to forward the IP packets only to an egress interface. This is helpful when the router is flooded with packets to process, or to both an egress interface and the Routing Engine.



NOTE: Any *firewall filter* that is configured on the Routing Engine loopback interface (lo0) cannot be applied to IP packets that are forwarded to the Routing Engine as a result of a targeted broadcast. This is because broadcast packets are forwarded as flood next-hop traffic and not as local next-hop traffic, and you can apply a firewall filter only to local next-hop routes for traffic directed towards the Routing Engine.

Understanding IP Directed Broadcast

IN THIS SECTION

- IP Directed Broadcast Overview | **149**
- IP Directed Broadcast Implementation | **149**
- When to Enable IP Directed Broadcast | **149**
- When Not to Enable IP Directed Broadcast | **150**

IP directed broadcast helps you implement remote administration tasks such as backups and wake-on-LAN (WOL) application tasks by sending broadcast packets targeted at the hosts in a specified destination subnet. IP directed broadcast packets traverse the network in the same way as unicast IP packets until they reach the destination subnet. When they reach the destination subnet and IP directed broadcast is enabled on the receiving switch, the switch translates (*explodes*) the IP directed broadcast packet into a broadcast that floods the packet on the target subnet. All hosts on the target subnet receive the IP directed broadcast packet.

This topic covers:

IP Directed Broadcast Overview

IP directed broadcast packets have a destination IP address that is a valid broadcast address for the subnet that is the target of the directed broadcast (the target subnet). The intent of an IP directed broadcast is to flood the target subnet with the broadcast packets without broadcasting to the entire network. IP directed broadcast packets cannot originate from the target subnet.

When you send an IP directed broadcast packet, as it travels to the target subnet, the network forwards it in the same way as it forwards a unicast packet. When the packet reaches a switch that is directly connected to the target subnet, the switch checks to see whether IP directed broadcast is enabled on the interface that is directly connected to the target subnet:

- If IP directed broadcast is enabled on that interface, the switch broadcasts the packet on that subnet by rewriting the destination IP address as the configured broadcast IP address for the subnet. The switch converts the packet to a link-layer broadcast packet that every host on the network processes.
- If IP directed broadcast is disabled on the interface that is directly connected to the target subnet, the switch drops the packet.

IP Directed Broadcast Implementation

You configure IP directed broadcast on a per-subnet basis by enabling IP directed broadcast on the Layer 3 interface of the subnet's VLAN. When the switch that is connected to that subnet receives a packet that has the subnet's broadcast IP address as the destination address, the switch broadcasts the packet to all hosts on the subnet.

By default, IP directed broadcast is disabled.

When to Enable IP Directed Broadcast

IP directed broadcast is disabled by default. Enable IP directed broadcast when you want to perform remote management or administration services such as backups or WOL tasks on hosts in a subnet that does not have a direct connection to the Internet.

Enabling IP directed broadcast on a subnet affects only the hosts within that subnet. Only packets received on the subnet's Layer 3 interface that have the subnet's broadcast IP address as the destination address are flooded on the subnet.

When Not to Enable IP Directed Broadcast

Typically, you do not enable IP directed broadcast on subnets that have direct connections to the Internet. Disabling IP directed broadcast on a subnet's Layer 3 interface affects only that subnet. If you disable IP directed broadcast on a subnet and a packet that has the broadcast IP address of that subnet arrives at the switch, the switch drops the broadcast packet.

If a subnet has a direct connection to the Internet, enabling IP directed broadcast on it increases the network's susceptibility to denial-of-service (DoS) attacks.

For example, a malicious attacker can spoof a source IP address (use a source IP address that is not the actual source of the transmission to deceive a network into identifying the attacker as a legitimate source) and send IP directed broadcasts containing Internet Control Message Protocol (ICMP) echo (ping) packets. When the hosts on the network with IP directed broadcast enabled receive the ICMP echo packets, they all send replies to the victim that has the spoofed source IP address. This creates a flood of ping replies in a DoS attack that can overwhelm the spoofed source address; this is known as a *smurf* attack. Another common DoS attack on exposed networks with IP directed broadcast enabled is a *fraggle* attack, which is similar to a smurf attack except that the malicious packet is a User Datagram Protocol (UDP) echo packet instead of an ICMP echo packet.

Configure Targeted Broadcast

IN THIS SECTION

- Configure Targeted Broadcast and Its Options | **150**
- Display Targeted Broadcast Configuration Options | **152**

The following sections explain how to configure targeted broadcast on an egress interface and its options:

Configure Targeted Broadcast and Its Options

You can configure targeted broadcast on an egress interface with different options.

Either of these configurations is acceptable:

- You can allow the IP packets destined for a Layer 3 broadcast address to be forwarded on the egress interface and to send a copy of the IP packets to the Routing Engine.
- You can allow the IP packets to be forwarded on the egress interface only.

Note that the packets are broadcast only if the egress interface is a LAN interface.

To configure targeted broadcast and its options:

1. Configure the physical interface.

```
[edit]
user@host# set interfaces interface-name
```

2. Configure the logical unit number at the `[edit interfaces interface-name` hierarchy level.

```
[edit interfaces interface-name]
user@host# set unit logical-unit-number
```

3. Configure the protocol family as inet at the `[edit interfaces interface-name unit interface-unit-number` hierarchy level.

```
[edit interfaces interface-name unit interface--unit-number]
user@host# set family inet
```

4. Configure targeted broadcast at the `[edit interfaces interface-name unit interface-unit-number family inet` hierarchy level.

```
[edit interfaces interface-name unit interface--unit-number family inet]
user@host# set targeted-broadcast
```

5. Allow IP packets to be forwarded on the egress interface only.

```
[edit interfaces interface-name unit interface-unit-number family inet targeted-broadcast]
user@host# set forward-only
```



NOTE: SRX devices do not support the targeted broadcast option `forward-and-send-to-re`.

Display Targeted Broadcast Configuration Options

IN THIS SECTION

- Example: Forward IP Packets on the Egress Interface and to the Routing Engine | **152**
- Example: Forward IP Packets on the Egress Interface Only | **153**

The following example topics display targeted broadcast configuration options:

Example: Forward IP Packets on the Egress Interface and to the Routing Engine

IN THIS SECTION

- Purpose | **152**
- Action | **152**

Purpose

Display the configuration when targeted broadcast is configured on the egress interface to forward the IP packets on the egress interface and to send a copy of the IP packets to the Routing Engine.

Action

To display the configuration, run the `show` command at the `[edit interfaces` *`interface-name`* `unit` *`interface-unit-number`* `family inet]` where the interface name is ge-2/0/0, the unit value is set to 0, and the protocol family is set to inet.

```
[edit interfaces interface-name unit interface-unit-number family inet]
user@host#show
targeted-broadcast {
    forward-only;
}
```

Example: Forward IP Packets on the Egress Interface Only

IN THIS SECTION

- Purpose | 153
- Action | 153

Purpose

Display the configuration when targeted broadcast is configured on the egress interface to forward the IP packets on the egress interface only.

Action

To display the configuration, run the `show` command at the `[edit interfaces` *`interface-name`* `unit` *`interface-unit-number`* `family inet]` where the interface name is ge-2/0/0, the unit value is set to 0, and the protocol family is set to inet.

```
[edit interfaces interface-name unit interface-unit-number family inet]
user@host#show
targeted-broadcast {
    forward-only;
}
```

Configuring IP Directed Broadcast (CLI Procedure)

Before you begin to configure IP directed broadcast:

- Ensure that the subnet on which you want broadcast packets using IP direct broadcast is not directly connected to the Internet.
- Configure a routed VLAN interface (RVI) for the subnet that will be enabled for IP direct broadcast. See Configuring Routed VLAN Interfaces on Switches (CLI Procedure).

NOTE: We recommend that you do not enable IP directed broadcast on subnets that have a direct connection to the Internet because of increased exposure to denial-of-service (DoS) attacks.

NOTE: This task uses Junos OS for EX Series switches that does not support the Enhanced Layer 2 Software (ELS) configuration style. For ELS details, see *Using the Enhanced Layer 2 Software CLI*.

You can use IP directed broadcast on an EX Series switch to facilitate remote network management by sending broadcast packets to hosts on a specified subnet without broadcasting to the entire network. IP directed broadcast packets are broadcast on only the target subnet. The rest of the network treats IP directed broadcast packets as unicast packets and forwards them accordingly.

To enable IP directed broadcast for a specified subnet:

1. Add the target subnet's logical interfaces to the VLAN:

```
[edit interfaces]
user@switch# set  ge-0/0/0.0 family ethernet-switching vlan members v1
user@switch# set  ge-0/0/1.0 family ethernet-switching vlan members v1
```

2. Configure the Layer 3 interface on the VLAN that is the target of the IP directed broadcast packets:

```
[edit interfaces]
user@switch# set  vlan.1 family inet address 10.1.2.1/24
```

3. Associate a Layer 3 interface with the VLAN:

```
[edit vlans]
user@switch# set  v1 l3-interface (VLAN) vlan.1
```

4. Enable the Layer 3 interface for the VLAN to receive IP directed broadcasts:

```
[edit interfaces]
user@switch# set  vlan.1 family inet targeted-
broadcast
```

SEE ALSO

Example: Configuring IP Directed Broadcast on a Switch | 155

Example: Configuring IP Directed Broadcast on a Switch | 155

Example: Configuring IP Directed Broadcast on a Switch

IN THIS SECTION

- Requirements | 155
- Overview and Topology | 156
- Configuring IP Directed Broadcast for non-ELS Switches | 157
- Configuring IP Directed Broadcast for Switches with ELS Support | 160

IP directed broadcast provides a method of sending broadcast packets to hosts on a specified subnet without broadcasting those packets to hosts on the entire network.

This example shows how to enable a subnet to receive IP directed broadcast packets so you can perform backups and other network management tasks remotely:

Requirements

This example uses the following software and hardware components:

- Junos OS Release 9.4 or later for EX Series switches or Junos OS Release 15.1X53-D10 for QFX10000 switches.
- One PC
- One EX Series switch or QFX10000 switch

Before you configure IP directed broadcast for a subnet:

- Ensure that the subnet does not have a direct connection to the Internet.
- Configure routed VLAN interfaces (RVIs) for the ingress and egress VLANs on the switch. For non-ELS, see Configuring Routed VLAN Interfaces on Switches (CLI Procedure) or Configuring VLANs for EX Series Switches (J-Web Procedure). For ELS, see l3-interface.

Overview and Topology

IN THIS SECTION

- Topology | 156

You might want to perform remote administration tasks such as backups and wake-on-LAN (WOL) application tasks to manage groups of clients on a subnet. One way to do this is to send IP directed broadcast packets targeted at the hosts in a particular target subnet.

The network forwards IP directed broadcast packets as if they were unicast packets. When the IP directed broadcast packet is received by a VLAN that is enabled for `targeted-broadcast`, the switch broadcasts the packet to all the hosts in its subnet.

In this topology (see Figure 8 on page 156), a host is connected to an interface on a switch to manage the clients in subnet `10.1.2.1/24`. When the switch receives a packet with the broadcast IP address of the target subnet as its destination address, it forwards the packet to the subnet's Layer 3 interface and broadcasts it to all the hosts within the subnet.

Figure 8: Topology for IP Directed Broadcast



Topology

Table 21 on page 156 shows the settings of the components in this example.

Table 21: Components of the IP Directed Broadcast Topology

Property	Settings
Ingress VLAN name	`v0`
Ingress VLAN IP address	`10.1.1.1/24`

Table 21: Components of the IP Directed Broadcast Topology *(Continued)*

Property	Settings
Egress VLAN name	v1
Egress VLAN IP address	10.1.2.1/24
Interfaces in VLAN v0	ge-0/0/3.0
Interfaces in VLAN v1	ge-0/0/0.0 and ge-0/0/1.0

Configuring IP Directed Broadcast for non-ELS Switches

IN THIS SECTION

- Procedure | 157

To configure IP directed broadcast on a subnet to enable remote management of its hosts:

Procedure

CLI Quick Configuration

To quickly configure the switch to accept IP directed broadcasts targeted at subnet 10.1.2.1/24, copy the following commands and paste them into the switch's terminal window:

```
[edit]
                        set interfaces ge-0/0/0.0 family ethernet-switching vlan members
v1
                        set interfaces ge-0/0/1.0 family ethernet-switching vlan members
v1
                        set interfaces vlan.1 family inet address
10.1.2.1/24
                        set interfaces ge-0/0/3.0 family ethernet-switching vlan members
v0
```

```
                        set interfaces vlan.0 family inet address
10.1.1.1/24
                        set vlans v1 l3-interface vlan.1
                        set vlans v0 l3-interface vlan.0
                        set interfaces vlan.1 family inet targeted-broadcast
```

Step-by-Step Procedure

To configure the switch to accept IP directed broadcasts targeted at subnet `10.1.2.1/24`:

1. Add logical interface `ge-0/0/0.0` to VLAN `v1`:

```
[edit interfaces]
user@switch# set  ge-0/0/0.0 family ethernet-switching vlan members v1
```

2. Add logical interface `ge-0/0/1.0` to VLAN `v1`:

```
[edit interfaces]
user@switch# set  ge-0/0/1.0 family ethernet-switching vlan members v1
```

3. Configure the IP address for the egress VLAN, `v1`:

```
[edit interfaces]
user@switch# set  vlan.1 family inet address 10.1.2.1/24
```

4. Add logical interface `ge-0/0/3.0` to VLAN `v0`:

```
[edit interfaces]
user@switch# set  ge-0/0/3.0 family ethernet-switching vlan members v0
```

5. Configure the IP address for the ingress VLAN:

```
[edit interfaces]
user@switch# set  vlan.0 family inet address 10.1.1.1/24
```

6. To route traffic between the ingress and egress VLANs, associate a Layer 3 interface with each VLAN:

```
[edit vlans]
user@switch# set  v1 l3-interface vlan.1
user@switch# set  v0 l3-interface vlan.0
```

7. Enable the Layer 3 interface for the egress VLAN to receive IP directed broadcasts:

```
[edit interfaces]
user@switch# set  vlan.1 family inet targeted-broadcast
user@switch# set  vlan.0 family inet targeted-broadcast
```

Results

Check the results:

```
user@switch# show
interfaces {
    ge-0/0/0 {
        unit 0 {
            family ethernet-switching {
                vlan {
                    members v1;
                }
            }
        }
    }
    ge-0/0/1 {
        unit 0 {
            family ethernet-switching {
                vlan {
                    members v1;
                }
            }
        }
    }
    ge-0/0/3 {
        unit 0 {
            family ethernet-switching {
```

```
                vlan {
                    members v0;
                }
            }
        }
    }
    vlan {
        unit 0 {
            family inet {
                targeted-broadcast;
                address 10.1.1.1/24;
            }
        }
        unit 1 {
            family inet {
                targeted-broadcast;
                address 10.1.2.1/24;
            }
        }
    }
vlans {
    default;
    v0 {
        l3-interface vlan.0;
    }
    v1 {
        l3-interface vlan.1;
    }
}
```

Configuring IP Directed Broadcast for Switches with ELS Support

IN THIS SECTION

- Procedure | 161

To configure IP directed broadcast on a subnet to enable remote management of its hosts:

Procedure

CLI Quick Configuration

To quickly configure the switch to accept IP directed broadcasts targeted at subnet `10.1.2.1/24`, copy the following commands and paste them into the switch's terminal window:

```
[edit]
                        set interfaces ge-0/0/0.0 family ethernet-switching vlan members
v1
                        set interfaces ge-0/0/1.0 family ethernet-switching vlan members
v1
                        set interfaces irb.1 family inet address
10.1.2.1/24
                        set interfaces ge-0/0/3.0 family ethernet-switching vlan members
v0
                        set interfaces irb.0 family inet address
10.1.1.1/24
                        set vlans v1 l3-interface irb.1
                        set vlans v0 l3-interface irb.0
                        set interfaces irb.1 family inet targeted-broadcast
```

Step-by-Step Procedure

To configure the switch to accept IP directed broadcasts targeted at subnet `10.1.2.1/24`:

1. Add logical interface `ge-0/0/0.0` to VLAN `v1`:

```
[edit interfaces]
user@switch# set  ge-0/0/0.0 family ethernet-switching vlan members v1
```

2. Add logical interface `ge-0/0/1.0` to VLAN `v1`:

```
[edit interfaces]
user@switch# set  ge-0/0/1.0 family ethernet-switching vlan members v1
```

3. Configure the IP address for the egress VLAN, `v1`:

```
[edit interfaces]
user@switch# set  irb.1 family inet address 10.1.2.1/24
```

4. Add logical interface `ge-0/0/3.0` to VLAN `v0`:

```
[edit interfaces]
user@switch# set  ge-0/0/3.0 family ethernet-switching vlan members v0
```

5. Configure the IP address for the ingress VLAN:

```
[edit interfaces]
user@switch# set  irb.0 family inet address 10.1.1.1/24
```

6. To route traffic between the ingress and egress VLANs, associate a Layer 3 interface with each VLAN:

```
[edit vlans]
user@switch# set  v1 l3-interface irb.1
user@switch# set  v0 l3-interface irb.0
```

7. Enable the Layer 3 interface for the egress VLAN to receive IP directed broadcasts:

```
[edit interfaces]
user@switch# set  irb.1 family inet targeted-broadcast
user@switch# set  irb.0 family inet targeted-broadcast
```

On QFX5000 Series, EX4300 Series, and EX4600 Series switches, the maximum number of targeted-broadcast supported is 63.

Results

Check the results:

```
user@switch# show
interfaces {
```

```
    ge-0/0/0 {
        unit 0 {
            family ethernet-switching {
                vlan {
                    members v1;
                }
            }
        }
    }
    ge-0/0/1 {
        unit 0 {
            family ethernet-switching {
                vlan {
                    members v1;
                }
            }
        }
    }
    ge-0/0/3 {
        unit 0 {
            family ethernet-switching {
                vlan {
                    members v0;
                }
            }
        }
    }
    vlan {
        unit 0 {
            family inet {
                targeted-broadcast;
                address 10.1.1.1/24;
            }
        }
        unit 1 {
            family inet {
                targeted-broadcast;
                address 10.1.2.1/24;
            }
        }
    }
vlans {
    default;
```

```
    v0 {
        l3-interface irb.0;
    }
    v1 {
        l3-interface irb.1;
    }
}
```

SEE ALSO

Configuring IP Directed Broadcast for Switches

Verifying IP Directed Broadcast Status

IN THIS SECTION

- Purpose | 164
- Action | 164

Purpose

Verify that IP directed broadcast is enabled and is working on the subnet.

Action

Use the `show vlans extensive` command to verify that IP directed broadcast is enabled and working on the subnet as shown in "Example: Configuring IP Directed Broadcast on a Switch" on page 155.

ARP

IN THIS SECTION

- Static ARP Table Entries Overview | 165
- Configuring Static ARP Table Entries For Mapping IP Addresses to MAC Addresses | 165
- Restricted and Unrestricted Proxy ARP Overview | 168
- Configuring Restricted and Unrestricted Proxy ARP | 171
- Configuring Gratuitous ARP | 172

Static address resolution protocol (ARP) table entries are reponded to by default when the destination address of the ARP is on the local network. These static ARP addresses can be configured for Ethernet or Gigabit Ethernet interfaces. The topics below discuss the overview of static ARP table entries, restricted and unrestricted proxy ARP, configuration details to map the IP addresses to the MAC addresses.

Static ARP Table Entries Overview

For Fast Ethernet, Gigabit Ethernet, Tri-Rate Ethernet copper, and 10-Gigabit Ethernet interfaces, you can configure static ARP table entries, defining mappings between IP and MAC addresses.

SEE ALSO

Ethernet Interfaces User Guide for Routing Devices

Configuring Static ARP Table Entries For Mapping IP Addresses to MAC Addresses

By default, the device responds to an Address Resolution Protocol (ARP) request only if the destination address of the ARP request is on the local network of the incoming interface. For Fast Ethernet or Gigabit Ethernet interfaces, you can configure static ARP entries that associate the IP addresses of

nodes on the same Ethernet subnet with their media access control (MAC) addresses. These static ARP entries enable the device to respond to ARP requests even if the destination address of the ARP request is not local to the incoming Ethernet interface.

Also, unlike dynamically learned ARP entries, static ARP entries do not age out. You can also configure static ARP entries in a troubleshooting situation or if your device is unable to learn a MAC address dynamically.



NOTE: By default, an ARP policer is installed that is shared among all the Ethernet interfaces on which you have configured the `family inet` statement. By including the `arp` statement at the `[edit interfaces` *`interface-name`* `unit` *`logical-unit-number`* `family inet policer]` hierarchy level, you can apply a specific ARP-packet policer to an interface. This feature is not available on EX Series switches.

To configure static ARP entries:

1. In the configuration mode, at the `[edit]` hierarchy level, configure the router interface on which the ARP table entries for the router is configured.

```
[edit]
user@host# edit interfaces interface-name
```

2. Configure the protocol family, the logical unit of the interface, and the interface address of the router interface at the `[edit interfaces` *`interface-name`*`]` hierarchy level. While configuring the protocol family, specify `inet` as the protocol family.



NOTE: When you need to conserve IP addresses, you can configure an Ethernet interface to be unnumbered by including the `unnumbered-address` statement at the `[edit interfaces` *`interface-name`* `unit` *`logical-unit-number`* `family inet]` hierarchy level.

```
[edit interfaces interface-name]
user@host# edit unit logical-unit-number family inet address interface-address
```

3. Configure a static ARP entry by specifying the IP address and the MAC address that are to be mapped to each other. The IP address specified must be part of the subnet defined in the enclosing `address` statement. The MAC address must be specified as hexadecimal bytes in the following formats:

nnnn.*nnnn*.*nnnn* or *nn*:*nn*:*nn*:*nn*:*nn*:*nn*format. For instance, you can use either `0011.2233.4455` or `00:11:22:33:44:55`.

```
[edit interfaces interface-name unit logical-unit-number family inet address interface-address
user@host# set arp ip-address mac mac-address
```

4. Configure another static ARP entry by specifying the IP address and the MAC address that are to be mapped to each other. You can also associate a multicast MAC address with a unicast IP address by including the `multicast-mac` option with the `arp` statement. You can optionally configure the router to respond to ARP requests for the specified IP address by using the `publish` option with the `arp` statement.

> **NOTE**: For unicast MAC addresses only, if you include the `publish` option, the router or switch replies to proxy ARP requests.

```
[edit interfaces interface-name unit logical-unit-number family inet address interface-address
user@host# set arp ip-address multicast-mac mac-address publish
```

> **NOTE**: The Junos OS supports the IPv6 static neighbor discovery cache entries, similar to the static ARP entries in IPv4.

SEE ALSO

- arp
- Management Ethernet Interface Overview
- *Applying Policers*
- Configuring an Unnumbered Interface
- Ethernet Interfaces User Guide for Routing Devices

Restricted and Unrestricted Proxy ARP Overview

IN THIS SECTION

- Restricted Proxy ARP | 168
- Unrestricted Proxy ARP | 168
- Topology Considerations for Unrestricted Proxy ARP | 169

By default, the Junos OS responds to an Address Resolution Protocol (ARP) request only if the destination address of the ARP request is local to the incoming interface.

For Ethernet Interfaces, you can configure the router or switches to proxy-reply to the ARP requests using the restricted or unrestricted proxy ARP configuration.

You might want to configure restricted or unrestricted proxy ARP for routers that act as provider edge (PE) devices in Ethernet Layer 2 LAN switching domains.



NOTE: From Junos OS Release 10.0 onward, Junos OS does not respond to proxy ARP requests with the default route 0.0.0.0. This behavior is in compliance with RFC 1027.

Restricted Proxy ARP

Restricted proxy ARP enables the router or switch to respond to the ARP requests in which the physical networks of the source and target are not the same and the router or switch has an active route to the target address in the ARP request. The router does not reply if the target address is on the same subnet and the same interface as the ARP requestor.

Unrestricted Proxy ARP

Unrestricted proxy ARP enables the router or switch to respond to any ARP request, on condition that the router has an active route to the destination address of the ARP request. The route is not limited to the incoming interface of the request, nor is it required to be a direct route.



WARNING: If you configure unrestricted proxy ARP, the proxy router replies to ARP requests for the target IP address on the same interface as the incoming ARP request. This behavior is appropriate for cable modem termination system (CMTS) environments,

but might cause Layer 2 reachability problems if you enable unrestricted proxy ARP in other environments.

When an IP client broadcasts the ARP request across the Ethernet wire, the end node with the correct IP address responds to the ARP request and provides the correct MAC address. If the unrestricted proxy ARP feature is enabled, the router response is redundant and might fool the IP client into determining that the destination MAC address within its own subnet is the same as the address of the router.

NOTE: While the destination address can be remote, the source address of the ARP request must be on the same subnet as the interface upon which the ARP request is received. For security reasons, this rule applies to both unrestricted and restricted proxy ARP.

Topology Considerations for Unrestricted Proxy ARP

In most situations, you should not configure the router or switch to perform unrestricted proxy ARP. Do so only for special situations, such as when cable modems are used. Figure 9 on page 170 and Figure 10 on page 170 show examples of situations in which you might want to configure unrestricted proxy ARP.

In Figure 9 on page 170, the edge device is not running any IP protocols. In this case, you configure the core router to perform unrestricted proxy ARP. The edge device is the client of the proxy.

In Figure 10 on page 170, the Broadband Remote Access Server (B-RAS) routers are not running any IP protocols. In this case, you configure unrestricted proxy ARP on the B-RAS interfaces. This allows the core device to behave as though it is directly connected to the end users.

Figure 9: Edge Device Case for Unrestricted Proxy ARP

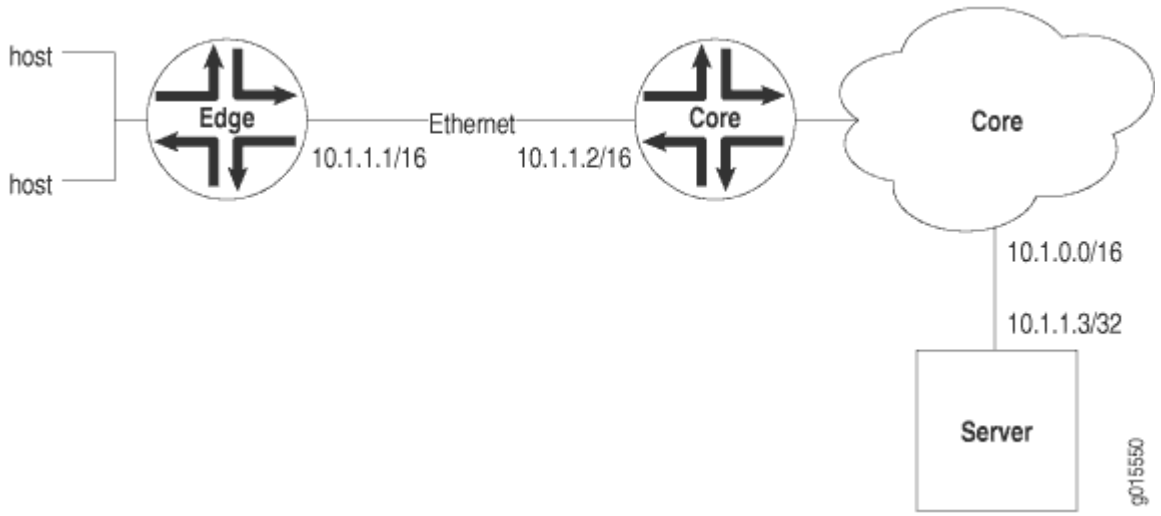
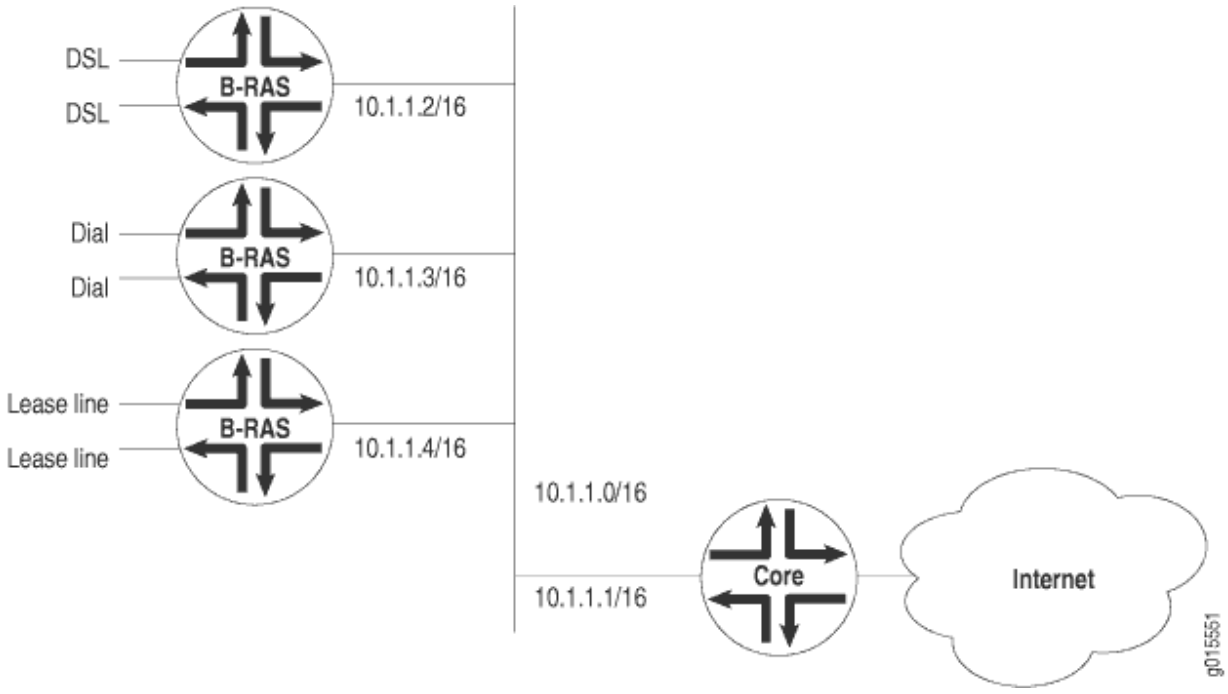


Figure 10: Core Device Case for Unrestricted Proxy ARP



SEE ALSO

Ethernet Interfaces User Guide for Routing Devices

Configuring Restricted and Unrestricted Proxy ARP

To configure restricted or unrestricted proxy ARP, include the `proxy-arp` statement:

```
proxy-arp  (restricted |unrestricted);
```

You can include this statement at the following hierarchy levels:

- `[edit interfaces` *`interface-name`* `unit` *`logical-unit-number`* `]`
- `[edit logical-systems` *`logical-system-name`* `interfaces` *`interface-name`* `unit` *`logical-unit-number`*`]`

To return to the default—that is, to disable restricted or unrestricted proxy ARP—delete the `proxy-arp` statement from the configuration:

```
[edit]
user@host# delete interfaces interface-name unit logical-unit-number proxy-arp
```

You can track the number of restricted or unrestricted proxy ARP requests processed by the router or switch by issuing the `show system statistics arp` operational mode command.

NOTE: When proxy ARP is enabled as default or unrestricted, the router or switch responds to any ARP request as long as the device has an active route to the target address of the ARP request. This gratuitous ARP behavior can result in an error when the receiving interface and target response interface are the same and the end device (for example, a client) performs a duplicate address check. To prevent this error, configure the router or switch interface with the `no-gratuitous-arp-request` statement. See "Configuring Gratuitous ARP" on page 172 for information about how to disable responses to gratuitous ARP requests.

SEE ALSO

Ethernet Interfaces User Guide for Routing Devices

Configuring Gratuitous ARP

Gratuitous Address Resolution Protocol (ARP) requests help detect duplicate IP addresses. A gratuitous ARP is a broadcast request for a router's own IP address. If a router or switch sends an ARP request for its own IP address and no ARP replies are received, the router- or switch-assigned IP address is not being used by other nodes. However, if a router or switch sends an ARP request for its own IP address and an ARP reply is received, the router- or switch-assigned IP address is already being used by another node.

Gratuitous ARP replies are reply packets sent to the broadcast MAC address with the target IP address set to be the same as the sender's IP address. When the router or switch receives a gratuitous ARP reply, the router or switch can insert an entry for that reply in the ARP cache. By default, updating the ARP cache on gratuitous ARP replies is disabled on the router or switch.

To enable updating of the ARP cache for gratuitous ARPs:

1. In configuration mode, go to the `[edit interfaces interface-name]` hierarchy level.

```
[edit]
user@host# edit interfaces interface-name
```

2. Include the `gratuitous-arp-reply` statement.

```
[edit interfaces interface-name]
user@host# set gratuitous-arp-reply
```

To restore the default behavior, that is, to disable updating of the ARP cache for gratuitous ARP, delete the `gratuitous-arp-reply` statement from the configuration:

```
[edit interfaces interface-name]
user@host# delete gratuitous-arp-reply;
```

By default, the router or switch responds to gratuitous ARP requests. However, on Ethernet interfaces, you can disable responses to gratuitous ARP requests.

To disable responses to gratuitous ARP requests:

1. In configuration mode, go to the `[edit interfaces interface-name]` hierarchy level.

```
[edit]
user@host# edit interfaces interface-name
```

2. Include the `no-gratuitous-arp-request` statement.

```
[edit interfaces interface-name]
user@host# set no-gratuitous-arp-request
```

To return to the default—that is, to respond to gratuitous ARP requests—delete the `no-gratuitous-arp-request` statement from the configuration:

```
[edit interfaces interface-name]
user@host# delete no-gratuitous-arp-request
```

SEE ALSO

- gratuitous-arp-reply
- *no-gratuitous-arp-request*
- *Ethernet Interfaces Overview*
- Ethernet Interfaces User Guide for Routing Devices

Use of Resillient Hashing to Minimize Flow Remapping

IN THIS SECTION

- Limitations and Caveats for Resilient Hashing | **176**
- Configuring Resilient Hashing for ECMP | **176**
- Configuring Resilient Hashing for Aggregated Ethernet Interfaces | **177**

In deployments between network endpoints, it is necessary to preserve established connections and associated Layer 2 and Layer 3 paths. If there is any change in the network, such as failure of a networking device or a server, the packets take a new path.

Resilient hashing reduces the impact of the network change. Each ECMP with resilient hashing is assigned a 256-entry region in the load-balancing table (also known as macro-flow table). Each entry in the table stores member link ID assigned to that macro-flow.

Resilient hashing works as described below:

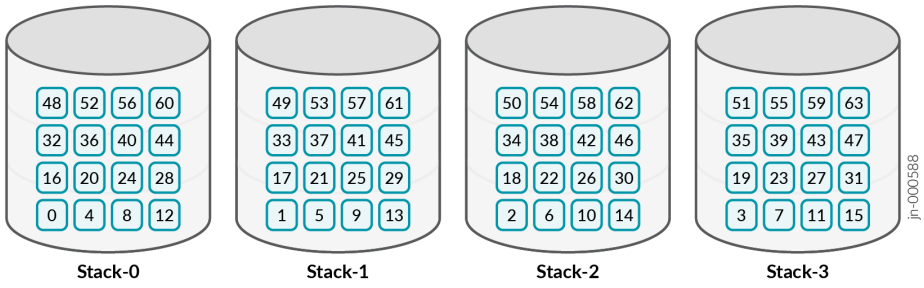
- Hash incoming packets to one of these macro-flow entries or buckets.
- You then link packets to the paths in the ECMP group.

If we use a "basket" to represent each member link/path, the resilient hashing operations can be modeled as putting buckets (macro flows) into one of the baskets.

If we have N buckets and P paths for ECMP group, use the following sequence:

1. The initial bucket mapping is generated using a round-robin method. Thus, all buckets are almost equally (N/P) distributed among the ECMP group members. Later, the buckets move around based on the path addition or deletion events.

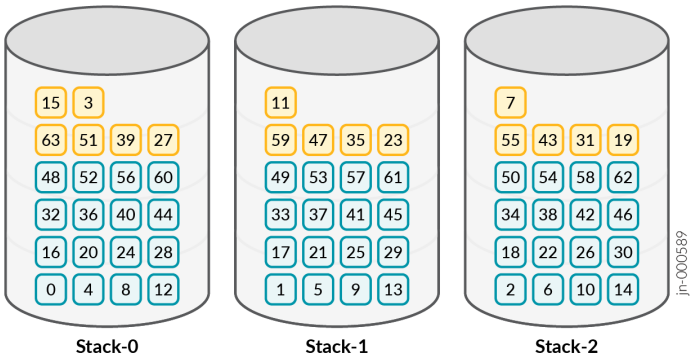
   If N=64 buckets and P=4 paths, you distribute all 64 buckets in a round-robin manner. Since you have 4 paths, there are 4 stacks. Each stack corresponds to one path. Each stack has the same number of buckets, N/P=16.



   Last_processed_path= 0 (refer to Step 5 of the algorithm).

2. If there is a path failure or removal, you suddenly remove all the buckets from the failed path/stack and push them into remaining paths/stacks in a circular round-robin manner.

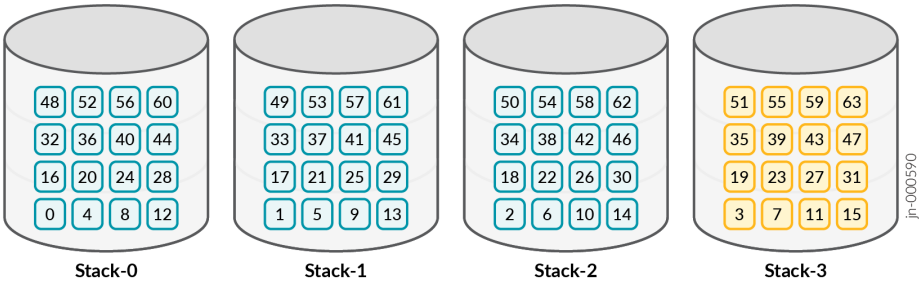
   If you remove path 3 (Stack 3 in the above image), you need to move all the buckets from Stack 3 (orange in the figure below) to remaining stacks.



3. If there is a path addition, you suddenly remove N/(P+1) buckets from the existing paths in circular round-robin manner and push them into the newly added path/stack.

   If you add a new path, you need to move N/P+1=64/4=16 buckets from existing stacks (stacks 0, 1, 2). All orange buckets are now back in stack 3, blue stacks are not moved and are intact.

   `Last_processed_path= 0`



4. Circular round-robin direction for Step 2 and Step 3 is opposite. It is important to determine the first stack from which circular round-robin starts. You keep an index pointer `last_processed_path` that provides the start stack index for Step 2 and before start stack for Step 3.

5. 5. To set `last_processed_path`, do the following:

   - When you push buckets as in Step 2, `last_processed_path` is the next stack of the last stack where you pushed the last bucket.
   - When you remove buckets as in Step 3, `last_processed_path` is the last stack from where the bucket was removed.

Limitations and Caveats for Resilient Hashing

- Resilient hashing is supported only on the equal-cost BGP routes based ECMP group. When you configure other protocols or static routes having higher priority than BGP routes, resilient hashing is not supported.
- Resilient hashing is not supported on mixed speed LAG.
- 128-way ECMP resilient hashing is not supported with current design. Only 64-way ECMP resilient hashing is supported.
- Mixed-Rate Aggregate Ethernet (AE) and Adaptive Load Balancing (ALB) AE are not supported with current resilient hashing design.

Configuring Resilient Hashing for ECMP

1. Enable resilient hashing for select ECMP routes. Create a separate routing policy to match incoming routes to one or more destination prefixes. See Configuring the Default Action in Routing Policies.
2. Apply the policy at the required level(s) of the BGP configuration hierarchy – global, group, or peer:

```
protocols {
        group BGP-GROUP-SERVERS {
            import POLICY-CHASH;
        }
        group BGP-GROUP-EXTERNAL-PEERS {
            neighbor 192.168.0.1 {
                import POLICY-CHASH;
            }
        }
        import POLICY-CHASH;
    }
}
```



NOTE: A peer-level import or export statement overrides a group import or export statement. A group-level import or export statement overrides a global BGP import or export statement. A key point is that in a configuration as shown above, only the most explicit policy is applied. A neighbor-level policy is more explicit than a group-level

policy, and a group-level policy than a global policy. (Although the same policy is applied at each level in the above example for illustration purposes, the result is unaffected.)

If you need a neighbor to perform the function of all the three policies, perform either of the following:

- You can write and apply a new neighbor-level policy that encompasses the functions of the other three.
- You can apply all three existing policies, as a chain, to this neighbor.

3. [Optional] Select packet fields used in the hash-key computation. The following examples are from PTX10001-36MR 22.2R1.12-Junos OS Evolved:

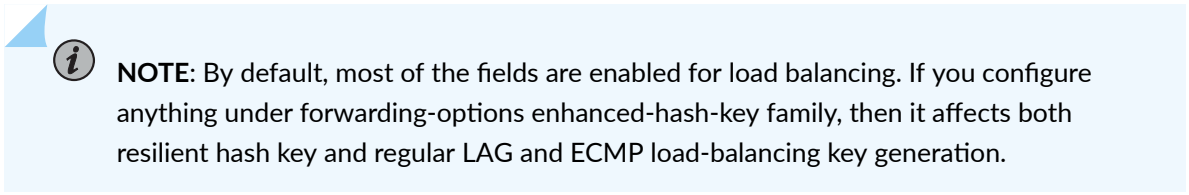
   Use the following commands to select packet fields:

   a. `user@router# set forwarding-options enhanced-hash-key family family`

      Here, family can take up `inet`, `inet6`, `mpls`, or `multiservice` values.

   b. `user@router# set forwarding-options enhanced-hash-key hash-seed`

   c. `user@router# set forwarding-options enhanced-hash-key resilient-hash-seed`



NOTE: By default, most of the fields are enabled for load balancing. If you configure anything under forwarding-options enhanced-hash-key family, then it affects both resilient hash key and regular LAG and ECMP load-balancing key generation.

Configuring Resilient Hashing for Aggregated Ethernet Interfaces

Use the following command to configure:

```
user@router# set interface ae1 aggregated-ehter-options resilient-hash
```

Generic Routing Encapsulation (GRE)

IN THIS SECTION

- Understanding Generic Routing Encapsulation | 178
- Configuring Generic Routing Encapsulation Tunneling | 183
- Verifying That Generic Routing Encapsulation Tunneling Is Working Correctly | 185

Generic routing encapsulation (GRE) is a virtual point to point link that encapsulates data traffic in a tunnel . The below topics discusses the tunneling of GRE, encapsulation and de-capsulation process, configuring GREs and verifying the working of GREs.

Understanding Generic Routing Encapsulation

IN THIS SECTION

- Overview of GRE | 178
- GRE Tunneling | 179
- Using a Firewall Filter to De-encapsulate GRE Traffic on a QFX5100, QFX10000, and OCX Series Switches | 181
- Configuration Limitations | 181

Generic routing encapsulation (GRE) provides a private path for transporting packets through an otherwise public network by encapsulating (or tunneling) the packets.

This topic describes:

Overview of GRE

GRE encapsulates data packets and redirects them to a device that de-encapsulates them and routes them to their final destination. This allows the source and destination switches to operate as if they

have a virtual point-to-point connection with each other (because the outer header applied by GRE is transparent to the encapsulated payload packet). For example, GRE tunnels allow routing protocols such as RIP and OSPF to forward data packets from one switch to another switch across the Internet. In addition, GRE tunnels can encapsulate multicast data streams for transmission over the Internet.

GRE is described in RFC 2784 (obsoletes earlier RFCs 1701 and 1702). The switches support RFC 2784, but not completely. (For a list of limitations, see "Configuration Limitations" on page 181.)

As a *tunnel source router*, the switch encapsulates a payload packet for transport through the tunnel to a destination network. The payload packet is first encapsulated in a GRE packet, and then the GRE packet is encapsulated in a delivery protocol. The switch performing the role of a *tunnel remote router* extracts the tunneled packet and forwards the packet to its destination. Note that you can use one firewall term to terminate many GRE tunnels on a QFX5100 switch.

GRE Tunneling

Data is routed by the system to the GRE endpoint over routes established in the route table. (These routes can be statically configured or dynamically learned by routing protocols such as RIP or OSPF.) When a data packet is received by the GRE endpoint, it is de-encapsulated and routed again to its destination address.

GRE tunnels are *stateless*--–that is, the endpoint of the tunnel contains no information about the state or availability of the remote tunnel endpoint. Therefore, the switch operating as a tunnel source router cannot change the state of the GRE tunnel interface to down if the remote endpoint is unreachable.

For details about GRE tunneling, see:

Encapsulation and De-Encapsulation on the Switch

Encapsulation—A switch operating as a tunnel source router encapsulates and forwards GRE packets as follows:

1. When a switch receives a data packet (payload) to be tunneled, it sends the packet to the tunnel interface.
2. The tunnel interface encapsulates the data in a GRE packet and adds an outer IP header.
3. The IP packet is forwarded on the basis of the destination address in the outer IP header.

De-encapsulation—A switch operating as a tunnel remote router handles GRE packets as follows:

1. When the destination switch receives the IP packet from the tunnel interface, the outer IP header and GRE header are removed.
2. The packet is routed based on the inner IP header.

Number of Source and Destination Tunnels Allowed on a Switch

QFX5100 and OCX Series switches support as many as 512 GRE tunnels, including tunnels created with a firewall filter. That is, you can create a total of 512 GRE tunnels, regardless of which method you use.

EX switches support as many as 500 GRE tunnels between switches transmitting IPv4 or IPv6 payload packets over GRE. If a passenger protocol in addition to IPv4 and IPv6 is used, you can configure up to 333 GRE tunnels between the switches.

An EX switch can have a maximum of 20 tunnel source IP addresses configured, and each tunnel source IP can be configured with up to 20 destination IP addresses on a second switch. As a result, the two connected switches can have a maximum of 400 GRE tunnels. If the first switch is also connected to a third switch, the possible maximum number of tunnels is 500.

Class of Service on GRE Tunnels

When a network experiences congestion and delay, some packets might be dropped. Junos OS *class of service* (CoS) divides traffic into classes to which you can apply different levels of throughput and packet loss when congestion occurs and thereby set rules for packet loss. For details about CoS, see Junos OS CoS for EX Series Switches Overview.

The following CoS components are available on a switch operating as a GRE tunnel source router or GRE tunnel remote router:

- At the GRE tunnel source—On a switch operating as a tunnel source router, you can apply CoS classifiers on an *ingress port* or on a *GRE port*, with the following results on CoS component support on tunneled packets:
  - Schedulers only—Based on the CoS classification on the ingress port, you can apply CoS schedulers on a GRE port of the switch to define output queues and control the transmission of packets through the tunnel after GRE encapsulation. However, you cannot apply CoS *rewrite rules* to these packets.
  - Schedulers and rewrite rules—Depending on the CoS classification on the GRE port, you can apply both schedulers and rewrite rules to the encapsulated packets transmitted through the tunnel.

  NOTE: You cannot configure BA classifiers on `gr-` interfaces. You must classify traffic on `gr-` interfaces using firewall filters (multifield classifiers).

- At the GRE tunnel endpoint—When the switch is a tunnel remote router, you can apply CoS classifiers on the GRE port and schedulers and rewrite rules on the egress port to control the transmission of a de-encapsulated GRE packet out from the egress port.

Applying Firewall Filters to GRE Traffic

Firewall filters provide rules that define whether to permit, deny, or forward packets that are transiting an interface on a switch. (For details, see Firewall Filters for EX Series Switches Overview.) Because of the encapsulation and de-encapsulation performed by GRE, you are constrained as to where you can apply a firewall filter to filter tunneled packets and which header will be affected. Table 22 on page 181 identifies these constraints.

Table 22: Firewall Filter Application Points for Tunneled Packets

Endpoint Type	Ingress Interface	Egress Interface
Source (encapsulating)	inner header	outer header
Remote (de-encapsulating)	Cannot filter packets on ingress interface	inner header

Using a Firewall Filter to De-encapsulate GRE Traffic on a QFX5100, QFX10000, and OCX Series Switches

You can also use a firewall filter to de-encapsulate GRE traffic on switches . This feature provides significant benefits in terms of scalability, performance, and flexibility because you don't need to create a tunnel interface to perform the de-encapsulation. For example, you can terminate many tunnels from multiple source IP addresses with one firewall term. See *Configuring a Firewall Filter to De-Encapsulate GRE Traffic* for information about how to configure a firewall filter for this purpose.

Configuration Limitations

Table 23 on page 181 lists features that are not supported with GRE.

Table 23: Features Not Supported with GRE

EX Switches	QFX Switches
MPLS over GRE tunnels	MPLS over GRE tunnels
GRE keepalives	GRE keepalives

GRE keys, payload packet fragmentation, and sequence numbers for fragmented packets	GRE keys, payload packet fragmentation, and sequence numbers for fragmented packets
BGP dynamic tunnels	BGP dynamic tunnels
Outer IP address must be IPv4	Outer IP address must be IPv4
	On QFX10002 , QFX10008 and QFX5K Series switches, If you configure GRE tunneling with the underlying ECMP next-hop instead of a Unicast next-hop, GRE tunnel encapsulation fails and network traffic is dropped
Bidirectional Forwarding Detection (BFD) protocol over GRE distributed mode	
OSPF limitation—Enabling OSPF on a GRE interface creates two equal-cost routes to the destination: one through the Ethernet network or uplink interface and the other through the tunnel interface. If data is routed through the tunnel interface, the tunnel might fail. To keep the interface operational, we recommend that you use a static route, disable OSPF on the tunnel interface, or configure the peer not to advertise the tunnel destination over the tunnel interface.	
	QFX series switches do not support configuring GRE interface and the underlying tunnel source interface in two different routing instances. If you try this configuration, it will result in a commit error.

SEE ALSO

Configuring a Firewall Filter to De-Encapsulate GRE Traffic

Configuring Generic Routing Encapsulation Tunneling

IN THIS SECTION

- Configuring a GRE Tunnel | 183

Generic routing encapsulation (GRE) provides a private path for transporting packets through an otherwise public network by encapsulating (or tunneling) the packets. GRE tunneling is accomplished through tunnel endpoints that encapsulate or de-encapsulate traffic.

You can also use a firewall filter to de-encapsulate GRE traffic on QFX5100 and OCX Series switches. This feature provides significant benefits in terms of scalability, performance, and flexibility because you don't need to create a tunnel interface to perform the de-encapsulation. For example, you can terminate many tunnels from multiple source IP addresses with one firewall term. For more information on this feature, see *Configuring a Firewall Filter to De-Encapsulate GRE Traffic*.

To configure a GRE tunnel port on a switch:

1. Determine the network port or uplink port on your switch to convert to a GRE tunnel port.
2. Configure the port as a tunnel port for GRE tunnel services:

```
[edit chassis]user@switch# set fpc slot pic pic-number  tunnel-port port-number tunnel-
services
```

NOTE: For QFX10000, gr-0/0/0 interface is created by default. Also, you need not configure the **set fpc** *slot* **pic** *pic-number* **tunnel-port** *port-number* **tunnel-services**statement.

This topic describes:

Configuring a GRE Tunnel

To configure a GRE tunnel interface:

1. Create a GRE interface with a unit number and address:

```
[edit interfaces]
user@switch# set gr-0/0/0 unit number family inet address
```

> **NOTE**: The base name of the interface must be `gr-0/0/0`.

This is a pseudo interface, and the address you specify can be any IP address. The routing table must specify `gr-0/0/0.x` as the outgoing interface for any packets that will be tunneled.

If you configure a GRE interface on a QFX5100 switch that is a member of a Virtual Chassis and later change the Virtual Chassis member number of the switch, the name of the GRE interface does not change in any way (because it is a pseudo interface). For example, if you change the member number from `0` to `5`, the GRE interface name does *not* change from `gr-0/0/0.x` to `gr-5/0/0.x`.

2. Specify the tunnel source address for the logical interface:

```
[edit interfaces]
user@switch# set gr-0/0/0 unit number tunnel source source-address
```

3. Specify the destination address:

```
[edit interfaces]
user@switch# set gr-0/0/0 unit number tunnel destination destination-address
```

The destination address must be reachable through static or dynamic routing. If you use static routing, you must get the destination MAC address (for example, by using `ping`) before user traffic can be forwarded through the tunnel.

> **NOTE**: On QFX10002 and QFX10008 switches, If you configure GRE tunneling with the underlying ECMP next-hop instead of Unicast next-hop, GRE tunnel encapsulation fails and the network traffic is dropped.

> **NOTE**: Indirect egress next-hops is currently not supported in the GRE implementation for QFX10000 switches.

RELATED DOCUMENTATION

Configuring a Firewall Filter to De-Encapsulate GRE Traffic

Verifying That Generic Routing Encapsulation Tunneling Is Working Correctly

IN THIS SECTION

- Purpose | 185
- Action | 185
- Meaning | 186

Purpose

Verify that the generic routing encapsulation (GRE) interface is sending tunneled traffic.

Action

Display status information about the specified GRE interface by using the command `show interfaces` .

```
user@switch> show interfaces gr-0/0/0.0
Physical interface: gr-0/0/0, Enabled, Physical link is Up
Interface index: 132, SNMP ifIndex: 26
  Type: GRE, Link-level type: GRE, MTU: Unlimited, Speed: 800mbps
  Device flags   : Present Running
  Interface flags: Point-To-Point SNMP-Traps
  Input rate     : 0 bps (0 pps)
  Output rate    : 0 bps (0 pps)

  Logical interface gr-0/0/0.0 (Index 68) (SNMP ifIndex 47)
    Flags: Point-To-Point SNMP-Traps 16384
    IP-Header 10.1.1.2:10.1.1.1:47:df:64:0000000000000000 Encapsulation: GRE-NULL
  Input packets : 0
  Output packets: 0
    Protocol inet, MTU: 1476
```

```
        Flags: None
        Addresses, Flags: Is-Primary
          Local: 10.0.0.0
```

Meaning

The output indicates that the GRE interface gr-0/0/0 is up. The output displays the name of the physical interface and the traffic statistics for this interface---the number of and the rate at which input and output bytes and packets are received and transmitted on the physical interface.

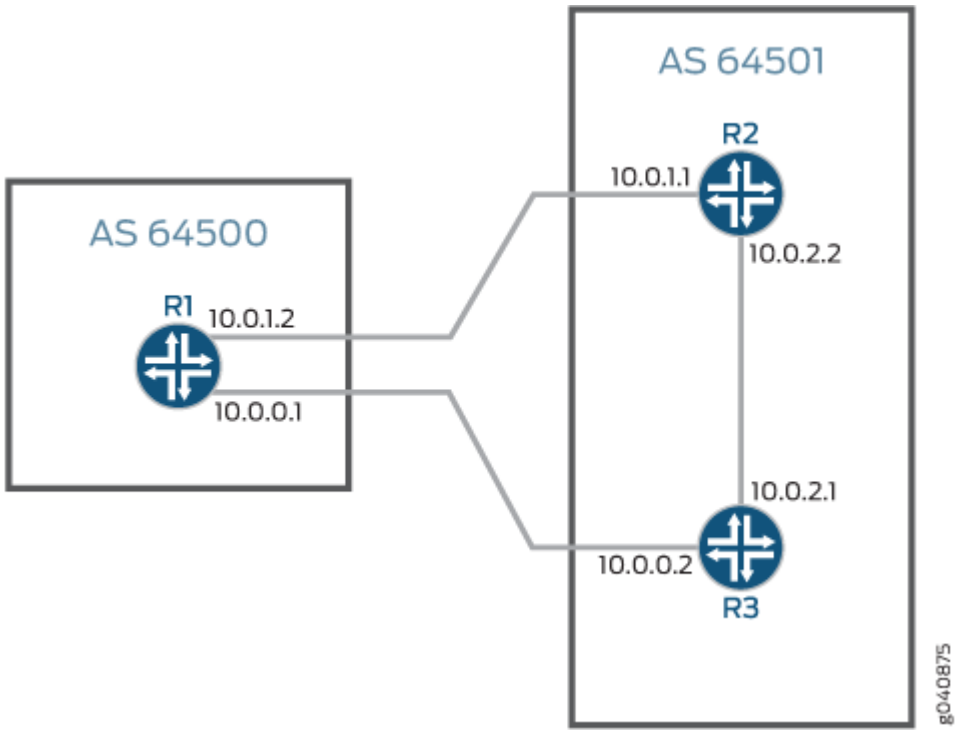
Understanding Per-Packet Load Balancing

By default, when there are multiple equal-cost paths to the same destination for the active route, Junos OS uses a hash algorithm to choose one of the next-hop addresses to install in the forwarding table. Whenever the set of next hops for a destination changes in any way, the next-hop address is re-chosen using the hash algorithm. Starting in Junos OS Release 18.3R1, for MX series routers, the default behavior for IPv6, GRE, and PPPoE packet hash computation was modified to include the flow-label field for improved load-balancing in certain cases (you can use the `no-payload` option to revert to the previous method for hash computation). See *Understanding the Algorithm Used to Load Balance Traffic on MX Series Routers* for details.

You can configure Junos OS so that, for the active route, all next-hop addresses for a destination are installed in the forwarding table. This feature is called *per-packet load balancing*. The naming may be counter-intuitive. However, Junos *per-packet* load balancing is functionally equivalent to what other vendors may term *per-flow* load balancing. You can use load balancing to spread traffic across multiple paths between routers.

Figure 11 on page 187 shows a simple load balancing scenario. Device R1 is in AS 64500 and is connected to both Device R2 and Device R3, which are in AS 64501. Device R1 can be configured to load balance traffic across the two links.

Figure 11: Simple Load Balancing Scenario



Starting in Junos OS 13.3R3, for MX Series 5G Universal Routing Platforms with modular port concentrators (MPCs) only, you can configure consistent load balancing, which prevents the reordering of all flows to active paths in an equal-cost multipath (ECMP) group when one or more next-hop paths fail. Only flows for paths that are inactive are redirected to another active next-hop path. Flows mapped to servers that remain active are maintained. This feature applies only to external BGP peers.

Starting in Junos OS Release 19.1R1,on QFX10000 switches, you can configure load balancing of IPv4 or IPv6 packets by using GPRS Tunneling Protocol-tunnel endpoint identifier (GTP-TEID) field hash calculations. The GTP-TEID hashing is added to the Layer 2 and Layer 3 field hashing that you have already configured. To enable this feature on QFX10000 switches, configure the `gtp-tunnel-endpoint-identifier` statement at the `[edit forwarding-options enhanced-hash-key family inet]` or the `[edit forwarding-options enhanced-hash-key family inet6]` hierarchy Level. GTP versions 1 and 2 are supported; they support only user data. You must use UDP port number 2152 for both GTP versions.

Change History Table

Feature support is determined by the platform and release you are using. Use Feature Explorer to determine if a feature is supported on your platform.

Release	Description
19.1R1	Starting in Junos OS Release 19.1R1,on QFX10000 switches, you can configure load balancing of IPv4 or IPv6 packets by using GPRS Tunneling Protocol-tunnel endpoint identifier (GTP-TEID) field hash calculations
18.3R1	Starting in Junos OS Release 18.3R1, for MX series routers, the default behavior for IPv6, GRE, and PPPoE packet hash computation was modified to include the flow-label field for improved load-balancing in certain cases (you can use the `no-payload` option to revert to the previous method for hash computation).
13.3R3	Starting in Junos OS 13.3R3, for MX Series 5G Universal Routing Platforms with modular port concentrators (MPCs) only, you can configure consistent load balancing, which prevents the reordering of all flows to active paths in an equal-cost multipath (ECMP) group when one or more next-hop paths fail.

RELATED DOCUMENTATION

- Example: Load Balancing BGP Traffic
- *Configuring Per-Packet Load Balancing*
- Configuring Load Balancing Based on MPLS Labels
- *Configuring Load Balancing for Ethernet Pseudowires*
- *Configuring Load Balancing Based on MAC Addresses*
- *Configuring VPLS Load Balancing Based on IP and MPLS Information*
- *Configuring VPLS Load Balancing on MX Series 5G Universal Routing Platforms*
- Configuring Consistent Load Balancing for ECMP Groups

Understanding ECMP Groups

SUMMARY

IN THIS SECTION

- Configuring Consistent Load Balancing for ECMP Groups | **189**
- Understanding Consistent Load Balancing Through Resilient Hashing on ECMP Groups | **192**

Configuring Consistent Load Balancing for ECMP Groups

Per-packet load balancing allows you to spread traffic across multiple equal-cost paths. By default, when a failure occurs in one or more paths, the hashing algorithm recalculates the next hop for all paths, typically resulting in the redistribution of all flows. *Consistent load balancing* enables you to override this behavior so that only flows for links that are inactive are redirected. All existing active flows are maintained without disruption. In a data center environment, the redistribution of all flows when a link fails potentially results in significant traffic loss or a loss of service to servers whose links remain active. Consistent load balancing maintains all active links and instead remaps only those flows affected by one or more link failures. This feature ensures that flows connected to links that remain active continue uninterrupted.

This feature applies to topologies where members of an equal-cost multipath (ECMP) group are external BGP neighbors in a single-hop BGP session. Consistent load balancing does not apply when you add a new ECMP path or modify an existing path in any way. To add a new path with minimal disruption, define a new ECMP group without modifying the existing paths. In this way, clients can be moved to the new group gradually without terminating existing connections.

- (On MX Series) Only Modular Port Concentrators (MPCs) are supported.
- Both IPv4 and IPv6 paths are supported.
- ECMP groups that are part of a virtual routing and forwarding (VRF) instance or other routing instance are also supported.
- Multicast traffic is not supported.

- Aggregated interfaces are supported, but consistent load balancing is not supported among members of the link aggregation (LAG) bundle. Traffic from active members of the LAG bundle might be moved to another active member when one or more member links fail. Flows are rehashed when one or more LAG member links fail.
- We strongly recommend that you apply consistent load balancing to no more than a maximum of 1,000 IP prefixes per router or switch.
- Layer 3 adjacency over integrated routing and bridging (IRB) interfaces is supported.

You can configure the BGP add-path feature to enable replacement of a failed path with a new active path when one or more paths in the ECMP group fail. Configuring replacement of failed paths ensures that traffic flow on the failed paths only are redirected. Traffic flow on active paths will remain unaltered.



NOTE:

- When you configure consistent load balancing on generic routing encapsulation (GRE) tunnel interfaces, you must specify the inet address of the far end GRE interface so that the Layer 3 adjacencies over the GRE tunnel interfaces are installed correctly in the forwarding table. However, ECMP fast reroute (FRR) over GRE tunnel interfaces is not supported during consistent load balancing. You can specify the destination address on the router configured with consistent load balancing at the `[edit interfaces interface name unit unit name family inet address address]` hierarchy level. For example:

  ```
  [edit interfaces]
  user@host# set interfaces gr-4/0/0 unit 21 family inet address 10.10.31.2/32
  destination 10.10.31.1
  ```

  For more information on generic routing encapsulation see "Configuring Generic Routing Encapsulation Tunneling" on page 183.

- Consistent load balancing does not support BGP multihop for EBGP neighbors. Therefore, do not enable the `multihop` option on devices configured with consistent load balancing.

To configure consistent load balancing for ECMP groups:

1. Configure BGP and enable the BGP group of external peers to use multiple paths.

2. Create a routing policy to match incoming routes to one or more destination prefixes.

```
[edit policy-options]
user@host# set policy-statement policy-statement-name from route-filter destination-prefix
orlonger
```

3. Apply consistent load balancing to the routing policy so that only traffic flows to one or more destination prefixes that experience a link failure are redirected to an active link.

```
[edit policy-options]
user@host# set policy-statement policy-statement-name then load-balance consistent-hash
```

4. Create a separate routing policy and enable per-packet load balancing.



NOTE: You must configure and apply a per-packet load-balancing policy to install all routes in the forwarding table.

```
[edit policy-options]
user@host# set policy-statement policy-statement-name then load-balance per-packet
```

5. Apply the routing policy for consistent load balancing to the BGP group of external peers.



NOTE: Consistent load balancing can be applied only to BGP external peers. This policy cannot be applied globally.

```
[edit protocols bgp]
user@host# set group group-name  import policy-statement-name
#This policy-statement-name refers to the policy created in Step 2.
```

6. (Optional) Enable bidirectional forwarding detection (BFD) for each external BGP neighbor.

```
[edit protocols bgp]
user@host# set group group-name neighbor ip-address bfd-liveness-detection milliseconds
```

NOTE: This step shows the minimum BFD configuration required. You can configure additional options for BFD.

7. Apply the per-prefix load-balancing policy globally to install all next-hop routes in the forwarding table.

```
[edit routing-options]
user@host# set forwarding table export policy-statement-name
#This policy-statement-name refers to the policy created in Step 4.
```

8. (Optional) Enable fast reroute for ECMP routes.

```
[edit routing-options]
user@host# set forwarding-table ecmp-fast-reroute
```

9. Verify the status of one or more ECMP routes for which you enabled consistent load balancing.

```
user@host> show route destination-prefix extensive
```

The output of the command displays the following flag when consistent load balancing is enabled:

```
State: <Active Ext LoadBalConsistentHash>
```

Understanding Consistent Load Balancing Through Resilient Hashing on ECMP Groups

You can use consistent load balancing to minimize flow remapping in an equal-cost multipath (ECMP) group.

By default, when there are multiple equal-cost paths to the same destination for the active route, Junos OS uses a hash algorithm to choose one of the next-hop addresses to install in the forwarding table. Whenever the set of next hops for a destination changes in any way, Junos OS rechooses the next-hop address by using the hash algorithm.

You can configure *consistent load balancing* on the switch to prevent the reordering of *all* flows to active paths in an ECMP group when one or more next-hop paths fail. Only flows for paths that are inactive are redirected to another active next-hop path. Flows mapped to servers that remain active are maintained.

This feature applies only to external BGP peers.

2

CHAPTER

Port Speed for Switches

IN THIS CHAPTER

- Port Speed Overview | **194**
- Configure Port Speed at Chassis Level and Interface Level | **196**
- Port Speed on EX Switches | **199**
- Port Speed on QFX Switches | **234**

Port Speed Overview

SUMMARY

Learn about the port speed on a switch or line card, channelization support, and the port speed configuration.

IN THIS SECTION

- Port Speed Channelization | **194**
- Port Speed Autonegotiation | **195**
- Interface Naming Conventions | **195**

Port speed is the maximum data that the line card transmits through a port in a second. You can measure port speed in kilobits per second (Kbps), gigabits per second (Gbps), or terabits per second (Tbps).

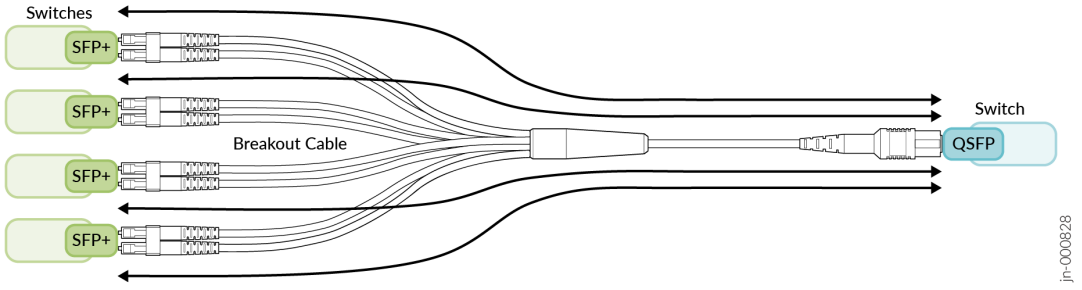
Port types:

- Switch ports: Devices plug into a switch port. You can also use the switch ports as uplink ports.
- Uplink ports: These ports are faster than switch ports. You can use uplink ports to connect two switches.
- Console ports: Use the console ports to:
  - Control a switch. The console port is usually on the rear side of a switch.
  - Program, configure, and turn on or turn off the switch.

Port Speed Channelization

You can split a high-speed port on a network equipment into several low-speed ports. For example, you can channelize a 100 Gbps port into four 25 Gbps ports.

Figure 12: Channelization of Speed



Port Speed Autonegotiation

The autonegotiation mechanism helps ports determine the optimal values of speed and duplex mode.

Autonegotiation is a signaling mechanism used by Ethernet over twisted pair. In autonegotiation, the connected devices first share their capabilities regarding these parameters and then choose the highest performance transmission mode they both support. Autonegotiation takes place when a device and a switch have different NIC cards.

Interface Naming Conventions

Each interface name includes a unique identifier and follows a naming convention. When you configure the interface, use the interface name. You can either configure a port as a single interface (non-channelized interface) or partition the port into smaller data channels or multiple interfaces (channelized interfaces).

When multiple interfaces are supported on a physical port, you use the colon (:) notation in the interface naming conventions as a delimiter to differentiate the multiple interfaces on a physical port. In the interface naming convention, xe-x/y/z:channel:

- x refers to the FPC slot number.
- y refers to the PIC slot number.
- z refers to the physical port number.
- `channel` refers to the number of channelized interfaces.

When the 40-Gigabit Ethernet interfaces (et-fpc/pic/port) are channelized as 10-Gigabit Ethernet interfaces, the interface appears in the xe-fpc/pic/port: channel format, and channel is a value of 0 through 3.

Table 24: Interface Naming Conventions

Interfaces	Non-channelized Interfaces Naming Formats	Channelized Interfaces Naming Formats
10-Gigabit Ethernet Interfaces	Prefix is xe-. The interface name appears in the xe-fpc/pic/port format.	Prefix is xe-. The interface name appears in the xe-fpc/pic/port:channel format.
25-Gigabit Ethernet Interfaces, 40-Gigabit Ethernet Interfaces, 100-Gigabit Ethernet Interfaces, 200-Gigabit Ethernet Interfaces, and 400-Gigabit Ethernet Interfaces.	Prefix is et-. The interface name appears in the et-fpc/pic/port format.	Prefix is et-. The interface name appears in the et-fpc/pic/port:channel format.

Configure Port Speed at Chassis Level and Interface Level

SUMMARY

Understand how to configure port speed at chassis and interface levels.

IN THIS SECTION

- Configure Port Speed at Chassis Level | 197
- Configure Speed at Interfaces Level | 198

Configure Port Speed at Chassis Level

Table 25: Port Speed Configuration at Chassis Level

Configuration Steps	Details	Example

Channelize Individual Port: Configure an individual port to operate at a specific channel speed. Specify a port number and channel speed.	`[edit chassis fpc fpc-slot pic pic-slot]`
	`user@host# set port port-number channel-speed speed`
		To configure an individual 40-Gigabit Ethernet (et) port to operate as 10-Gigabit Ethernet (xe) ports, specify a port number and channel speed.
		`[edit chassis fpc 0 pic 0]`
		`user@host# set port 3 channel-speed 10g`

Channelize Block of Ports: Channelize a block of ports. Specify a port range and channel speed.	`[edit chassis fpc fpc-slot pic pic-slot]`
	`user@host# set port-range port-range-low port-range-high channel-speed speed`
		To configure ports 0 through 3 on PIC 0 to operate as 50-Gigabit Ethernet ports:
		`[edit chassis fpc 0 pic 0]`
		`user@host# set port-range 0 3 channel-speed 50g`

Configure Speed per Quad: Configure port speeds only per quad (group of 4 ports) and not individually. Specify the speed for the first port of the quad ports. All ports operate at a single speed within the quad.	`[edit chassis fpc fpc-slot pic pic-slot]`
	`user@host# set port port-number speed speed`
		To configure ports 4 through 7 to operate as 25-Gigabit Ethernet ports, you must configure port 4 to operate as 25-Gigabit Ethernet ports.
		`[edit chassis fpc 0 pic 0]`
		`user@host# set port 4 speed 25g`

Configure Speed on an Individual Port	`[edit chassis fpc` *`fpc-slot`* `pic` *`pic-slot`*`]` `user@host# set port` *`port-number`* `speed` *`speed`*	`[edit chassis fpc 0 pic 0]` `user@host# set port 3 speed 25g`

Configure Speed at Interfaces Level

Table 26: Port Speed Configuration at Interfaces Level

Configuration Steps	Non-Channelized Interfaces	Channelized Interfaces

Step 1: To indicate the speed at which the ports operate, configure the speed statement for the desired interfaces.	`[edit interfaces interface-name]`
	`user@host# set speed (10G | 25G | 40G | 50G| 100G | 400G)`
	For example:
	`[edit interfaces et-1/0/3]`
	`user@host# set speed 100g`
		`[edit interfaces interface-name]`
		`user@host# set speed (10G | 25G | 40G | 50G| 100G | 400G)`
		For example:
		`[edit interfaces et-1/0/3]`
		`user@host# set speed 100g`

Step 2: To configure the speed for a group of ports.	`[edit ] user@host# wildcard range set interfaces interface-name speed speed`
	For example:
	`[edit ] user@host# wildcard range set interfaces et-1/0/[0-5] speed 100g`
		`[edit ] user@host# wildcard range set interfaces interface-name speed speed`
		For example:
		`[edit ] user@host# wildcard range set interfaces et-1/0/[7-12] speed 100g`

Step 3: To specify the number of interfaces you want to configure per port.	Not applicable	`[edit interfaces interface-name]`
		`user@host# set number-of-sub-ports number-of-sub-ports`
		For example:
		`[edit interfaces et-1/0/3]`
		`user@host# set number-of-sub-ports 4`
		In this example, in Step 1 and Step 2, you configure 4x100GE channelized interfaces.

Table 26: Port Speed Configuration at Interfaces Level *(Continued)*

Configuration Steps	Non-Channelized Interfaces	Channelized Interfaces
Step 4: (Optional) To control the number of interfaces created on a physical port, use the unused statement. If you configure a port as unused, no interfaces are created for that port irrespective of the port profile configuration for that port.	**[edit] user@host# set interfaces interface-name unused** **For example:** [edit] user@host# set interfaces et-2/0/3 unused In this example, no interfaces (channelized or non-channelized) are created on port 3 of the line card installed in the FPC slot 2.	[edit] user@host# set interfaces interface-name unused For example: [edit] user@host# set interfaces et-2/0/4 unused In this example, no interfaces (channelized or non-channelized) are created on port 4 of the line card installed in the FPC slot 2.
Step 5: Verify the configuration.	et-x/y/z { speed 100g; unit 0 { ... } ... unit N { ... } } ... et-x/y/z { unused;	et-x/y/z { speed 100g; number-of-sub-ports 4; et-x/y/z:0 { unit 0{ ... } } et-x/y/z:1 { unit 0{ ... } } et-x/y/z:2 { unit 0{ ... } } et-x/y/z:3 { unit 0{ ... } } ... et-x/y/z:6 { unused;
Step 6: Commit the configuration.		

Port Speed on EX Switches

IN THIS SECTION

- Port Speed on EX4650-48Y Switches | **200**
- Port Speed on EX4400 Switches | **203**

- Port Speed on EX4100 Switches | **216**
- Port Speed on EX4100-H Switches | **229**
- Port Speed on EX4000-12MP, EX4000-24MP, and EX4000-48MP Switches | **231**

Port Speed on EX4650-48Y Switches

For information about EX4650-48Y Switches, see EX4650 Switch Hardware Guide.

For information about platform support, see Hardware Compatibility Tool (HCT).

Table 27: EX4650-48Y Details and Description

Details	Description
FPC/PIC	FPC 0 and PIC 0; one FPC and one PIC.
QSFP/QSFP28 and SFP+ ports	Total number of ports- 56; 48 SFP+ ports and eight extension module ports.
Auto speed detection mode (Enabled by default)	If you have disabled auto-channelization, then to channelize the ports, manually configure the port speed using the `set chassis fpc` *`slot-number`* `port` *`port-number`* `channel-speed` *`speed`* command, where the speed can be set to 10 GbE or 25 GbE If a 100-Gigabit Ethernet transceiver is connected, you can only set the speed to 25GbE For the SFP+ ports, you can set the speed to 25 GbE or 1 G. There is no commit check for this, however. On EX4650 switches, the extension module ports support auto-channelization.

Table 28 on page 201 summarizes the supported port speeds on EX4650-48Y.

Table 28: Port Speed for EX4650-48Y

PIC	Port Number	Port Speed Supported	Default Speed
PIC 0	(Labeled 0 through 47) 48 SFP+ ports	1-Gigabit Ethernet 10-Gigabit Ethernet 25-Gigabit Ethernet You can configure the SFP and SFP28 port speeds only per quad (group of 4 ports) and not individually. The interface will not get created automatically on inserting 1-Gigabit Ethernet or 25-Gigabit Ethernet transceivers. You must use the CLI to configure the port speed to 1-Gigabit Ethernet or 25-Gigabit Ethernet mode manually.	10 Gbps
	(Labeled 48 through 55) 8 extension module ports	100-Gigabit Ethernet (QSFP28 ports) 40-Gigabit Ethernet (QSFP+ ports) 4x10 GbE 4x25 GbE	100 Gbps (for QSFP28 ports) 40 Gbps (for QSFP+ ports)

EX4650-48Y does not support autonegotiation when 1-gigabit fiber SFP transceiver is plugged in. In such cases, we recommend to disable auto-negotiation on the remote end device. But, EX4650-48Y switches with 1-gigabit copper SFP transceiver supports autonegotiation, as the physical layer within the transceiver handles autonegotiation.

Table 29 on page 201 lists the interface naming conventions of SFP+ ports (labeled 0 through 47) for the EX4650-48Y switch.

Table 29: Interface Naming Convention for the EX4650-48Y Switch (SFP+ Ports)

PIC	1-Gigabit Ethernet Interface	10-Gigabit Ethernet Interface	25-Gigabit Ethernet Interface
0	ge-*0*/0/0	xe-*0*/0/0	et-*0*/0/0
	ge-*0*/0/1	xe-*0*/0/1	et-*0*/0/1

Table 29: Interface Naming Convention for the EX4650-48Y Switch (SFP+ Ports) *(Continued)*

PIC	1-Gigabit Ethernet Interface	10-Gigabit Ethernet Interface	25-Gigabit Ethernet Interface
	ge-*0*/0/2	xe-*0*/0/2	et-*0*/0/2
	ge-*0*/0/3	xe-*0*/0/3	et-*0*/0/3
	ge-*0*/0/4	xe-*0*/0/4	et-*0*/0/4

Table 30 on page 202 lists the interface naming conventions of extension module ports (labeled 48 through 55) for the EX4650-48Y switch.

Table 30: Interface Naming Convention for the EX4650-48Y Switch (Extension Module Ports)

PIC	10-Gigabit Ethernet Interface	25-Gigabit Ethernet Interface	40-Gigabit Ethernet Interface	100-Gigabit Ethernet Interface
0	xe-0/0/48:*[0-3]*	et-0/0/48:*[0-3]*	et-0/0/48	et-0/0/48
	xe-0/0/49:*[0-3]*	et-0/0/49:*[0-3]*	et-0/0/49	et-0/0/49
	xe-0/0/50:*[0-3]*	et-0/0/50:*[0-3]*	et-0/0/50	et-0/0/50
	xe-0/0/51:*[0-3]*	et-0/0/51:*[0-3]*	et-0/0/51	et-0/0/51
	xe-0/0/52:*[0-3]*	et-0/0/52:*[0-3]*	et-0/0/52	et-0/0/52
	xe-0/0/53:*[0-3]*	et-0/0/53:*[0-3]*	et-0/0/53	et-0/0/53
	xe-0/0/54:*[0-3]*	et-0/0/54:*[0-3]*	et-0/0/54	et-0/0/54
	xe-0/0/55:*[0-3]*	et-0/0/55:*[0-3]*	et-0/0/55	et-0/0/55

Port Speed on EX4400 Switches

For information about EX4400 Switches, see EX4400 Switch Hardware Guide.

To view the supported transceivers, optical interfaces, and DAC cables on EX4400 switches, see Hardware Compatibility Tool (HCT).

Table 31: Port Speed for EX4400-24T and EX4400-24P

PIC	Ports	Supported Port Speeds	Default Speed
PIC 0	24 RJ-45 built-in ports (Numbered 0 through 23)	10 Mbps, 100 Mbps, and 1 Gbps Autonegotiation is supported and enabled by default.	1 Gbps
PIC 1	2x100 GbE QSFP28 ports (Numbered 0 and 1)	100 Gbps	100 Gbps
		40 Gbps	40 Gbps
		4x25 Gbps	
		4x10 Gbps	
PIC 2	1x100 GbE QSFP28 extension module (model number: EX4400-EM-1C) The extension module port can operate as a VCP using HGoE. To learn about HGoE mode, see EX4400 Switches in a Virtual Chassis. **NOTE**: EX4400 switches except EX4400-24X require System CPLD Firmware 1.0 or later installed in them to support the 1x100 GbE QSFP28 extension module. There is no CPLD	100 Gbps	100 Gbps
		40 Gbps	40 Gbps
		4x25 Gbps	

Table 31: Port Speed for EX4400-24T and EX4400-24P *(Continued)*

PIC	Ports	Supported Port Speeds	Default Speed
	upgrade that is required on EX4400-24X to support the 1x100 GbE QSFP28 extension module". See Installing and Upgrading Firmware and request system firmware upgrade for steps to upgrade the firmware.	4x10 Gbps	
	4x25 GbE SFP28 extension module (model number: EX4400-EM-4Y) When the extension module ports operate at 25 Gbps speed, you can configure them to operate as VCPs using HGoE. To learn about HGoE mode, see EX4400 Switches in a Virtual Chassis.	25 Gbps 10 Gbps 1 Gbps	25 Gbps 4x25G extension module can also support 10 GbE or 1 GbE upon setting the first port to 10 GbE/ 1 GbE using the command: `set chassis fpc <fpc-slot> pic-slot 2 port 0 speed <10g/1g>` You can revert to the default 25-gigabit mode by using `set chassis fpc <fpc-slot> pic-slot 2 port 0 speed 25g` or `delete chassis fpc <fpc-slot> pic-slot 2 port 0 speed <1g/10g>` command. All the ports in the extension module operate in the same speed.
	4x10 GbE SFP+ extension module (model number: EX4400-EM-4S)	10 Gbps	10 Gbps or 1 Gbps. The default speed depends on the plugged-in transceiver and is the default behavior. In Junos OS 24.2R2, 24.4R2, and 25.1R1 Release onwards, EX4400-EM-4S extension module supports a mix of 10G and 1G transceivers and interfaces in the default mode of operation. You do not need additional configuration to support 1G. The system automatically detects the presence of 10G or 1G transceiver

Table 31: Port Speed for EX4400-24T and EX4400-24P *(Continued)*

PIC	Ports	Supported Port Speeds	Default Speed
		1 Gbps	and creates a physical interface of the corresponding speed. To set speed explicitly to 1G or 10G, use the following command: `set chassis fpc` *`fpc-slot`* `pic` *`pic-number`* `port` *`port-number`* `speed` *`port speed`*. For example: `set chassis fpc 0 pic 2 port 0 speed 1G`. The `set chassis` command overrides the default behavior and all the ports in the PIC operates in the configured speed.

Table 32: Port Speed for EX4400-48T and EX4400-48P

PIC	Ports	Supported Port Speeds	Default Speed
PIC 0	48 RJ-45 built-in ports (Numbered 0 through 47)	10 Mbps, 100 Mbps, and 1 Gbps Autonegotiation is supported and enabled by default.	1 Gbps
PIC 1	2x100 GbE QSFP28 ports (Numbered 0 and 1)	100 Gbps	100 Gbps
		40 Gbps	40 Gbps
		4x25 Gbps	
		4x10 Gbps	

Table 32: Port Speed for EX4400-48T and EX4400-48P *(Continued)*

PIC	Ports	Supported Port Speeds	Default Speed
PIC 2	1x100 GbE QSFP28 extension module (model number: EX4400-EM-1C) The extension module port can operate as a VCP using HGoE. To learn about HGoE mode, see EX4400 Switches in a Virtual Chassis.	100 Gbps	100 Gbps
		40 Gbps	40 Gbps
		4x25 Gbps	
		4x10 Gbps	
	4x25 GbE SFP28 extension module (model number: EX4400-EM-4Y) When the extension module ports operate at 25 Gbps speed, you can configure them to operate as VCPs using HGoE. To learn about HGoE mode, see EX4400 Switches in a Virtual Chassis.	25 Gbps	25 Gbps 4x25G extension module can also support 10 GbE or 1 GbE upon setting the first port to 10 GbE/1 G using the command: `set chassis fpc <fpc-slot> pic-slot 2 port 0 speed <10g/1g>` You can revert to the default 25-gigabit mode by using `set chassis fpc <fpc-slot> pic-slot 2 port 0 speed 25g` or `delete chassis fpc <fpc-slot> pic-slot 2 port 0 speed <1g/10g>` command. All the ports in the extension module operate in the same speed.
		10 Gbps	
		1 Gbps	
	4x10 GbE SFP+ extension module (model number: EX4400-EM-4S)	10 Gbps	10 Gbps or 1 Gbps. The default speed depends on the plugged-in transceiver and is the default behavior. In Junos OS 24.2R2, 24.4R2, and 25.1R1 Release onwards, EX4400-EM-4S extension module supports a mix of 10G and 1G transceivers and interfaces in the default mode of operation. You do not need additional configuration to support 1G. The system automatically detects the presence of 10G or 1G transceiver and creates a physical interface of the corresponding speed.

Table 32: Port Speed for EX4400-48T and EX4400-48P *(Continued)*

PIC	Ports	Supported Port Speeds	Default Speed
		1 Gbps	To set speed explicitly to 1G or 10G, use the following command: `set chassis fpc` *`fpc-slot`* `pic` *`pic-number`* `port` *`port-number`* `speed` *`port speed`*. For example: `set chassis fpc 0 pic 2 port 0 speed 1G`. The `set chassis` command overrides the default behavior and all the ports in the PIC operates in the configured speed.

Table 33: Port Speed for EX4400-48F

PIC	Ports	Supported Port Speeds	Default Speed
PIC 0	36 built-in (SFP) ports (Numbered 0 through 35) Auto-negotiation is supported with both copper and fiber SFPs.	1 Gbps	1 Gbps
	12 built-in (SFP+) ports (Numbered 36 through 47) Auto-negotiation is only supported with copper SFPs for 1 Gbps speed.	10 Gbps	10 Gbps
		1 Gbps	
PIC 1	2x100 GbE QSFP28 ports Numbered 0 and 1	100 Gbps	100 Gbps
		40 Gbps	40 Gbps
		4x25 Gbps	
		4x10 Gbps	

Table 33: Port Speed for EX4400-48F *(Continued)*

PIC	Ports	Supported Port Speeds	Default Speed
PIC 2	1x100 GbE QSFP28 extension module (model number: EX4400-EM-1C) The extension module port can operate as a VCP using HGoE. To learn about HGoE mode, see EX4400 Switches in a Virtual Chassis.	100 Gbps	100 Gbps
		40 Gbps	40 Gbps
		4x25 Gbps	
		4x10 Gbps	
	4x25 GbE SFP28 extension module (model number: EX4400-EM-4Y) When the extension module ports operate at 25 Gbps speed, you can configure them to operate as VCPs using HGoE. To learn about HGoE mode, see EX4400 Switches in a Virtual Chassis.	25 Gbps	25 Gbps 4x25G extension module can also support 10 GbE or 1 GbE upon setting the first port to 10 GbE/1 GbE using the command: `set chassis fpc <fpc-slot> pic-slot 2 port 0 speed <10g/1g>` You can revert to the default 25-gigabit mode by using `set chassis fpc <fpc-slot> pic-slot 2 port 0 speed 25g` or `delete chassis fpc <fpc-slot> pic-slot 2 port 0 speed <1g/10g>` command. All the ports in the extension module operate in the same speed.
		10 Gbps	
		1 Gbps	
	4x10 GbE SFP+ extension module (model number: EX4400-EM-4S)	10 Gbps	10 Gbps or 1 Gbps. The default speed depends on the plugged-in transceiver and is the default behavior. In Junos OS 24.2R2, 24.4R2, and 25.1R1 Release onwards, EX4400-EM-4S extension module supports a mix of 10G and 1G transceivers and interfaces in the default mode of operation. You do not need additional configuration to support 1G. The system automatically detects the presence of 10G or 1G transceiver and creates a physical interface of the corresponding speed. To set speed explicitly to 1G or 10G, use the following command:

Table 33: Port Speed for EX4400-48F *(Continued)*

PIC	Ports	Supported Port Speeds	Default Speed
		1 Gbps	`set chassis fpc` *`fpc-slot`* `pic` *`pic-number`* `port` *`port-number`* `speed` *`port speed`*. For example: `set chassis fpc 0 pic 2 port 0 speed 1G`. The `set chassis` command overrides the default behavior and all the ports in the PIC operates in the configured speed.

Table 34: Port Speed for EX4400-24X

PIC	Port Number/Module	Port Speed Supported	Default Speed
PIC 0	24 (0-23) fixed ports	1 GbE	1 GbE
		10 GbE (SFP+ ports)	10 GbE
PIC 1	2x100G (network ports or virtual chassis ports)	40 GbE (QSFP28 ports)	40 GbE
		100 GbE (QSFP28 ports)	100 GbE
PIC 2 (Extension module)	4x10 GbE extension module	1 GbE and 10 GbE	10 GbE
	4x25 GbE extension module	1 GbE, 10 GbE, and 25 GbE	25 GbE

Table 34: Port Speed for EX4400-24X *(Continued)*

PIC	Port Number/Module	Port Speed Supported	Default Speed
	4x10 GbE SFP+ extension module (model number: EX4400-EM-4S)	10GbE, 1 GbE	10 Gbps or 1 Gbps. The default speed depends on the plugged-in transceiver and is the default behavior. In Junos OS 24.2R2, 24.4R2, and 25.1R1 Release onwards, EX4400-EM-4S extension module supports a mix of 10G and 1G transceivers and interfaces in the default mode of operation. You do not need additional configuration to support 1G. The system automatically detects the presence of 10G or 1G transceiver and creates a physical interface of the corresponding speed. To set speed explicitly to 1G or 10G, use the following command: `set chassis fpc` *`fpc-slot`* `pic` *`pic-number`* `port` *`port-number`* `speed` *`port speed`*. For example: `set chassis fpc 0 pic 2 port 0 speed 1G`. The `set chassis` command overrides the default behavior and all the ports in the PIC operates in the configured speed.

Note the following guidelines for EX4400-24X switches:

- You can configure PIC 1 port 0 in 100 GbE virtual chassis mode and port 1 in 100GbE network mode simultaneously and vice versa.
- You can configure PIC 1 port 0 in 40 GbE virtual chassis mode and port 1 in 40 GbE network mode simultaneously and vice versa.
- You can channelize 100GbE ports into 4x25G network ports and 40GbE network ports into 4x10G.
- Virtual chassis ports do not support channelization.
- When you change the speed of port 0 to 1 GbE, 10GbE, or 25 GbE in PIC 2, all the four ports change to the same speed.

Table 35: Port Speed for EX4400-24MP

PIC	Ports	Port Speeds Supported	Default Speed
PIC 0	24 RJ-45 built-in ports (Numbered 0 through 23)	10 Gbps 5 Gbps 2.5 Gbps 1 Gbps 100 Mbps	10 Gbps
PIC 1	2x100 GbE QSFP28 ports (Numbered 0 and 1)	100 Gbps	100 Gbps
		40 Gbps	40 Gbps
		4x25 Gbps	
		4x10 Gbps	
PIC 2	1x100 GbE QSFP28 extension module (model number: EX4400-EM-1C) The extension module port can operate as a VCP using HGoE. To	100 Gbps	100 Gbps
		40 Gbps	40 Gbps
		4x25 Gbps	

Table 35: Port Speed for EX4400-24MP *(Continued)*

PIC	Ports	Port Speeds Supported	Default Speed
	learn about HGoE mode,EX4400 Switches in a Virtual Chassis.	4x10 Gbps	
	4x25 GbE SFP28 extension module (model number: EX4400-EM-4Y) When the extension module ports operate at 25 Gbps speed, you can configure them to operate as VCPs using HGoE. To learn about HGoE mode, see EX4400 Switches in a Virtual Chassis.	25 Gbps 10 Gbps 1 Gbps	25 Gbps 4x25 GbE extension module can also support 10 GbE or 1 GbE upon setting the first port to 10 GbE/1 GbE using the command: `set chassis fpc <fpc-slot> pic-slot 2 port 0 speed <10g/1g>` You can revert to the default 25-gigabit mode by using `set chassis fpc <fpc-slot> pic-slot 2 port 0 speed 25g` or `delete chassis fpc <fpc-slot> pic-slot 2 port 0 speed <1g/10g>` command. All the ports in the extension module operate in the same speed.
	4x10 GbE SFP+ extension module (model number: EX4400-EM-4S)	10 Gbps	10 Gbps or 1 Gbps. The default speed depends on the plugged-in transceiver and is the default behavior. In Junos OS 24.2R2, 24.4R2, and 25.1R1 Release onwards, EX4400-EM-4S extension module supports a mix of 10G and 1G transceivers and interfaces in the default mode of operation. You do not need additional configuration to support 1G. The system automatically detects the presence of 10G or 1G transceiver and creates a physical interface of the corresponding speed. To set speed explicitly to 1G or 10G, use the following command: `set chassis fpc` *`fpc-slot`* `pic` *`pic-number`* `port` *`port-number`* `speed` *`port speed`*. For example: `set chassis fpc 0 pic 2 port 0 speed 1G`.

Table 35: Port Speed for EX4400-24MP *(Continued)*

PIC	Ports	Port Speeds Supported	Default Speed
		1 Gbps	The `set chassis` command overrides the default behavior and all the ports in the PIC operates in the configured speed.

Table 36: Port Speed for EX4400-48MP

PIC	Ports	Supported Port Speeds	Default Speed
PIC 0	36 RJ45 built-in ports (Numbered 0 through 35)	2.5 Gbps 1 Gbps 100 Mbps	2.5 Gbps
	12 built-in (SFP+) ports (Numbered 36 through 47)	10 Gbps 5 Gbps 2.5 Gbps 1 Gbps 100 Mbps	10 Gbps
PIC 1	2x100 GbE QSFP28 ports (Numbered 0 and 1)	100 Gbps	100 Gbps
		40 Gbps	40 Gbps
		4x25 Gbps	
		4x10 Gbps	
PIC 2	1x100 GbE QSFP28 extension module (model number: EX4400-EM-1C)	100 Gbps	100 Gbps
		40 Gbps	40 Gbps

Table 36: Port Speed for EX4400-48MP *(Continued)*

PIC	Ports	Supported Port Speeds	Default Speed
	The extension module port can operate as a VCP using HGoE. To learn about HGoE mode, see EX4400 Switches in a Virtual Chassis.	4x25 Gbps	
		4x10 Gbps	
	4x25 GbE SFP28 extension module (model number: EX4400-EM-4Y) When the extension module ports operate at 25 Gbps speed, you can configure them to operate as VCPs using HGoE. To learn about HGoE mode, see EX4400 Switches in a Virtual Chassis.	25 Gbps	25 Gbps 4x25G extension module can also support 10 GbE or 1 GbE upon setting the first port to 10 GbE/1 GbE using the command: `set chassis fpc <fpc-slot> pic-slot 2 port 0 speed <10g/1g>` You can revert to the default 25-gigabit mode by using `set chassis fpc <fpc-slot> pic-slot 2 port 0 speed 25g` or `delete chassis fpc <fpc-slot> pic-slot 2 port 0 speed <1g/10g>` command. All the ports in the extension module operate in the same speed.
		10 Gbps	
		1 Gbps	
	4x10 GbE SFP+ extension module (model number: EX4400-EM-4S)	10 Gbps	10 Gbps or 1 Gbps. The default speed depends on the plugged-in transceiver and is the default behavior. In Junos OS 24.2R2, 24.4R2, and 25.1R1 Release onwards, EX4400-EM-4S extension module supports a mix of 10G and 1G transceivers and interfaces in the default mode of operation. You do not need additional configuration to support 1G. The system automatically detects the presence of 10G or 1G transceiver and creates a physical interface of the corresponding speed. To set speed explicitly to 1G or 10G, use the following command: `set chassis fpc *fpc-slot* pic *pic-number* port *port-number* speed *port speed*.`

Table 36: Port Speed for EX4400-48MP *(Continued)*

PIC	Ports	Supported Port Speeds	Default Speed
		1 Gbps	For example: `set chassis fpc 0 pic 2 port 0 speed 1G`. The `set chassis` command overrides the default behavior and all the ports in the PIC operates in the configured speed.

NOTE:

- When the two 100 GbE QSFP28 ports of PIC 1 of EX4400 switches are configured to operate as network ports, they support channelization.
- You can channelize the 100 GbE/40 GbE ports to 4x25G and 4x10G using CLI configuration for both PIC 1 and PIC 2. See Configure Port Speed at Chassis Level and Interface Level to configure channelization.
- You can configure one port at 100 Gbps and the other at 40 Gbps at the same time, if needed.
- By default, each of the two QSFP28 ports in PIC1 of EX4400 (except EX4400-24X) is configured as two logical 50-Gbps VCP interfaces.

Use the `request virtual-chassis mode network-port` command to convert 1x100 GbE VCPs to network mode and reboot the system. Use the `request virtual-chassis mode network-port disable` command to disable network-port mode and reboot the system.

Table 37: Naming Formats for EX4400-48MP and EX4400-24MP Switches

Interfaces	Interfaces Naming Formats
100-Mbps, 1-Gbps, 2.5-Gbps, 5-Gbps, and 10-Gbps Interfaces	mge-0/0/0 mge-0/0/1

Port Speed on EX4100 Switches

The EX4100 family of switches contains the following:

- Switches: EX4100-48P, EX4100-48T, EX4100-24P, and EX4100-24T.
- Fixed switches: EX4100-F-48P, EX4100-F-48T, EX4100-F-24P, EX4100-F-24T, EX4100-F-12P, and EX4100-F-12T.
- Multigigabit switch models: EX4100-24MP and EX4100-48MP

In the switches, you can replace the power modules and fans, where in the fixed switches you cannot replace.

For information about EX4100 and 4100-F Switches, see EX4100 and EX4100-F Switch Hardware Guide.

For information about platform support, see Hardware Compatibility Tool (HCT).

Table 38 on page 217, Table 39 on page 219, Table 43 on page 225, Table 42 on page 223, Table 40 on page 221, Table 41 on page 222, and Table 44 on page 227 summarizes the supported port speeds on EX4100 switches.

Table 38: Port Speed for EX4100-48P and EX4100-48T

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
PIC 0	Downlink ports (48 ports) Port 0–47	10-Megabit Ethernet 100-Megabit Ethernet 1-Gigabit Ethernet	1-Gigabit Ethernet	Downlink ports support 1-Gbps speed with full-duplex. Both full-duplex and half-duplex are supported on 100-Mbps and 10-Mbps speeds. By default, the ports come up with 1-Gbps speed. You can configure the other port speeds at the `[edit chassis]` hierarchy level. Autonegotiation is supported and enabled by default.

Table 38: Port Speed for EX4100-48P and EX4100-48T *(Continued)*

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
PIC 1	VCPs (4 ports) Port 0-3	4x10 Gbps 4x25 Gbps- 4x1 Gbps (This speed is supported when VCP port is converted into network port along with 25GbE and 10G).	25-Gigabit Ethernet	Switches support 25-Gbps and 10-Gbps speed in both virtual chassis and network modes. See "Virtual Chassis Ports and Network Ports" on page 228. 1-Gbps speed is supported only in network mode. We support FEC74 and FEC108 (RS-FEC) standards on 25-Gigabit Ethernet network port. By default, 25 Gbps network ports have FEC74 to support the legacy EX devices, where FEC108 is configurable. Autonegotiation is not supported. The `show interfaces interface-name` command displays incorrect autonegotiation status when you disable autonegotiation. You can configure the network ports in mixed speed. For example, 2x10 GbE on ports 0 and 1, 1x1 GbE on port 2,

Table 38: Port Speed for EX4100-48P and EX4100-48T *(Continued)*

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
				and 1x25 GbE on port 3.
PIC 2	Extension module ports (4 ports) Port 0-3	4x10 Gbps 4x1 Gbps 4x100 Mbps	10-Gigabit Ethernet	Autonegotiation is not supported. You can configure the extension module ports in mixed speed. For example, port 0 with 1x100Mbps, port 1 with 1x10G, and ports 2 and 3 with 2x1G.

Table 39: Port Speed for EX4100-24P and EX4100-24T

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
PIC 0	Downlink ports (24 ports) Port 0–23	10-Mbps 100 Mbps 1 Gbps	1 Gbps	Downlink ports support 1 Gbps speed with full-duplex. Both full-duplex and half-duplex are supported on 100 Mbps and 10 Mbps speeds. By default, the ports come up with 1 Gbps speed. You can configure the other port speeds at the `[edit chassis]` hierarchy level. Autonegotiation is supported and enabled by default.

Table 39: Port Speed for EX4100-24P and EX4100-24T *(Continued)*

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
PIC 1	VCPs (4 ports) Port 0-3	4x10 Gbps 4x25 Gbps 4x1 Gbps (This speed is supported when VCP port is converted into network port along with 25 GbE and 10G).	25-Gigabit Ethernet	Switches support 25 Gbps and 10 Gbps speed in both VC and network modes. See "Virtual Chassis Ports and Network Ports" on page 228. 1 Gbps speed is supported only in network mode. We support FEC74 and FEC108 (RS-FEC) standards on 25-Gigabit Ethernet network port. By default, 25 Gbps network ports have FEC74 to support the legacy EX devices, where FEC108 is configurable. Autonegotiation is not supported. The `show interfaces interface-name` command displays incorrect autonegotiation status when you disable autonegotiation. You can configure the network ports in mixed speed. For example, 2x10G on ports 0 and 1, 1x1G

Table 39: Port Speed for EX4100-24P and EX4100-24T *(Continued)*

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
				on port 2, and 1x25G on port 3.
PIC 2	Extension module ports (4 ports) Port 0-3	4x10 Gbps 4x1 Gbps 4x100 Mbps	10 Gbps	Autonegotiation is not supported. You can configure the extension module ports in mixed speed. For example, port 0 with 1x100Mbps, port 1 with 1x10G, and ports 2 and 3 with 2x1G.

Table 40: Port Speed for EX4100-24MP

PIC	Port Number	Port Speed Supported	Default Speed
PIC 0	Downlink ports (0 - 23)	• Ports (0 - 7) are multi-rate ports that support 100 Mbps, 1 Gbps, 2.5 Gbps, 5 Gbps, and 10 Gbps. • Ports (8-23) are GigE ports that support 10 Mbps, 100 Mbps, and 1 Gbps speed.	• For multi-rate ports - 10 Gbps • For GigE ports - 1 Gbps
PIC 1	VCPs (0 - 3)	4x25 Gbps 4x10 Gbps 4x1 Gbps (This speed is supported when VCP port is converted into network port along with 25 GbE and 10 GbE. You can configure the network ports in mixed speed. For example, 2x10G on port 0 and 1, 1x1G on port 2, and 1x25G on port 3.) If you convert the VCPs to network ports, ports 0 through 3 on PIC1 support 1 Gbps, 10 Gbps, or 25 Gbps speed. See "Virtual Chassis Ports and Network Ports" on page 228.	No default value

Table 40: Port Speed for EX4100-24MP *(Continued)*

PIC	Port Number	Port Speed Supported	Default Speed
PIC 2	Extension module ports (0 - 3)	4x10 Gbps, 4x1 Gbps, or 4x100 Mbps speed. You can configure the extension module ports in mixed speed. For example, port 0 with 1x100Mbps, port 1 with 1x10G, and ports 2 and 3 with 2x1G.	No default value

Table 41: Port Speed for EX4100-48MP

PIC	Port Number	Port Speed Supported	Default Speed
PIC 0	Downlink ports (0 - 47)	• Ports (0 - 15) are multi-rate ports that support 100 Mbps, 1 Gbps, and 2.5 Gbps. • Ports (16-47) are GigE ports that support 10 Mbps, 100 Mbps, and 1 Gbps speed.	• For multi-rate ports - 2.5 Gbps • For GigE ports - 1 Gbps
PIC 1	VCPs ports (0 - 3)	4x25 Gbps 4x10 Gbps 4x1 Gbps (This speed is supported when VCP port is converted into network port along with 25 GbE and 10 GbE You can configure the network ports in mixed speed. For example, 2x10G on port 0 and 1, 1x1G on port 2, and 1x25G on port 3.) If you convert the VCPs to network ports, ports 0 through 3 on PIC1 support 1 Gbps, , 10 Gbps, or 25 Gbps speed. See "Virtual Chassis Ports and Network Ports" on page 228.	No default value
PIC 2	Extension module ports (0 - 3)	4x10-Gbps, 4x1-Gbps, or 4x100 Mbps speed. You can configure the extension module ports in mixed speed. For example, port 0 with 1x100Mbps, port 1 with 1x10G, and ports 2 and 3 with 2x1G.	No default value

On EX4100-48MP and EX4100-24MP switches, when you disable automatic MDI-X by using the `no-auto-mdix` option, automatic MDI-X is not disabled. The `show interfaces interface-name` displays incorrect auto MDI-X status when you disable auto MDI-X.

The 4xSFP28 (PIC 1) ports in EX4100 can be network ports or Virtual Chassis ports (VCPs), but not both at the same time. See EX4100 and EX4100-F System Overview for EX4100 and EX4100-F PIC terminology. If PIC 1 is in VC mode, PIC 2 can be in network mode.

Table 42: Port Speed for EX4100-F-24P and EX4100-F-24T (Fixed Switches)

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
PIC 0	Downlink ports (24 ports) Port 0–23	10-Megabit Ethernet 100-Megabit Ethernet 1-Gigabit Ethernet	1-Gigabit Ethernet	Downlink ports support 1-Gbps speed with full-duplex. Both full-duplex and half-duplex are supported on 100-Mbps and 10-Mbps speeds. By default, the ports come up with 1-Gbps speed. You can configure the other port speeds at the `[edit chassis]` hierarchy level. Autonegotiation is supported and enabled by default.

Table 42: Port Speed for EX4100-F-24P and EX4100-F-24T (Fixed Switches) *(Continued)*

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
PIC 1	VCPs (4 ports) Port 0-3	4x10 Gbps 4x1 Gbps (This speed is supported when VCP port is converted into network port along with 10 GbE)	10-Gigabit Ethernet	Fixed switches support 10-Gbps speed in both VC and network modes. See "Virtual Chassis Ports and Network Ports" on page 228. 1-Gbps speed is supported only in network mode. We support FEC74 and FEC108 (RS-FEC) standards on 25-Gigabit Ethernet network port. By default, 25-Gigabit Ethernet network ports have FEC74 to support the legacy EX devices, where FEC108 is configurable. Autonegotiation is not supported. The `show interfaces interface-name` command displays incorrect autonegotiation status when you disable autonegotiation. You can configure the network ports in mixed speed. For example, you can configure 2x10G on ports 0 and 1 and

Table 42: Port Speed for EX4100-F-24P and EX4100-F-24T (Fixed Switches) *(Continued)*

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
				2x1G on ports 2 and 3.
PIC 2	Extension module ports (4 ports) Port 0-3	100 Mbps 4x10 Gbps 4x1 Gbps	10 Gbps	Autonegotiation is not supported.

Table 43: Port Speed for EX4100-F-48P and EX4100-F-48T (Fixed Switches)

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
PIC 0	Downlink ports (48 ports) Port 0–47	10 Mbps 100 Mbps 1 Gbps	1 Gbps	Downlink ports support 1-Gbps speed with full-duplex. Both full-duplex and half-duplex are supported on 100-Mbps and 10-Mbps speeds. By default, the ports come up with 1-Gbps speed. You can configure the other port speeds at the `[edit chassis]` hierarchy level. Autonegotiation is supported and enabled by default.

Table 43: Port Speed for EX4100-F-48P and EX4100-F-48T (Fixed Switches) *(Continued)*

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
PIC 1	VCPs (4 ports) Port 0-3	4x10 Gbps 4x1 Gbps (This speed is supported when VCP port is converted into network port along with 10 GbE)	10 Gbps	Fixed switches support 10-Gbps speed in both VC and network modes. See "Virtual Chassis Ports and Network Ports" on page 228. 1-Gbps speed is supported only in network mode. We support FEC74 and FEC108 (RS-FEC) standards on 25-Gigabit Ethernet network port. By default, 25-Gigabit Ethernet network ports have FEC74 to support the legacy EX devices, where FEC108 is configurable. Autonegotiation is not supported. The `show interfaces interface-name` command displays incorrect autonegotiation status when you disable autonegotiation. You can configure the network ports in mixed speed. For example, you can configure 2x10G on ports 0 and 1 and

Table 43: Port Speed for EX4100-F-48P and EX4100-F-48T (Fixed Switches) *(Continued)*

PIC/Ports	Port Number	Speeds Supported	Default Speed	Description
				2x1G on ports 2 and 3.
PIC 2	Extension module ports (4 ports) Port 0-3	100 Mbps 4x10 Gbps 4x1 Gbps (only when you convert VCPs to extension module ports)	10 Gbps	Autonegotiation is not supported.

Table 44: Port Speed for EX4100-F-12P and EX4100-F-12T

PIC	Port Number	Port Speed Supported	Default Speed
PIC 0	Downlink ports (0 - 11)	Ports (0 - 11) are GigE ports that support 10 Mbps, 100 Mbps, and 1 Gbps speed	1 Gbps
PIC 1	VCPs ports (0 - 3)	4x10 Gbps 4x1 Gbps (This speed is supported when VCP port is converted into network port along with 10 GbE You can configure the network ports in mixed speed. For example, you can configure 2x10G on ports 0 and 1 and 2x1G on ports 2 and 3). If you convert the VCPs to network ports, ports 0 through 3 on PIC1 support 1-Gbps or 10-Gbps speed. See "Virtual Chassis Ports and Network Ports" on page 228.	No default value
PIC 2	Extension module ports (0 and 1)	2x100 Mbps or 1 Gbps speeds, 2.5 Gbps, 5 Gbps, and 10 Gbps speed.	10 Gbps

The 4xSFP+ ports in EX4100-F can be network ports or VCPs, but not both at the same time. See EX4100 and EX4100-F System Overview for EX4100 and EX4100-F PIC terminology.

The maximum MTU size supported on EX4100 switches is 9216 bytes. Packets above MTU+8 bytes are marked as oversized frames and the packets between MTU+4 and MTU+8 bytes with invalid errors.

Virtual Chassis Ports and Network Ports

You can use the `request virtual-chassis mode network-port` command to enable network port mode, which converts the default VCPs on the switch into network ports. After executing this command, you must reboot the switch for this command to take effect.

To disable network port mode and return these ports to their default settings as VCPs, use the network-port and disable options with the `request virtual-chassis mode` command. You must reboot the switch for network port mode changes to take effect, so you can include the reboot option in the same command. For example:

```
request virtual-chassis mode network-port disable reboot
```

The following are some of the guidelines for configuring the VCPs:

- In the EX4100 switches family, all four dedicated ports on PIC 1 are VCPs by default. You can convert the VCPs to network ports. The ports work either as VCPs or network ports, where mixed mode is not supported. On conversion of VC ports to network ports, all the ports dynamically detect the port speed.
- The default speed of VCPs in EX4100-48P, EX4100-48T, EX4100-24P, EX4100-24T, EX4100-24MP, and EX4100-48MP switches is 4x25G. You can convert the VCPs of these switches to network ports that operates at the speed of 25G, 10G, and 1G. You can configure the network ports in mixed speed. For example, port 0 with 1x25G, port 1 with 1x10G, and ports 2 and 3 with 2x1G.
- The default speed of VCPs in EX4100-F-48P, EX4100-F-48T, EX4100-F-24P, EX4100-F-24T, EX4100-F-12P, and EX4100-F-12T switches is 4x10G. You can convert the VCPs of EX4100-F-48P, EX4100-F-48T, EX4100-F-24P, EX4100-F-24T, EX4100-F-12P, and EX4100-F-12T switches to network ports that operates at the speed of 4x10G or 4x1G. The ports work either as VCPs or network ports, where mixed mode is not supported. But you can configure the network ports in mixed speed. For example, port 0 with 1x10G, port 2 with 1x10G, and ports 2 and 3 with 2x1G.
- In the EX4100 switches family, we support 1G speed on PIC1 in network mode only and not as a virtual chassis port.

Interface Naming Conventions

Table 45: Interface Naming Formats for EX4100 Switches

Interfaces	Interfaces Naming Formats
10-Megabit Ethernet interfaces, 100-Megabit Ethernet interfaces, and 1-Gigabit Ethernet Interfaces.	ge-0/0/x

Table 45: Interface Naming Formats for EX4100 Switches *(Continued)*

Interfaces	Interfaces Naming Formats
25-Gigabit Ethernet Interfaces	et-0/1/x
10-Gigabit Ethernet Interfaces	xe-0/2/x

SEE ALSO

request virtual-chassis mode

Port Speed on EX4100-H Switches

SUMMARY

Provides port speed, autonegotiation, and channelization information of EX4100-H switches.

To view the supported transceivers, optical interfaces, and DAC cables on EX4100-H, see Hardware Compatibility Tool (HCT).

EX4100-H-12MP includes three PICs with speeds as given below:

- PIC 0 with four 2.5 Gbps and eight 1 Gbps ports (downlink ports)
- PIC 1 with two 1 Gbps/10 Gbps ports
- PIC 2 with two 1 Gbps/10 Gbps ports (uplink ports)

Table 1 summarizes the supported port speeds on EX4100-H switches.

Table 46: Port Speed for EX4100-H Switches

PIC	Port Number and Type of Ports	Port Speed Supported	Default Speed
PIC 0	4 RJ45 ports (0-3)	100 Mbps, 1 Gbps, and 2.5 Gbps	2.5 Gbps
	8 RJ45 ports (4-11)	10 Mbps, 100 Mbps, and 1 Gbps	1 Gbps
PIC 1	2 SFP+ ports	1 Gbps and 10 Gbps	Depends on the plugged-in transceiver.
PIC 2	2 SFP+ ports (uplink ports)	1 Gbps and 10 Gbps	Depends on the plugged-in transceiver.

Table 47: Interface Naming Conventions

PIC	Interface Type	Interfaces
PIC 0	RJ45	mge-0/0/0 – mge-0/0/3
		ge-0/0/4 – ge-0/0/11
PIC 1	SFP	ge-0/1/0 – ge-0/1/1
	SFP+	xe-0/1/0 - xe-0/1/1
PIC 2	SFP	ge-0/2/0 - ge-0/2/1
	SFP+	xe-0/1/0 – xe-0/1/1

Follow these guidelines when you configure the port speed:

- You must configure applicable speeds on mge interfaces. Unsupported speeds do not reflect on the link speed of interfaces.

- Only network mode supports 1 Gbps on PIC 1. Virtual chassis (VC) mode does not support 1 Gbps.
- Always enable flow control on MACsec configured ports. Packets above MTU+8 bytes are marked as oversized frames and dropped.
- Supports maximum MTU size of 9216 bytes.
- EX4100-H-12 MP switches do not support mixed speed aggregated Ethernet (AE) link aggregation group (LAG). You can form AE LAG only if all ports in the LAG are of the same speed. For example: ge ports, mge ports, or uplink ports are of the 1G speed.
- EX4100-H-12 MP switches do not support interface hold timer in subseconds.
- The `mge` ports in PIC 0 do not support `auto-MDIX disable`.
- PIC 2 ports are not functional when PIC 1 operates in HiGig (HG) mode to form VC.

Port Speed on EX4000-12MP, EX4000-24MP, and EX4000-48MP Switches

SUMMARY

Learn about port speed, and auto negotiation information of EX4000-12MP, EX4000-24MP, and EX4000-48MP switches.

To view the supported transceivers, optical interfaces, and DAC cables on EX4000-12MP, EX4000-24MP, and EX4000-48MP, see Hardware Compatibility Tool (HCT).

Network Interfaces Support. EX4000-12MP, EX4000-24MP, and EX4000-48MP includes two PICs with speeds as given below:

- PIC 0

  Ports 0 - 11 for EX4000-12MP. The first 4 are multigigabit ports that support 100-Mbps, 1-Gbps, and 2.5-Gbps. The remaining 8 ports are GigE ports that support 10-Mbps, 100-Mbps, and 1Gbps.

  Ports 0 - 23 for EX4000-24MP. The first 4 are multigigabit ports that support 100-Mbps, 1-Gbps, and 2.5-Gbps. The remaining 20 ports are GigE ports that support 10-Mbps, 100-Mbps, and 1Gbps.

  Ports 0 - 47 for EX4000-48MP. The first 8 are multigigabit ports that support 100-Mbps, 1-Gbps, and 2.5-Gbps. The remaining 40 ports are GigE ports that support 10-Mbps, 100-Mbps, and 1Gbps.

- PIC 1

  Ports 0-3 for EX4000-12MP, EX4000-24MP and EX4000-48MP) support 1-Gbps and 10-Gbps speeds on network mode.

Table 48: Port speed support on EX4000-12MP, EX4000-24MP, and EX4000-48MP switches.

PIC	Switches	Port Number and Type of Ports	Supported Port Speed
PIC 0	EX4000-12MP	4 RJ45 Multi-Rate ports (ports 0-3)	100 Mbps, 1 Gbps, and 2.5 Gbps
		8 RJ45 GigE ports (ports 4-11)	10 Mbps, 100 Mbps, and 1 Gbps
	EX4000-24MP	4 RJ45 Multi-Rate ports (ports 0-3)	100 Mbps, 1 Gbps, and 2.5 Gbps
		20 RJ45 GigE ports (ports 4-23)	10 Mbps, 100 Mbps, and 1 Gbps
	EX4000-48MP	8 RJ45 Multi-Rate ports (ports 0-7)	100 Mbps, 1 Gbps, and 2.5 Gbps
		40 RJ45 GigE ports (ports 0-47)	10 Mbps, 100 Mbps, and 1 Gbps
PIC 1	EX4000-12MP, EX4000-24MP, and EX4000-48MP	4 SFP/SFP+ ports (ports 0-3)	1 Gbps and 10 Gbps (network mode) 10 Gbps (vcp mode)

Table 49: Interface Naming Conventions for EX4000-12MP, EX4000-24MP, and EX4000-48MP switches.

PIC	Switches	Interface Type	Interfaces
PIC 0	EX4000-12MP	RJ45	mge-0/0/0 – mge-0/0/3

			ge-0/0/4 – ge-0/0/11
	EX4000-24MP	RJ45	mge-0/0/0 – mge-0/0/3
			ge-0/0/4 – ge-0/0/23
	EX4000-48MP	RJ45	mge-0/0/0 – mge-0/0/7
			ge-0/0/8 – ge-0/0/47
PIC 1	EX4000-12MP, EX4000-24MP, and EX4000-48MP	SFP+	vcp-0/1/0 - vcp-0/1/1
		SFP/SFP+	xe-0/1/2 - xe-0/1/3
			ge-0/1/2 - ge-0/1/3

Follow these guidelines when you configure the port speed:

- You must configure applicable speeds on mge interfaces. Unsupported speeds don't reflect on the link speed of interfaces.
- Packets above MTU+8 bytes are marked as oversized frames and dropped.
- Supports maximum MTU size of 9216 bytes.
- EX4000-12MP, EX4000-24MP, and EX4000-48MP switches do not support mixed speed aggregated Ethernet (AE), link aggregation group (LAG). You can form AE LAG only if all ports in the LAG are of the same speed. For example: ge ports, mge ports, or uplink ports are of the 1G speed.
- EX4000-12MP, EX4000-24MP, and EX4000-48MP switches do not support interface hold timer in sub seconds.
- On a SFP-T port with no-auto-negotiation configuration, the user must configure appropriate MDIX settings.

SEE ALSO

Network Interfaces for EX Series

Port Speed on QFX Switches

IN THIS SECTION

- Port Speed on QFX5100-24Q Switches | 234
- Port Speed on QFX5110-48S Switches | 236
- Port Speed on QFX5120-32C Switches | 237
- Port Speed on QFX5120-48T Switches | 238
- Port Speed on QFX5120-48Y Switches | 239
- Port Speed on QFX5120-48YM Switches | 240
- Port Speed on QFX5130-32CD Switches | 241
- Port Speed on QFX5130-48C/QFX5130-48CM Switches | 243
- Port Speed on QFX5200-32C Switches | 250
- Port Speed on QFX5210-64C Switches | 250
- Port Speed on QFX5230-64CD Switches | 251
- Port Speed on QFX5240 Switches | 255
- Port Speed on QFX5700 Switches | 259

Port Speed on QFX5100-24Q Switches

To view the supported transceivers, optical interfaces, and DAC cables on QFX5100-24Q, see Hardware Compatibility Tool (HCT).

For more information on QFX5100-24Q switches, see QFX5100 Switch Hardware Guide.

Table 50: Port Speed on QFX5100-24Q Switches

PIC	Port Number	Port Speeds Supported
PIC 0	0-23 (QSFP+ ports)	40 Gbps QSFP+ supports channelization into 4x10 Gbps using breakout cables.

Use the speed command to set the speed on tri-rate copper SFP port. For information on how to configure the speed at the PIC level, see Table 2. For information on how to configure the speed at the port level, see Table 3.

Guidelines:

- In standalone mode, any of the 24 ports 0 through 23 can be configured as either uplink or access ports.
- You can use 40-Gigabit Ethernet QSFP+ transceivers and QSFP+ direct attach copper cables in any downstream port.
- You can configure up to 4 of the 40 Gbps ports as uplinks.
- The QFX5100-24Q device has two module bays for the optional expansion modules, QFX-EM-4Q or EX4600-EM-8F. QFX-EM-4Q can add a total of 8 additional QSFP+ ports to the chassis and EX4600-EM-8F can provide 8 additional 10 Gbps Enhanced SFP+ ports. The QFX-EM-4Q ports can also be configured as either access ports or as uplink ports, but only ports 0 and 2 can be channelized using port mode.
- When fully populated with two QFX-EM-4Q Expansion Modules, the QFX5100-24Q device has 128 physical ports. However, only 104 logical ports can be used for port channelization. Depending on the system mode you configure for channelization, different ports are restricted. If you attempt to channelize a restricted port, the configuration is ignored.
- Virtual Chassis and Virtual Chassis Fabric: The QFX5100-24Q device operates as a standalone switch, a member of a QFX Virtual Chassis, or as a spine or leaf device in a QFX5100 Virtual Chassis Fabric (VCF). QFX Virtual Chassis support up to 10 members. QFX5100 VCF supports 20 QFX5100 and EX4300 devices, of which four QFX5100 devices can be configured as spines.

Port Speed on QFX5110-48S Switches

IN THIS SECTION

- Virtual Chassis and Virtual Chassis Fabric | 237

To view the supported transceivers, optical interfaces, and DAC cables on QFX5110-48S, see Hardware Compatibility Tool (HCT).

For more information on QFX5110-48S switches, see QFX5110 Switch Hardware Guide.

Table 51: Port Speed on QFX5110-48S Switches

PIC	Port Number	Port Speeds Supported
PIC 0	0-47 (SFP+ ports)	100 Mbps 1 Gbps 10 Gbps
		Starting in Junos OS release 20.1R1, in addition to 1 Gbps, 10 Gbps, 40 Gbps, 100 Gbps speeds, now you can also configure 100-Mbps speed using the set interfaces interface-name speed 100M command. With QFX-SFP-1GE-T connected, you can also configure 100 Mbps on QFX5110-48S switches.

Table 51: Port Speed on QFX5110-48S Switches *(Continued)*

PIC	Port Number	Port Speeds Supported
	48-51 (QSFP28 ports)	40 Gbps 100 Gbps You can configure each port as an independent 100-GbE port or as an independent 40-GbE port. QSFP28 ports support channelization of 40 Gbps into 4x10 Gbps interfaces using breakout cables.

Use the speed command to set the speed on tri-rate copper SFP port. For information on how to configure the speed at the PIC level, see Table 2. For information on how to configure the speed at the port level, see Table 3.

Guidelines:

- On QFX5110-48S standalone switches, the FPC value is always 0.
- You cannot configure channelized interfaces to operate as Virtual Chassis ports.

Virtual Chassis and Virtual Chassis Fabric

To connect QFX5110 switches as members in a QFX5110 Virtual Chassis, you need a pair of dedicated ports on each switch and cables that link each member in the Virtual Chassis into a ring topology. Each member in the ring has at least one direct Virtual Chassis port (VCP) conn;c□on to a upstream and downstream member. QFX5110 switches are recommended in the primary, backup, or line card role. You may only mix QFX5100 members with QFX5110 members in a QFX5110 Virtual Chassis; no other QFX Series or EX Series switches are supported.

Port Speed on QFX5120-32C Switches

For information about platform support, see Hardware Compatibility Tool (HCT).

For more information on QFX5120-32C switches, see QFX5120 Switch Hardware Guide.

Table 52: Port Speed on QFX5120-32C

PIC	Port Number	Port Speed Supported
PIC 0	0-31 (QSFP28 ports)	40 Gbps 100 Gbps Supports channelization of 100 Gbps into 2x50 Gbps interfaces or 4x25 Gbps using breakout cables. Supports channelization of 40 Gbps into 4x10 Gbps interfaces using breakout cables.

Use the speed command to set the speed on tri-rate copper SFP port. For information on how to configure the speed at the PIC level, see Table 2. For information on how to configure the speed at the port level, see Table 3.

Port Speed on QFX5120-48T Switches

To view the supported transceivers, optical interfaces, and DAC cables on QFX5120-48T, see Hardware Compatibility Tool (HCT).

For more information on QFX5120-48T switches, see QFX5120 Switch Hardware Guide.

Table 53: Port Speed on QFX5120-48T Switches

PIC	Port Number	Port Speeds Supported
PIC 0	0-47 (RJ-45 ports)	1 Gbps 10 Gbps RJ-45 ports does not support channelization

Table 53: Port Speed on QFX5120-48T Switches *(Continued)*

PIC	Port Number	Port Speeds Supported
	48-53 (QSFP28 ports)	40 Gbps 100 Gbps Supports channelization of 100 Gbps into 2x50 Gbps or 4x25 Gbps interfaces. Also, 40 Gbps into 4x10 Gbps interfaces.

Use the speed command to set the speed on tri-rate copper SFP port. For information on how to configure the speed at the PIC level, see Table 2. For information on how to configure the speed at the port level, see Table 3.

Guidelines:

- Port 50 and 51 supports either 4x10G or 4x25G based on the optic used.

Port Speed on QFX5120-48Y Switches

To view the supported transceivers, optical interfaces, and DAC cables on QFX5120-48Y, see Hardware Compatibility Tool (HCT).

For more information on QFX5120-48Y switches, see QFX5120 Swich Hardware Guide.

Table 54: Port Speed on QFX5120-48Y Switches

PIC	Port Number	Port Speeds Supported
PIC 0	0-47 (SFP+ ports)	1 Gbps 10 Gbps 25 Gbps SFP+ ports do not support channelization.

Table 54: Port Speed on QFX5120-48Y Switches *(Continued)*

PIC	Port Number	Port Speeds Supported
	48-55 (QSFP28 ports)	40 Gbps 100 Gbps Supports channelization of 100 Gbps into 4x25 and 40 Gbps into 4x10 interfaces. QSFP28 ports are uplink ports and support auto channelization.

Use the speed command to set the speed on tri-rate copper SFP port. For information on how to configure the speed at the PIC level, see Table 2. For information on how to configure the speed at the port level, see Table 3.

Guidelines:

- You cannot configure channelized interfaces to operate as Virtual Chassis ports.
- QFX5120-48Y does not support autonegotiation when 1-Gbps fiber SFP transceiver is plugged in. In such cases, it is recommended to disable autonegotiation on the remote end device. But, QFX5120-48Y switches with 1-Gbps copper SFP transceiver supports autonegotiation, as the physical layer within the transceiver handles autonegotiation.

Port Speed on QFX5120-48YM Switches

To view the supported transceivers, optical interfaces, and DAC cables on QFX5120-48YM, see Hardware Compatibility Tool (HCT).

For more information on QFX5120-48YM switches, see QFX5120 Swich Hardware Guide.

Table 55: Port Speed on QFX5120-48YM Switches

PIC	Port Number	Port Speed Supported
PIC 0	0-47 (SFP28 ports)	1 Gbps 10 Gbps 25 Gbps
	48-55 (QSFP28 ports)	40 Gbps 100 Gbps
		Ports 50 and 52 support channelization. QSFP28 ports are uplink ports and support auto channelization.

Use the speed command to set the speed on tri-rate copper SFP port. For information on how to configure the speed at the PIC level, see Table 2. For information on how to configure the speed at the port level, see Table 3.

Guidelines:

- The SFP28 ports are grouped in quads (groups of four) and you can configure the speed of the ports only in quads; you cannot configure the speed for a single SFP28 port.
- Auto-channelization does not support Virtual chassis ports.
- System reboot is not required after port channelization.
- QFX5120-48YM does not support autonegotiation when 1-gigabit fiber SFP transceiver is plugged in. In such cases, it is recommended to disable auto-negotiation on the remote end device.

Port Speed on QFX5130-32CD Switches

To view the supported transceivers, optical interfaces, and DAC cables on QFX5130-32CD, see Hardware Compatibility Tool (HCT).

For more information on QFX5130-32CD switches, see QFX5130-32CD Swich Hardware Guide.

Table 56: Port Speed on QFX5130-32CD Switches

PIC	Port Number	Port Speeds Supported
PIC 0	0-31 (QSFP/QSFP28 ports)	40 Gbps 100 Gbps 200 Gbps 400 Gbps Supports channelization of 400 Gbps into 4x100 Gbps, or 2x200 Gbps, or 8x50 interfaces. Supports channelization of 200 Gbps into 2x100 Gbps interfaces. Supports channelization of 100 Gbps into 2x50 Gbps or 4x25 Gbps interfaces. Supports channelization of 40 Gbps into 4x10 Gbps interfaces.
	32-32 (SFP+ ports)	10 Gbps

Use the speed command to set the speed on tri-rate copper SFP port. For information on how to configure the speed at the PIC level, see Table 2. For information on how to configure the speed at the port level, see Table 3.

Guidelines:

- All the QSFP ports operate in 400 Gbps by default.
- QFX5130-32CD supports upto 16x400G-ZR; but when you use 400G-ZR or high power optics, you must configure the adjacent ports Unused.

Port Speed on QFX5130-48C/QFX5130-48CM Switches

IN THIS SECTION

- Interface Naming Conventions | 245
- Channelization | 245
- Supported FEC Modes | 249

To view the supported transceivers, optical interfaces, and DAC cables on QFX5130-48C/48CM, see Hardware Compatibility Tool (HCT).

QFX5130-48C/48CM supports the following port configurations:

- 48x100GbE / 50GbE / 25GbE / 10GbE on SFP-DD ports
- 8x400GbE / 200GbE / 100GbE / 40GbE on QSFP-DD ports
- 2x10GbE on SFP+ ports

See Table 57 on page 243 for details.

The QSFP-DD ports support the following channelizations:

- 4x100GbE
- 2x200GbE
- 8x50GbE
- 4x25GbE
- 4x10GbE

The SFP-DD ports support 2x50G channelization.

Table 57: Port Speed for QFX5130-48C

PIC	Ports	Optic Device	Interface Speed
PIC 0	Port 0-47(channelized Mode)	By default, all the active ports operate in 100-Gigabit Ethernet mode.	

Table 57: Port Speed for QFX5130-48C *(Continued)*

PIC	Ports	Optic Device	Interface Speed
		SFP DD 100GbE	1x100GbE
		SFP DD 50GbE	1x50GbE
		SFP DD 25GbE	1x25GbE
		SFP DD 10GbE	1x10GbE
	Port 48 – 55 (Channelized mode)	By default, all the active ports operate in 400 GbE mode.	
		QSFP DD 400GbE	1x400GbE
			4x100GbE
			2x200GbE
			8x50GbE
		QSFP+ 40GbE	1x40GbE
			4x10GbE
		QSFP28 100GbE	1x100GbE
			4x25GbE
	Ports 56 and 57 (Non-channelized mode)	By default, all active ports operate in 10 GbE mode.	

Interface Naming Conventions

Table 58: Interface Naming Conventions

PIC	Interface Type	Interfaces
PIC 0	100GbE/50GbE/25GBE/10GBE SFP-DD ports (0-47)	et-0/0/0 – et-0/0/47
	400GbE/200GbE/100GbE/40GbE QSFP-DD ports (48-55)	et-0/0/48 – et-/0/0/55
	10GbE SFP+ ports (56-57)	et-0/0/56 – et-0/0/57

Channelization

Follow the guidelines below to channelize port speeds:

QSFP-DD:

- To channelize QSFP-DD port into 8x50G, mark five SFP-DD ports as *unused*.

  Example: To channelize et-0/0/48, mark et-0/0/0 to et-0/0/4 as *unused*.

- To channelize QSFP-DD port into 4x100G, 4x25G, or 4x10G, mark one SFP-DD port as *unused*.

  Example:

  - To channelize et-0/0/49, mark et-0/0/5 as *unused*.
  - To channelize et-0/0/52, mark et-0/024 as *unused*.

- To channelize QSFP-DD port into 2x200G, do not mark any port as *unused*.

Table 59: Unused SFP Ports for 8x50G QSFP Port Channelization

Port to be channelized	Ports Unused
48	0
	1

Table 59: Unused SFP Ports for 8x50G QSFP Port Channelization *(Continued)*

Port to be channelized	Ports Unused
	2
	3
	4
49	5
	6
	7
	8
	9
50	12
	13
	14
	15
	16
51	17
	18

Table 59: Unused SFP Ports for 8x50G QSFP Port Channelization *(Continued)*

Port to be channelized	Ports Unused
	19
	20
	21
52	24
	25
	26
	27
	28
53	29
	30
	31
	32
	33
54	36
	37

Table 59: Unused SFP Ports for 8x50G QSFP Port Channelization *(Continued)*

Port to be channelized	Ports Unused
	38
	39
	40
55	41
	42
	43
	44
	45

Table 60: Unused SFP Ports for 4x100G, 4x25G, and 4x10G QSFP Port Channelization

Port to be channelized	Ports Unused
48	0
49	5
50	12
51	17
52	24
53	29

Table 60: Unused SFP Ports for 4x100G, 4x25G, and 4x10G QSFP Port Channelization *(Continued)*

Port to be channelized	Ports Unused
54	36
55	41

See Table 57 on page 243 for details.

Supported FEC Modes

See Table 61 on page 249 for supported FEC modes on different transceivers.

Table 61: FEC Modes Supported on Transceivers

Optical Transceiver	FEC Mode
SFP-DD-100GbE	FEC91-RS544
SFP-DD-50GbE	FEC91-RS544
SFP-DD-25GbE	FEC91
SFP-DD-10GbE	None
QSFP56-DD-400GbE	FEC 119
QSFP56-200GbE	FEC91-RS544
QSFP28-100GbE	FEC91
QSFP+-40GbE	None
SFP+-10GbE	None\

Port Speed on QFX5200-32C Switches

To view the supported transceivers, optical interfaces, and DAC cables on QFX5200-32C, see Hardware Compatibility Tool (HCT).

For more information on QFX5200-32C switches, see QFX5200 Swich Hardware Guide.

Table 62: Port Speed on QFX5200-32C Switches

PIC	Port Number	Port Speed Supported
PIC 0	0-31	40 Gbps 100 Gbps Supports channelization of 100 Gbps into 2x50 Gbps interfaces or 4x25 Gbps using breakout cables. Supports channelization of 40 Gbps into 4x10 Gbps interfaces using breakout cables.

Use the speed command to set the speed on tri-rate copper SFP port. For information on how to configure the speed at the PIC level, see Table 2. For information on how to configure the speed at the port level, see Table 3.

Guidelines:

- You cannot configure channelized interfaces to operate as Virtual Chassis ports.
- You can use any port as either 100 Gbps Ethernet or 40 Gbps Ethernet interfaces.
- On QFX5110-48S standalone switches, the FPC value is always 0.

Port Speed on QFX5210-64C Switches

To view the supported transceivers, optical interfaces, and DAC cables on QFX5210-64C, see Hardware Compatibility Tool (HCT).

For more information on QFX5210-64C switches, see QFX5210 Switch Hardware Guide.

Table 63: Port Speed on QFX5210-64C Switches

PIC	Port Number	Port Speeds Supported
PIC 0	0-63 (QSFP28 ports)	40 Gbps 100 Gbps Supports channelization of 100 Gbps into 2x50 Gbps interfaces or 4x25 Gbps using breakout cables. Supports channelization of 40 Gbps into 4x10 Gbps interfaces using breakout cables.

Use the speed command to set the speed on tri-rate copper SFP port. For information on how to configure the speed at the PIC level, see Table 2. For information on how to configure the speed at the port level, see Table 3.

Guidelines:

- QSFP28 ports are divided into two ranges; 0-31 as lower order ports, and 32-63 as higher order ports.
- Channelization is supported only on lower order ports 0-31.
- You can use any port as either 100 Gbps or 40 Gbps interfaces.
- The port channelization on QFX5210 switches occurs automatically when the total number of ports does not exceed 128 BCM ports and when the number of port per pipe does not exceed 32 BCM ports.

Port Speed on QFX5230-64CD Switches

SUMMARY

IN THIS SECTION

- Interface Naming Conventions | 253
- Channelization | 254

Supported FEC Modes | 254

To view the supported transceivers, optical interfaces, and DAC cables on QFX5230-64CD, see Hardware Compatibility Tool.

QFX5230-64CD has 64x400GbE/200GbE/100GbE/40GbE on QSFP56-DD ports and two SFP+ ports with 10 GbE. See Table 64 on page 252 for details.

Table 64: Port Speed for QFX5230-64CD

PIC	Ports	Optic Device	Interface Speed
PIC 0	Port 0-63 (Channelized Mode)	By default, all the active ports operate in 400-Gigabit Ethernet mode.	
		QSFP-DD	1x400GbE
			4x100GbE
			2x200GbE
			1x200GbE
			2x100GbE
		QSFP	2x100GbE
			1x200GbE
			1x100GbE
			2x50GbE
			1x50GbE

Table 64: Port Speed for QFX5230-64CD *(Continued)*

PIC	Ports	Optic Device	Interface Speed
			4x50GbE
			4x25GbE
			1x40GbE
			4x10GbE
	Ports 64 and 65 (Non-channelized mode)	By default, all active ports operate in 10 GbE mode.	
		SFP+ ports	1x10 GbE

Follow these guidelines when you configure the port speed:

- Use only the bottom 16 ports for the 400G-ZR and 400G-ZR-M. You can use the ports 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, and 63. If you use 400G-ZR and 400G-ZR-M optics on ports other than the bottom 16, you get `High power optics cannot be supported on the port` alarm.
- Optic ports that support 400G-ZR-M (33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, and 63) support 3x100 GbE channelization.

Interface Naming Conventions

Table 65 on page 253 lists the interface naming conventions for QFX5230-64CD switches.

Table 65: Interface Naming Conventions

PIC	Interface Type	Interfaces
PIC 0	400GbE/200GbE/100GbE/40GbE QSFP-DD ports (0-63)	et-0/0/0 – et-0/0/63
	10 GbE SFP+ ports (64-65)	et-0/0/64 – et-0/0/65

Channelization

QFX5230-64CD supports the following channelizations:

- 1x400GbE
- 4x100GbE
- 1x200GbE
- 2x200GbE
- 4x100GbE
- 2x100GbE
- 2x50GbE
- 1x100GbE
- 1x50GbE
- 4x25GbE
- 4x10GbE
- 1x40GbE

See Table 64 on page 252 for details.

Supported FEC Modes

Table 66 on page 254 lists the supported FEC modes

Table 66: FEC Modes Supported on Transceivers

Optical Transceiver	FEC Mode
QSFP56-DD-400GbE	FEC119
QSFP56-DD-200GbE	FEC91-RS544
QSFP28-DD-100GbE	FEC91

Table 66: FEC Modes Supported on Transceivers *(Continued)*

Optical Transceiver	FEC Mode
QSP+-40GbE	None
SFP+-10GbE	None

Port Speed on QFX5240 Switches

SUMMARY

IN THIS SECTION

- Channelization | **256**

To view the supported transceivers, optical interfaces, and DAC cables on QFX5240-64OD, see Hardware Compatibilty Tool.

QFX5240-64OD supports the following speeds:

- 800 Gbps on octal small form factor pluggable (OSFP) ports
- 10 Gbps on SFP ports

Table 67: Port Speed Details and Description for QFX5240-64OD

PIC	Number and Type of Ports	Port Speed Supported	Default Speed
PIC 0	0-64 (OSFP ports).	800 Gbps 2x400 Gbps 4x200 Gbps 8x100 Gbps	800 Gbps

Table 67: Port Speed Details and Description for QFX5240-64OD *(Continued)*

PIC	Number and Type of Ports	Port Speed Supported	Default Speed
	64-65 (SFP ports).	10 Gbps	10 Gbps

Table 68: Port Speed Details and Description for QFX5240-64QD

PIC	Number and Type of Ports	Port Speed Supported	Default Speed
PIC 0	0-64 (QSFP-DD ports).	800 Gbps 2x400 Gbps 4x200 Gbps 8x100 Gbps	800 Gbps
	64-65 (SFP ports).	10 Gbps	10 Gbps

Channelization

See Channelize Block of Ports or Individual Port.

You can channelize OSFP ports into:

- 1x800 Gbps
- 2x400 Gbps
- 4x200 Gbps
- 8x100 Gbps

Table 69: Interface Naming Conventions for QFX5240-64OD

PIC	Interface Type	Interface Speed	Interfaces
PIC 0	OSFP	1x800 Gbps	et-0/0/0-et-0/0/63

Table 69: Interface Naming Conventions for QFX5240-64OD *(Continued)*

PIC	Interface Type	Interface Speed	Interfaces
		2x400 Gbps	et-0/0/0:0 et-0/0/0:1
		4x200 Gbps	et-0/0/0:0 et-0/0/0:1 et-0/0/0:2 et-0/0/0:3
		8x100 Gbps	et-0/0/0:0 et-0/0/0:1 et-0/0/0:2 et-0/0/0:3 et-0/0/0:4 et-0/0/0:5 et-0/0/0:6 et-0/0/0:7
	SFP	1x10 Gbps	et-0/0/0

You can channelize QSFP-DD ports into:

- 1x800 Gbps
- 2x400 Gbps
- 4x200 Gbps
- 8x100 Gbps

Table 70: Interface Naming Conventions for QFX5240-64QD

PIC	Interface Type	Interface Speed	Interfaces
PIC 0	QSFP-DD	1x800 Gbps	et-0/0/0-et-0/0/63
		2x400 Gbps	et-0/0/0:0 et-0/0/0:1
		4x200 Gbps	et-0/0/0:0 et-0/0/0:1 et-0/0/0:2 et-0/0/0:3
		8x100 Gbps	et-0/0/0:0 et-0/0/0:1 et-0/0/0:2 et-0/0/0:3 et-0/0/0:4 et-0/0/0:5 et-0/0/0:6 et-0/0/0:7
	SFP	1x10 Gbps	et-0/0/0

Guidelines:

- You need not mark the corresponding pair port as "unused" to channelize one interface to 8x100 Gbps.
- You can channelize only even numbered ports into 8x100 Gbps mode.

- If you channelize any even port into 8x100 Gbps, then its paired port works only in 1x800 Gbps or 2x400 mode.

Port Speed on QFX5700 Switches

SUMMARY

Describes supported port speeds, default speeds, and channelization.

The QFX5700 switches have QFX5K-FPC-20Y line card (20x10G/25G FPC), QFX5K-FPC-16C line card (16x100GE FPC) and QFX5K-FPC-4CD line card (4x400G FPC) with support of 20x25GE SFP28 ports, 16x100GE QSFP28 and 4x400GE QSFP56-DD ports.

For information about the line card, see QFX5K-FPC-20Y, QFX5K-FPC-4CD and QFX5K-FPC-16C for QFX5700 Switches.

For information about platform support, see Hardware Compatibility Tool (HCT).

For information on QFX5700 switches, see QFX5700 Switch Hardware Guide.

3

CHAPTER

Configuring Aggregated Ethernet Interfaces

IN THIS CHAPTER

- Aggregated Ethernet Interfaces | **261**
- Load Balancing for Aggregated Ethernet Interfaces | **368**
- Dynamic Load Balancing (DLB) | **385**
- Global Load Balancing (GLB) | **413**

Aggregated Ethernet Interfaces

IN THIS SECTION

- Understanding Aggregated Ethernet Interfaces and LACP for Switches | **262**
- Forcing LAG Links or Interfaces with Limited LACP Capability to Be Up | **268**
- Configuring an Aggregated Ethernet Interface | **268**
- Configuring Tagged Aggregated Ethernet Interfaces | **270**
- Configuring Untagged Aggregated Ethernet Interfaces | **270**
- Configuring the Number of Aggregated Ethernet Interfaces on the Device (Enhanced Layer 2 Software) | **271**
- Example: Configuring Aggregated Ethernet Interfaces | **272**
- Deleting an Aggregated Ethernet Interface | **274**
- Understanding Local Link Bias | **274**
- Configuring Local Link Bias | **276**
- Understanding Local Minimum Links | **277**
- Troubleshooting an Aggregated Ethernet Interface | **280**
- Configuring Link Aggregation | **283**
- Aggregated Ethernet Link Protection | **288**
- Configure the Aggregated Ethernet Link Speed | **291**
- Configuring Periodic Rebalancing of Subscribers in an Aggregated Ethernet Interface | **294**
- Example: Configuring Aggregated Ethernet High-Speed Uplinks Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch | **295**
- Example: Configuring Link Aggregation Between a QFX Series Product and an Aggregation Switch | **305**
- Configuring Aggregated Ethernet LACP | **313**
- Configuring LACP Link Protection of Aggregated Ethernet Interfaces for Switches | **323**
- Configuring LACP Hold-UP Timer to Prevent Link Flapping on LAG Interfaces | **329**
- Verifying That LACP Is Configured Correctly and Bundle Members Are Exchanging LACP Protocol Packets | **330**
- Example: Configuring Aggregated Ethernet High-Speed Uplinks with LACP Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch | **332**
- Example: Configuring Link Aggregation with LACP Between a QFX Series Product and an Aggregation Switch | **340**

- Understanding Independent Micro BFD Sessions for LAG | **346**
- Configuring Micro BFD Sessions for LAG | **348**
- Understanding the Algorithm Used to Hash LAG Bundle and Egress Next-Hop ECMP Traffic | **355**
- Configuring the Fields in the Algorithm Used To Hash LAG Bundle and ECMP Traffic (CLI Procedure) | **363**

The below topics discuss the overview aggregated ethernet interfaces, configuration details of link aggregation and aggregated Ethernet interfaces, troubleshooting and verification of aggregated Ethernet Interfaces.

Understanding Aggregated Ethernet Interfaces and LACP for Switches

IN THIS SECTION

- Link Aggregation Group | **263**
- Link Aggregation Control Protocol (LACP) | **267**

IEEE 802.3ad link aggregation enables you to group Ethernet interfaces to form a single link layer interface, also known as a *link aggregation group (LAG)* or *bundle*.

Aggregating multiple links between physical interfaces creates a single logical point-to-point trunk link or a LAG. The LAG balances traffic across the member links within an aggregated Ethernet bundle and effectively increases the uplink bandwidth. Another advantage of link aggregation is increased availability, because the LAG is composed of multiple member links. If one member link fails, the LAG continues to carry traffic over the remaining links.

NOTE: On QFX5100, QFX5120, EX4600, QFX10002 standalone switches, and on a QFX5100 Virtual Chassis and EX4600 Virtual Chassis, you can configure a mixed rate of link speeds for the aggregated Ethernet bundle. Link speeds of 10G, 40G, and 100G are supported. QFX5200 and QFX5210 switches support mixed link speeds. QFX5200 and QFX5210 switches also support load balancing with the mixed link speeds. Load balancing does not work if you configure link speeds that are not supported.

NOTE: You can configure port channel using different SFP models between two endpoints keeping the same bandwidth.

For example:

```
switch 1 gig0/1 (SFP-10G-SR-S) --------- MX 1 gig0/1 (SFP-10G-SR-S)

switch 1 gig0/2 (SFP-10G-LR-S) --------- MX 1 gig0/2 (SFP-10G-LR-S)
```

Link Aggregation Control Protocol (LACP) is a subcomponent of the IEEE 802.3ad standard and is used as a discovery protocol.

NOTE: To ensure load balancing across the aggregated Ethernet (AE) interfaces on a redundant server Node group, the members of the AE must be equally distributed across the redundant server Node group.

NOTE: During a network Node group switchover, traffic might be dropped for a few seconds.

Link Aggregation Group

You configure a LAG by specifying the link number as a physical device and then associating a set of interfaces (ports) with the link. All the interfaces must have the same speed and be in full-duplex mode. Juniper Networks Junos operating system (Junos OS) for EX Series Ethernet Switches assigns a unique ID and port priority to each interface. The ID and priority are not configurable.

The number of interfaces that can be grouped into a LAG and the total number of LAGs supported on a switch varies according to switch model. Table 71 on page 264 lists the EX Series switches and the maximum number of interfaces per LAG and the maximum number of LAGs they support.

LAGs with member links of different interface types, for example, ge and mge are not supported on multirate switches.

NOTE: For Junos OS Evolved, the software does not impose a limit on the maximum number of AE interfaces in a mixed-rate AE bundle. Because all child logical interfaces belong to same AE physical interface and share the same selector, using much less load balance memory, mixed-rate AE interface configurations should go through even if they exceed 64 logical interfaces.

Table 71: Maximum Interfaces per LAG and Maximum LAGs per Switch (EX Series Switches)

Switch	Maximum Interfaces per LAG	Maximum LAGs
EX2200	8	32
EX2300	8	128
EX3200	8	32
EX3300 and EX3300 *Virtual Chassis*	8	32
EX3400	16	128
EX4100-F Virtual Chassis	8	128
EX4200 and EX4200 Virtual Chassis	8	111
EX4300 and EX4300 Virtual Chassis	16	128
EX4500, EX4500 Virtual Chassis, EX4550, and EX4550 Virtual Chassis	8	111
EX4400	16	128
EX4600	32	128
EX4650 Virtual Chassis	64	72
EX6200	8	111
EX8200	12	255

Table 71: Maximum Interfaces per LAG and Maximum LAGs per Switch (EX Series Switches) ***(Continued)***

Switch	Maximum Interfaces per LAG	Maximum LAGs
EX8200 Virtual Chassis	12	239
EX9200	64	150

Table 72: Maximum Interfaces per LAG and Maximum LAGs per Switch (QFX Series Switches)

Switch	Maximum Interfaces per LAG	Maximum LAGs
QFX5100	64	96
QFX5110	64	96
QFX5120	64	72
QFX5130	64	128
QFX5200	64	128
QFX5700	128	144
QFX10002	64	150
QFX10008	64	1000
QFX10016	64	1000

NOTE: On QFX Series switches, if you try to commit a configuration containing more than 64 Ethernet interfaces in a LAG, you will receive an error message saying that the group limit of 64 has been exceeded, and the configuration checkout has failed.

To create a LAG:

1. Create a logical aggregated Ethernet interface.
2. Define the parameters associated with the logical aggregated Ethernet interface, such as a logical unit, interface properties, and Link Aggregation Control Protocol (LACP).
3. Define the member links to be contained within the aggregated Ethernet interface—for example, two 10-Gigabit Ethernet interfaces.
4. Configure LACP for link detection.

Keep in mind these hardware and software guidelines:

- For Junos OS Evolved, when a new interface is added as a member to the aggregated Ethernet bundle, a link flap event is generated. When you add an interface to the bundle, the physical interface is deleted as a regular interface and then added back as a member. During this time, the details of the physical interface are lost.
- Up to 32 Ethernet interfaces can be grouped to form a LAG on a redundant server Node group, a server Node group, and a network Node group on a QFabric system. Up to 48 LAGs are supported on redundant server Node groups and server Node groups on a QFabric system, and up to 128 LAGs are supported on network Node groups on a QFabric system. You can configure LAGs across Node devices in redundant server Node groups, server Node groups, and network Node groups.

  NOTE: On a Qfabric system, if you try to commit a configuration containing more than 32 Ethernet interfaces in a LAG, you will receive an error message saying that the group limit of 32 has been exceeded, and the configuration checkout has failed.

- Up to 64 Ethernet interfaces can be grouped to form a LAG, and In a Junos Fusion, up to 1,000 LAGs are supported on QFX10002 switches acting as aggregation devices.
- The LAG must be configured on both sides of the link.
- The interfaces on either side of the link must be set to the same speed and be in full-duplex mode.

  NOTE: Junos OS assigns a unique ID and port priority to each port. The ID and priority are not configurable.

- QFabric systems support a special LAG called an FCoE LAG, which enables you to transport FCoE traffic and regular Ethernet traffic (traffic that is not FCoE traffic) across the same link aggregation bundle. Standard LAGs use a hashing algorithm to determine which physical link in the LAG is used for a transmission, so communication between two devices might use different physical links in the LAG for different transmissions. An FCoE LAG ensures that FCoE traffic uses the same physical link in the LAG for requests and replies in order to preserve the virtual point-to-point link between the

FCoE device converged network adapter (CNA) and the FC SAN switch across a QFabric system Node device. An FCoE LAG does not provide load balancing or link redundancy for FCoE traffic. However, regular Ethernet traffic uses the standard hashing algorithm and receives the usual LAG benefits of load balancing and link redundancy in an FCoE LAG. See *Understanding FCoE LAGs* for more information.

Link Aggregation Control Protocol (LACP)

LACP is one method of bundling several physical interfaces to form one logical aggregated Ethernet interface. By default, Ethernet links do not exchange LACP protocol data units (PDUs), which contain information about the state of the link. You can configure Ethernet links to actively transmit LACP PDUs, or you can configure the links to passively transmit them, sending out LACP PDUs only when the Ethernet link receives them from the remote end. The LACP mode can be active or passive. The transmitting link is known as the *actor*, and the receiving link is known as the *partner*. If the actor and partner are both in passive mode, they do not exchange LACP packets, and the aggregated Ethernet links do not come up. If either the actor or partner is active, they do exchange LACP packets. By default, LACP is in passive mode on aggregated Ethernet interfaces. To initiate transmission of LACP packets and response to LACP packets, you must enable LACP active mode. You can configure both VLAN-tagged and untagged aggregated Ethernet interfaces without LACP enabled. LACP is defined in IEEE 802.3ad, *Aggregation of Multiple Link Segments*.

LACP was designed to achieve the following:

- Automatic addition and deletion of individual links to the LAG without user intervention.
- Link monitoring to check whether both ends of the bundle are connected to the correct group.

In a scenario where a dual-homed server is deployed with a switch, the network interface cards form a LAG with the switch. During a server upgrade, the server might not be able to exchange LACP PDUs. In such a situation, you can configure an interface to be in the `up` state even if no PDUs are exchanged. Use the `force-up` statement to configure an interface when the peer has limited LACP capability. The interface selects the associated LAG by default, whether the switch and peer are both in active or passive mode. When PDUs are not received, the partner is considered to be working in the passive mode. Therefore, LACP PDU transmissions are controlled by the transmitting link.

If the remote end of the LAG link is a security device, LACP might not be supported because security devices require a deterministic configuration. In this case, do not configure LACP. All links in the LAG are permanently operational unless the switch detects a link failure within the Ethernet physical layer or data link layers.

When LACP is configured, it detects misconfigurations on the local end or the remote end of the link. Thus, LACP can help prevent communication failure:

- When LACP is not enabled, a local LAG might attempt to transmit packets to a remote single interface, which causes the communication to fail.

- When LACP is enabled, a local LAG cannot transmit packets unless a LAG with LACP is also configured on the remote end of the link.

SEE ALSO

Verifying the Status of a LAG Interface

Forcing LAG Links or Interfaces with Limited LACP Capability to Be Up

A link without Link Access Control Protocol (LACP) configuration remains down and cannot be accessed by the provider edge (PE) devices in the topology. You can configure the force-up feature in LACP on a PE device for which you need connectivity.

To ensure that the peer with limited LACP capability is up and accessible on the LAG network, configure one of the aggregated Ethernet links or interfaces on a PE device to be up by using the appropriate hierarchy level on your device:

- `set interfaces interface-name ether-options 802.3ad lacp force-up`
- `set interfaces interface-name aggregated-ether-options lacp force-up`

By default, only one link of a LAG can be in the FUP state at any time.

In a standalone or a virtual chassis environment configured with Aggregated Ethernet (AE) :

- if an aggregated Ethernet interface (AE) on a switch has multiple member links and one member link in that AE is in the force-up state with its peer's LACP down, and then if LACP comes up partially—that is, if LACP is established with a non-force-up member link—force-up is disabled on the member link on which force-up has been set, and that member link is ready for connection establishment through LACP. Force-up is eligible only if the server-side interface has LACP issues.

Configuring an Aggregated Ethernet Interface

You can associate a physical interface with an aggregated Ethernet interface.

To configure an aggregated Ethernet interface:

1. Specify that you want to configure the link aggregation group interface.

```
user@host# edit interfaces interface-name
```

2. Configure the aggregated Ethernet interface.

```
[edit interfaces interface-name]
user@host# set ether-options 802.3ad aex
```

You specify the interface instance number *x* to complete the link association; You must also include a statement defining `aex` at the `[edit interfaces]` hierarchy level. You can optionally specify other physical properties that apply specifically to the aggregated Ethernet interfaces; for details, see Ethernet Interfaces Overview.

NOTE: In general, aggregated Ethernet bundles support the features available on all supported interfaces that can become a member link within the bundle. As an exception, Gigabit Ethernet IQ features and some newer Gigabit Ethernet features are not supported in aggregated Ethernet bundles.

Gigabit Ethernet IQ and SFP interfaces can be member links, but IQ- and SFP-specific features are not supported on the aggregated Ethernet bundle even if all the member links individually support those features.

You need to configure the correct link speed for the aggregated Ethernet interface to eliminate any warning message.

NOTE: Before you commit an aggregated Ethernet configuration, ensure that link mode is not configured on any member interface of the aggregated Ethernet bundle; otherwise, the configuration commit check fails.

SEE ALSO

Aggregated Ethernet Interfaces Overview

Configuring Tagged Aggregated Ethernet Interfaces

To specify aggregated Ethernet interfaces, include the `vlan-tagging` statement at the `[edit interfaces ae`*x*`]` hierarchy level:

```
[edit interfaces aex]
vlan-tagging;
```

You must also include the `vlan-id` statement:

```
vlan-id number;
```

You can include this statement at the following hierarchy levels:

- `[edit interfaces` *interface-name* `unit` *logical-unit-number*`]`
- `[edit logical-systems` *logical-system-name* `interfaces` *interface-name* `unit` *logical-unit-number*`]`

For more information about the `vlan-tagging` and `vlan-id` statements, see 802.1Q VLANs Overview.

SEE ALSO

> *vlan-id*
>
> *vlan-tagging*

Configuring Untagged Aggregated Ethernet Interfaces

When you configure an untagged Aggregated Ethernet interface, the existing rules for untagged interfaces apply. These rules are as follows:

- You can configure only one logical interface (unit 0) on the port. The logical unit 0 is used to send and receive LACP or marker protocol data units (PDUs) to and from the individual links.
- You cannot include the `vlan-id` statement in the configuration of the logical interface.

Configure an untagged aggregated Ethernet interface by omitting the`vlan-tagging` and `vlan-id` statements from the configuration:

```
[edit interfaces]
ge-1/1/1 {
    ether-options {
        802.3ad ae0;
    }
}
ae0 {
    # vlan-tagging; OMIT FOR UNTAGGED AE CONFIGURATIONS
    unit 0 {
        # vlan-id 100; OMIT FOR UNTAGGED AE CONFIGURATIONS
        family inet {
            address 10.0.0.1/24 {
                vrrp-group 0 {
                    virtual-address 192.168.110.0;
                    priority 200;
                }
            }
        }
    }
}
```

SEE ALSO

Ethernet Interfaces User Guide for Routing Devices

Configuring the Number of Aggregated Ethernet Interfaces on the Device (Enhanced Layer 2 Software)

By default, no aggregated Ethernet interfaces are created. You must set the number of aggregated Ethernet interfaces on the routing device before you can configure them.

1. Specify that you want to access the aggregated Ethernet configuration on the device.

```
user@host# edit chassis aggregated-devices ethernet
```

2. Set the number of aggregated Ethernet interfaces.

```
[edit chassis aggregated-devices ethernet]
user@host# set device-count number
```

You must also specify the constituent physical links by including the `802.3ad` statement at the `[edit interfaces interface-name ether-options]` hierarchy level.

SEE ALSO

- Ethernet Interfaces User Guide for Routing Devices
- Junos OS Administration Library for Routing Devices

Example: Configuring Aggregated Ethernet Interfaces

Aggregated Ethernet interfaces can use interfaces from different FPCs, DPCs, or PICs. The following configuration is sufficient to get an aggregated Gigabit Ethernet interface up and running.

```
[edit chassis]
aggregated-devices {
    ethernet {
        device-count 15;
    }
}
```

```
[edit interfaces]
ge-1/3/0 {
    gigether-options {
        802.3ad ae0;
    }
}
ge-2/0/1 {
```

```
    gigether-options {
        802.3ad ae0;
    }
}
ae0 {
    aggregated-ether-options {
        link-speed 1g;
        minimum-links 1;
    }
    }
    vlan-tagging;
    unit 0 {
        vlan-id 1;

        family inet {
            address 10.0.0.1/24;
        }
    }
    unit 1 {
        vlan-id 1024;
        family inet {
            address 10.0.0.2/24;
        }
    }

    unit 2 {
        vlan-id 1025;
        family inet {
            address 10.0.0.3/24;
        }
    }
    unit 3 {
        vlan-id 4094;

        family inet {
            address 10.0.0.4/24;
        }
    }
}
```

SEE ALSO

Configure 'link-speed' for Gigabit Ethernet based Aggregate Ethernet interface bundles

Deleting an Aggregated Ethernet Interface

There are two approaches to deleting an aggregated Ethernet interface:

- You can delete an aggregated Ethernet interface from the interface configuration. The Junos OS removes the configuration statements related to ae*x* and sets this interface to down state.
- You can also permanently remove the aggregated Ethernet interface from the device configuration by deleting it from the device-count on the routing device.

To delete an aggregated Ethernet interface:

1. Delete the aggregated Ethernet configuration.

   This step changes the interface state to down and removing the configuration statements related to ae*x*.

```
[edit]
user@host# delete interfaces aex
```

2. Delete the interface from the device count.

```
[edit]
user@host# delete chassis aggregated-devices ethernet device-count
```

SEE ALSO

Load Balancing on Aggregated Ethernet Interfaces

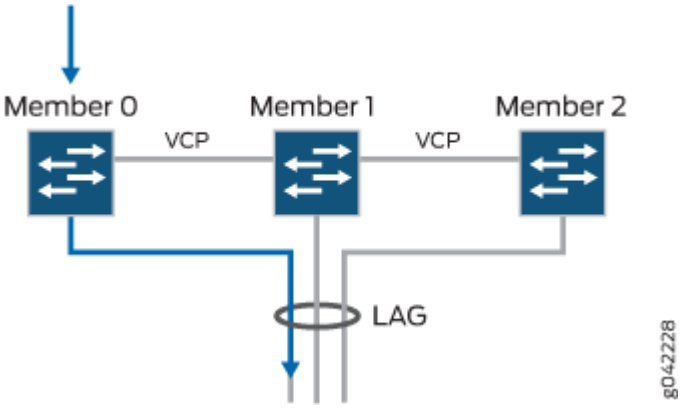
Periodic Packet Management

Understanding Local Link Bias

Local link bias conserves bandwidth on Virtual Chassis ports (VCPs) by using local links to forward unicast traffic exiting a Virtual Chassis or Virtual Chassis Fabric (VCF) that has a Link Aggregation group

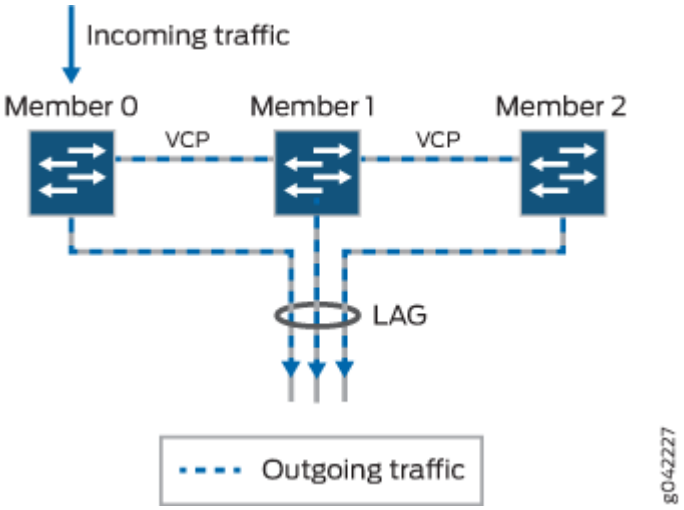
(LAG) bundle composed of member links on different member switches in the same Virtual Chassis or VCF. A local link is a member link in the LAG bundle that is on the member switch that received the traffic. Because traffic is received and forwarded on the same member switch when local link bias is enabled, no VCP bandwidth is consumed by traffic traversing the VCPs to exit the Virtual Chassis or VCF using a different member link in the LAG bundle. The traffic flow of traffic exiting a Virtual Chassis or VCF over a LAG bundle when local link bias is enabled is illustrated in Figure 13 on page 275.

Figure 13: Egress Traffic Flow with Local Link Bias



When local link bias is disabled, egress traffic exiting a Virtual Chassis or VCF on a LAG bundle can be forwarded out of any member link in the LAG bundle. Traffic forwarding decisions are made by an internal algorithm that attempts to load-balance traffic between the member links in the bundle. VCP bandwidth is frequently consumed by egress traffic when local link bias is disabled because the egress traffic traverses the VCPs to reach the destination egress member link in the LAG bundle. The traffic flow of traffic exiting a Virtual Chassis or VCF over a LAG bundle when local link bias is disabled is illustrated in Figure 14 on page 276.

Figure 14: Egress Traffic Flow without Local Link Bias



Starting in Junos OS Release 14.1X53-D25, local link bias can be enabled globally for all LAG bundles in a Virtual Chassis or VCF, or individually per LAG bundle in a Virtual Chassis. In prior Junos OS releases, local link bias could be enabled individually per LAG bundle only.

A Virtual Chassis or VCF that has multiple LAG bundles can contain bundles that have and have not enabled local link bias. Local link bias only impacts the forwarding of unicast traffic exiting a Virtual Chassis or VCF; ingress traffic handling is not impacted by the local link bias setting. Egress multicast, unknown unicast, and broadcast traffic exiting a Virtual Chassis or VCF over a LAG bundle is not impacted by the local link bias setting and is always load-balanced among the member links. Local link bias is disabled, by default.

You should enable local link bias if you want to conserve VCP bandwidth by always forwarding egress unicast traffic on a LAG bundle out of a local link. You should not enable local link bias if you want egress traffic load-balanced across the member links in the LAG bundle as it exits the Virtual Chassis or VCF.

Configuring Local Link Bias

Local link bias is used to conserve bandwidth on Virtual Chassis ports (VCPs) by using local links to forward unicast traffic exiting a Virtual Chassis or Virtual Chassis Fabric (VCF) that has a Link Aggregation group (LAG) bundle composed of member links on different member switches in the same Virtual Chassis or VCF. A local link is a member link in the LAG bundle that is on the member switch that received the traffic. Because traffic is received and forwarded on the same member switch when local link bias is enabled, no VCP bandwidth is consumed by traffic traversing the VCPs to exit the Virtual Chassis or VCF on a different member link in the LAG bundle.

You should enable local link bias if you want to conserve VCP bandwidth by always forwarding egress unicast traffic on a LAG out of a local link. You should not enable local link bias if you want egress traffic load-balanced as it exits the Virtual Chassis or VCF.

Local link bias can be enabled or disabled globally or per LAG bundle on a Virtual Chassis or VCF. In cases where local link bias is enabled at both the global and per LAG bundle levels, the per LAG bundle configuration takes precedence. For instance, if local link bias is enabled globally but disabled on a LAG bundle named **ae1**, local link bias is disabled on the LAG bundle named **ae1**.

To enable local link bias on a LAG bundle:

```
[edit]
user@switch# set interface aex aggregated-ether-options local-bias
```

where aex is the name of the aggregated Ethernet link bundle.

For instance, to enable local link bias on aggregated Ethernet interface ae0:

```
[edit]
user@switch# set interface ae0 aggregated-ether-options local-bias
```

Understanding Local Minimum Links

IN THIS SECTION

- Configuring Local Minimum Links | **279**
- Local Minimum Links Effect on LAG Minimum Links | **280**
- Local Minimum Links and Local Link Bias | **280**

NOTE: When describing the local minimum links feature, *member links* are links that are part of an aggregated Ethernet bundle (LAG), *member switches* are chassis that are members in a Virtual Chassis or Virtual Chassis Fabric (VCF), and *local member links* (or

simply *local links*) are member links of the same LAG that are local to a particular Virtual Chassis or VCF member switch.

A link aggregation group (LAG) can include member links on different chassis, and multiple local member links on member switches in a Virtual Chassis or VCF. If member links in the LAG fail, the LAG continues to carry traffic over the remaining member links that are still active. When multiple member links are local to one chassis and one or more of those links fail, LAG traffic coming into that chassis will be redistributed over the remaining local links. However, the remaining active local links can suffer traffic loss if the failed links result in sufficiently reduced total bandwidth through the chassis.

Introduced in Junos OS Release 14.1X53-D40, the local minimum links feature helps avoid traffic loss due to asymmetric bandwidth on LAG forwarding paths through a Virtual Chassis or VCF member switch when one or more local member links have failed.



NOTE: The local minimum links feature is supported on Virtual Chassis or VCFs with QFX5100 member switches only.

Based on a user-configured threshold value, when one or more member links fail, this feature marks any remaining active local links as "down," forcing LAG traffic to be redistributed only through member links on *other* chassis. To enable this feature on a particular aggregated Ethernet interface (ae*x*), you set the `local-minimum-links-threshold` configuration statement with a threshold value that represents the percentage of local member links that must be up on a chassis for *any* local member links on that chassis to continue to be active in the aggregated Ethernet bundle.

The configured threshold value:

- Applies to a specified aggregated Ethernet interface.
- Applies to any chassis that has links in the specified aggregated Ethernet bundle.
- Represents a percentage of active local member links out of the total number of local member links for the chassis.

When the local minimum links feature is enabled for a LAG, if one or more member links on a chassis fail, the feature compares the percentage of local member links that are still up to the threshold. If the percentage of "up" links is less than the threshold, the feature forces down the remaining active local links, and no traffic for the aggregated Ethernet interface will be forwarded through the member links on that chassis. If the percentage of links that are "up" is greater than or equal to the threshold, the status of the active links remains unchanged, and LAG traffic will continue to be distributed over available member links on that chassis.

For example, consider a member switch in a Virtual Chassis Fabric that has four links that are active member links of a LAG, and the local minimum links feature is enabled with the threshold set to 60:

- If one member link goes down, 75 percent (three out of four) of the links are still up, which is greater than the threshold (60 percent), so the remaining links stay up.
- If two member links go down, only 50 percent (two out of four) of the links are “up”, so the local minimum links feature forces the remaining two active links “down.” The same is true if three member links fail, the remaining link is forced down as well.

The local minimum links feature tracks whether links are down because the link failed or the link was forced down, as well as when active, failed, or forced-down member links are added or removed. As a result, the feature can respond dynamically when:

- Failed local member links come back up.
- You change the configured threshold value, or you disable the local minimum links feature.
- Adding or removing local member links changes the total number of local member links, or changes the ratio of “up” links to total local member links as compared to the threshold.

For example, if a failed member link causes all local member links to be forced down, then that link comes back up and brings the percentage of “up” links above the current threshold, the system adjusts the status of the forced-down links to mark them up again as well.

You should enable this feature only if your system closely manages ingress and egress traffic forwarding paths on LAGs for individual chassis in a Virtual Chassis and VCFs, especially where local link bias is also enabled.

Configuring Local Minimum Links

The local minimum links feature is disabled by default. To enable this feature for a LAG bundle (which then applies to any chassis that has local member links in the LAG), simply configure a threshold value for the LAG interface, as follows:

```
[edit interfaces]
user@switch# set aggregated-ether-options aex local-minimum-links-threshold threshold-value
```

To update the threshold value, use the same command with the new threshold value.

To disable the local minimum links feature, delete the `local-minimum-links-threshold` statement from the configuration. Any links that were forced down by this feature are automatically brought up again within a few seconds.

Local Minimum Links Effect on LAG Minimum Links

The per-chassis local minimum links threshold is similar to the *minimum-links* setting for a LAG bundle, which configures the minimum number of member links in the bundle that should be up for the aggregated Ethernet interface as a whole to be considered "up." Local member links that fail or are forced down by the local minimum links feature contribute to the count of "up" links for the LAG as a whole. As a result, this feature can cause the entire LAG to be brought down if enough local links are forced down. Enabling and configuring the local minimum links feature is independent of LAG minimum links configuration, but you should carefully consider the combined potential effect on the LAG as a whole when configuring both features.

Local Minimum Links and Local Link Bias

The local minimum links and local link bias features operate independently, but can influence each other's traffic forwarding results. For example, when local link bias is enabled and would otherwise favor forwarding traffic out of local links in the aggregated Ethernet bundle, but those links are down because the local minimum links threshold is not currently met, outgoing traffic will be redirected through the VCPs to other Virtual Chassis or VCF member switches for forwarding. In that case, unanticipated increased VCP traffic can impact Virtual Chassis or VCF performance.

Troubleshooting an Aggregated Ethernet Interface

IN THIS SECTION

- Show Interfaces Command Shows the LAG is Down | **281**
- Logical Interface Statistics Do Not Reflect All Traffic | **281**
- IPv6 Interface Traffic Statistics Are Not Supported | **282**
- SNMP Counters ifHCInBroadcastPkts and ifInBroadcastPkts Are Always 0 | **282**

Troubleshooting issues for aggregated Ethernet interfaces:

Show Interfaces Command Shows the LAG is Down

IN THIS SECTION

- Problem | 281
- Solution | 281

Problem

Description

The `show interfaces terse` command shows that the LAG is down.

Solution

Check the following:

- Verify that there is no configuration mismatch.
- Verify that all member ports are up.
- Verify that a LAG is part of family ethernet—switching (Layer 2 LAG) or family inet (Layer 3 LAG).
- Verify that the LAG member is connected to the correct LAG at the other end.
- Verify that the LAG members belong to the same switch (or the same Virtual Chassis).

Logical Interface Statistics Do Not Reflect All Traffic

IN THIS SECTION

- Problem | 282
- Solution | 282

Problem

Description

The traffic statistics for a logical interface do not include all of the traffic.

Solution

Traffic statistics fields for logical interfaces in `show interfaces` commands show only control traffic; the traffic statistics do not include data traffic. You can view the statistics for all traffic only per physical interface.

IPv6 Interface Traffic Statistics Are Not Supported

IN THIS SECTION

- Problem | 282
- Solution | 282

Problem

Description

The `IPv6 transit statistics` in the `show interfaces` command display all `0` values.

Solution

EX Series switches do not support the collection and reporting of IPv6 transit statistics.

SNMP Counters ifHCInBroadcastPkts and ifInBroadcastPkts Are Always 0

IN THIS SECTION

- Problem | 283
- Solution | 283

Problem

Description

The values for the SNMP counters ifHCInBroadcastPkts and ifInBroadcastPkts are always 0.

Solution

The SNMP counters ifHCInBroadcastPkts and ifInBroadcastPkts are not supported for aggregated Ethernet interfaces on EX Series switches.

RELATED DOCUMENTATION

- Verifying the Status of a LAG Interface

Configuring Link Aggregation

IN THIS SECTION

- Creating an Aggregated Ethernet Interface | **284**
- Configuring the VLAN Name and VLAN ID Number | **285**
- Configuring Aggregated Ethernet LACP (CLI Procedure) | **285**

Use the link aggregation feature to aggregate one or more links to form a virtual link or aggregation group. The MAC client can treat this virtual link as if it were a single link. Link aggregation increases bandwidth, provides graceful degradation as failure occurs, and increases link availability.

NOTE: An interface with an already configured IP address cannot form part of the aggregation group.

NOTE: On QFX5100, QFX5120, QFX5200, EX4600, QFX10002, and QFX10008 standalone switches and on QFX5100 Virtual Chassis and EX4600 Virtual Chassis, you can configure a mixed rate of link speeds for the aggregated Ethernet bundle. Load

balancing will not work if you configure link speeds that are not supported. (Platform support depends on the Junos OS release in your installation.)

Creating an Aggregated Ethernet Interface

To create an aggregated Ethernet interface:

1. Specify the number of aggregated Ethernet interfaces to be created:

```
[edit chassis]
user@switch# set aggregated-devices interfaces device-count device-count
```

For example, to specify 5:

```
[edit chassis]
user@switch#  set aggregated-devices interfaces device-count 5
```

2. Specify the minimum number of links for the aggregated Ethernet interface (ae*x*), that is, the defined bundle, to be labeled “up”:

NOTE: By default only one link must be up for the bundle to be labeled “up”.

```
[edit interfaces]
user@switch# set interface-name aggregated-ether-options minimum-links minimum-links
```

For example, to specify 5:

```
[edit interfaces]
user@switch#  set interface-name aggregated-ether-options minimum-links 5
```

3. Specify the link speed for the aggregated Ethernet bundle:

```
[edit interfaces]
user@switch# set interface-name aggregated-ether-options link-speed link-speed
```

For example, to specify 10g:

```
[edit interfaces]
user@switch#  set interface-name aggregated-ether-options link-speed 10g
```

4. Specify the members to be included within the aggregated Ethernet bundle:

```
[edit interfaces]
user@switch# set interface-name ether-options 802.3ad aex
user@switch# set interface-name ether-options 802.3ad aex
```

Configuring the VLAN Name and VLAN ID Number



NOTE: VLANs are not supported on OCX Series switches.

```
[edit vlans]
user@switch# set vlan-name vlan-id vlan-id-number
```

For example, 100.



NOTE: When you add or remove a vlan from a LAG interface, the interface goes down and comes back (flaps). The flapping happens when a low speed SFP is plugged into a relatively high speed port. To avoid flapping, configure the port speed to match the speed of the SFP.

Configuring Aggregated Ethernet LACP (CLI Procedure)

For aggregated Ethernet interfaces on EX Series switches, you can configure the Link Aggregation Control Protocol (LACP). LACP is one method of bundling several physical interfaces to form one logical interface. You can configure aggregated Ethernet interfaces with or without LACP enabled.

LACP was designed to achieve the following:

- Automatic addition and deletion of individual links to the bundle without user intervention
- Link monitoring to check whether both ends of the bundle are connected to the correct group



NOTE: You can also configure LACP link protection on aggregated Ethernet interfaces. For information, see "Configuring LACP Link Protection of Aggregated Ethernet Interfaces for Switches" on page 323.

The Junos OS implementation of LACP provides link monitoring but not automatic addition and deletion of links.

Before you configure LACP for EX Series, be sure you have:

- Configured the aggregated Ethernet bundles—also known as link aggregation groups (LAGs). See Configuring Aggregated Ethernet Links (CLI Procedure)

When LACP is enabled, the local and remote sides of the aggregated Ethernet links exchange protocol data units (PDUs), which contain information about the state of the link. You can configure Ethernet links to actively transmit PDUs, or you can configure the links to passively transmit them (sending out LACP PDUs only when they receive them from another link). One side of the link must be configured as `active` for the link to be up.



NOTE: Do not add LACP to a LAG if the remote end of the LAG link is a security device, unless the security device supports LACP. Security devices often do not support LACP because they require a deterministic configuration.

To configure LACP:

1. Enable the LACP mode:

```
[edit interfaces]
user@switch# set aex aggregated-ether-options lacp mode
```

For example, to specify the mode as active, execute the following command:

```
[edit interfaces]
user@switch# set aex aggregated-ether-options lacp active
```



NOTE: LACP decides active and back up state of links. When configuring LACP, state of the backup link should not be configured manually as down. The following command is not supported if LACP is configured:`set interfaces` **ae0** `aggregated-ether-options link-protection backup-state` *down*

2. Specify the interval and speed at which the interfaces send LACP packets:

```
[edit interfaces]
user@switch# set aex aggregated-ether-options lacp periodic interval
```

For example, to specify the interval as fast, execute the following command:

```
[edit interfaces]
user@switch# set aex aggregated-ether-options lacp periodic fast
```

3. (Optional) A link without Link Access Control Protocol (LACP) configuration remains down and cannot be accessed by the provider edge (PE) devices in the topology. Configure the `force-up` feature in LACP on a PE device for which you need connectivity.

```
[edit interfaces]
user@switch# set interfaces aex aggregated-ether-options lacp force-up
```

NOTE: The LACP process exists in the system only if you configure the system in either active or passive LACP mode.

SEE ALSO

- Configuring Aggregated Ethernet Links (CLI Procedure)
- Configuring LACP Link Protection of Aggregated Ethernet Interfaces for Switches | 323
- Configuring Aggregated Ethernet Interfaces (J-Web Procedure)
- *Example: Configuring Aggregated Ethernet High-Speed Uplinks with LACP Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch*
- *Example: Configuring Aggregated Ethernet High-Speed Uplinks Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch*
- Verifying the Status of a LAG Interface

RELATED DOCUMENTATION

- Understanding Interface Naming Conventions | 11
- *Configuring an FCoE LAG*

Example: Configuring Link Aggregation Between a QFX Series Product and an Aggregation Switch | 305

Verifying the Status of a LAG Interface

Verifying That LACP Is Configured Correctly and Bundle Members Are Exchanging LACP Protocol Packets | 330

show lacp statistics interfaces (View)

Aggregated Ethernet Link Protection

IN THIS SECTION

- Configuring Link Protection for Aggregated Ethernet Interfaces | 289
- Configuring Primary and Backup Links for Link Aggregated Ethernet Interfaces | 289
- Reverting Traffic to a Primary Link When Traffic is Passing Through a Backup Link | 290
- Disabling Link Protection for Aggregated Ethernet Interfaces | 291

You can configure link protection for aggregated Ethernet interfaces to provide QoS on the links during operation.

On aggregated Ethernet interfaces, you designate a primary and backup link to support link protection. Egress traffic passes only through the designated primary link. This includes transit traffic and locally generated traffic on the router or switch. When the primary link fails, traffic is routed through the backup link. Because some traffic loss is unavoidable, egress traffic is not automatically routed back to the primary link when the primary link is reestablished. Instead, you manually control when traffic should be diverted back to the primary link from the designated backup link.

Use Feature Explorer to confirm platform and release support for specific features.

Use the following table to review platform-specific behaviors for your platform:

Platform-Specific Link Protection Behavior

Platform	Difference
ACX Series	• ACX7000 Series routers that support link protection can operate in revertive mode of link protection. Enable auto-revertive operation using the following command: `set interfaces ae1 aggregated-ether-options link-protection revertive`

Configuring Link Protection for Aggregated Ethernet Interfaces

Aggregated Ethernet interfaces support link protection to ensure QoS on the interface.

To configure link protection:

1. Specify that you want to configure the options for an aggregated Ethernet interface.

```
user@host# edit interfaces aex aggregated-ether-options
```

2. Configure the link protection mode.

```
[edit interfaces aex aggregated-ether-options]
user@host# set link-protection
```

SEE ALSO

link-protection

aggregated-ether-options

Configuring Primary and Backup Links for Link Aggregated Ethernet Interfaces

To configure link protection, you must specify a primary and a secondary, or backup, link.

To configure a primary link and a backup link:

1. Configure the primary logical interface.

```
[edit interfaces interface-name]
user@host# set (fastether-options | gigether-options) 802.3ad aex primary
```

2. Configure the backup logical interface.

```
[edit interfaces interface-name]
user@host# set (fastether-options | gigether-options) 802.3ad aex backup
```

SEE ALSO

> *802.3ad*

Reverting Traffic to a Primary Link When Traffic is Passing Through a Backup Link

On aggregated Ethernet interfaces, you designate a primary and backup link to support link protection. Egress traffic passes only through the designated primary link. This includes transit traffic and locally generated traffic on the router or switch. When the primary link fails, traffic is routed through the backup link. Because some traffic loss is unavoidable, egress traffic is not automatically routed back to the primary link when the primary link is reestablished. Instead, you manually control when traffic should be diverted back to the primary link from the designated backup link.

To manually control when traffic should be diverted back to the primary link from the designated backup link, enter the following operational command:

```
user@host> request interface revert aex
```

SEE ALSO

> *request interface (revert | switchover) (Aggregated Ethernet Link Protection)*

Disabling Link Protection for Aggregated Ethernet Interfaces

To disable link protection, issue the `delete interfaces ae`*x* `aggregated-ether-options link-protection` configuration command.

```
user@host# delete interfaces aex aggregated-ether-options link-protection
```

SEE ALSO

> *request interface (revert | switchover) (Aggregated Ethernet Link Protection)*

Configure the Aggregated Ethernet Link Speed

IN THIS SECTION

- Platform-Specific LAG Behavior | 292

On aggregated Ethernet interfaces, you can set the required link speed for all interfaces included in the bundle.

Some devices support mixed rates and mixed modes. For example, you could configure the following on the same aggregated Ethernet (AE) interface:

- Member links of different modes (WAN and LAN) for 10-Gigabit Ethernet links
- Member links of different rates: 10-Gigabit Ethernet, 25-Gigabit Ethernet, 40-Gigabit Ethernet, 50-Gigabit Ethernet, 100-Gigabit Ethernet, 400-Gigabit Ethernet, and OC192 (10-Gigabit Ethernet WAN mode)

Use Feature Explorer to confirm platform and release support for specific features.

Review the "Platform-Specific LAG Behavior" on page 292 section for notes related to your platform.

NOTE:

- You can only configure 50-Gigabit Ethernet member links using the 50-Gigabit Ethernet interfaces of 100-Gigabit Ethernet PIC with CFP (PD-1CE-CFP-FPC4).
- You can only configure 100-Gigabit Ethernet member links using the two 50-Gigabit Ethernet interfaces of a 100-Gigabit Ethernet PIC with CFP. You can include this 100-Gigabit Ethernet member link in an aggregated Ethernet link that includes member links of other interfaces as well.

Platform-Specific LAG Behavior

Platform	Difference

ACX Series	• ACX7000 Series routers that support LAG can operate in mixed mode. Configure the following on the same aggregated Ethernet interface:
	  • Member links of different modes (WAN and LAN) with the same speed
	  • Member links of different modes (WAN and LAN) with different speeds
	• ACX7000 Series routers support two modes of LAG configuration:
	  • Maximum AE children 16 - 256 AE bundles
	  • Maximum AE children 64 - 64 AE bundles
	• ACX7000 Series routers use `ether-options` instead of `gigether-options`.

1. Specify that you want to configure the aggregated Ethernet options for the aggregated Ethernet interface.

```
[edit]
user@host# edit interfaces interface-name aggregated-ether-options
```

For example:

```
[edit]
user@host# edit interfaces ae0 aggregated-ether-options
```

2. Configure the link speed.

```
[edit interfaces interface-name aggregated-ether-options]
user@host# set link-speed speed
```

For example, to set the link speed of all member links of the aggregated Ethernet interface to 10 Gbps:

```
[edit interfaces ae0 aggregated-ether-options]
user@host# set link-speed 10g
```

3. (Optional) If you plan to configure the link speed of the member links to be different speeds, set the link speed for the aggregated Ethernet interface to `mixed`.

```
[edit interfaces interface-name aggregated-ether-options]
user@host# set link-speed mixed
```

For example:

```
[edit interfaces ae0 aggregated-ether-options]
user@host# set link-speed mixed
```

NOTE: The QFX5000 line of switches does not support mixed link speed for aggregated Ethernet interfaces.

You can configure Aggregated Ethernet interfaces on the M120 router to operate at one of the following speeds:

- `100m`—Links are 100 Mbps.
- `10g`—Links are 10 Gbps.
- `1g`—Links are 1 Gbps.

- `oc192`—Links are OC192 or STM64c.

You can configure aggregated Ethernet links on EX Series switches to operate at one of the following speeds:

- `10m`—Links are 10 Mbps.
- `100m`—Links are 100 Mbps.
- `1g`—Links are 1 Gbps.
- `10g`—Links are 10 Gbps.
- `50g`—Links are 50 Gbps.

You can configure aggregated Ethernet links on MX Series, and PTX Series routers and on QFX5100, QFX5120, QFX10002, QFX10008, and QFX10016 switches to operate at one of the following speeds:

- `100g`—Links are 100 Gbps.
- `100m`—Links are 100 Mbps.
- `10g`—Links are 10 Gbps.
- `1g`—Links are 1 Gbps.
- `40g`—Links are 40 Gbps.
- `50g`—Links are 50 Gbps.
- `80g`—Links are 80 Gbps.
- `8g`—Links are 8 Gbps.
- `mixed`—Links are of various speeds.
- `oc192`—Links are OC192.

Configuring Periodic Rebalancing of Subscribers in an Aggregated Ethernet Interface

If subscribers are frequently logging in and logging out of your network, you can configure the system to periodically rebalance the links based on a specific time and interval.

To configure periodic rebalancing:

1. Access the aggregated Ethernet interface for which you want to configure periodic rebalancing.

```
edit
user@host# edit interfaces ae number aggregated-ether-options
```

2. Configure the rebalancing parameters for the interface, including the time and the interval between rebalancing actions.

```
[edit interfaces aenumber aggregated-ether-options]
user@host# rebalance-periodic time hour:minute <interval hours>
```

SEE ALSO

Verifying the Distribution of Demux Subscribers in an Aggregated Ethernet Interface

Configuring the Distribution Type for Demux Subscribers on Aggregated Ethernet Interfaces

Distribution of Demux Subscribers in an Aggregated Ethernet Interface

Example: Configuring Aggregated Ethernet High-Speed Uplinks Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch

IN THIS SECTION

- Requirements | 296
- Overview and Topology | 296
- Configuration | 299
- Verification | 303
- Troubleshooting | 304

EX Series switches allow you to combine multiple Ethernet links into one logical interface for higher bandwidth and redundancy. The ports that are combined in this manner are referred to as a link

aggregation group (LAG) or bundle. The number of Ethernet links you can combine into a LAG depends on your EX Series switch model.

This example describes how to configure uplink LAGs to connect a Virtual Chassis access switch to a Virtual Chassis distribution switch:

Requirements

This example uses the following software and hardware components:

- Junos OS Release 9.0 or later for EX Series switches
- Two EX4200-48P switches
- Two EX4200-24F switches
- Four XFP uplink modules

Before you configure the LAGs, be sure you have:

- Configured the Virtual Chassis switches. See *Configuring an EX4200, EX4500, or EX4550 Virtual Chassis (CLI Procedure)*.
- Configured the uplink ports on the switches as trunk ports. See Configuring Gigabit Ethernet Interfaces (CLI Procedure).

Overview and Topology

For maximum speed and resiliency, you can combine uplinks between an access switch and a distribution switch into LAGs. Using LAGs can be particularly effective when connecting a multimember Virtual Chassis access switch to a multimember Virtual Chassis distribution switch.

The Virtual Chassis access switch in this example is composed of two member switches. Each member switch has an uplink module with two 10-Gigabit Ethernet ports. These ports are configured as trunk ports, connecting the access switch with the distribution switch.

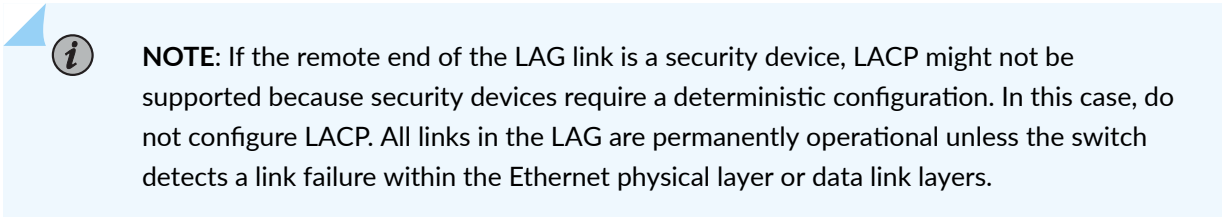
Configuring the uplinks as LAGs has the following advantages:

- Link Aggregation Control Protocol (LACP) can optionally be configured for link negotiation.
- It doubles the speed of each uplink from 10 Gbps to 20 Gbps.
- If one physical port is lost for any reason (a cable is unplugged or a switch port fails, or one member switch is unavailable), the logical port transparently continues to function over the remaining physical port.

The topology used in this example consists of one Virtual Chassis access switch and one Virtual Chassis distribution switch. The access switch is composed of two EX4200-48P switches (SWA-0 and SWA-1), interconnected to each other with their Virtual Chassis ports (VCPs) as member switches of Host-A. The

distribution switch is composed of two EX4200-24F switches (SWD-0 and SWD-1), interconnected with their VCPs as member switches of Host-D.

Each member of the access switch has an uplink module installed. Each uplink module has two ports. The uplinks are configured to act as trunk ports, connecting the access switch with the distribution switch. One uplink port from SWA-0 and one uplink port from SWA-1 are combined as LAG `ae0` to SWD-0. This link is used for one VLAN. The remaining uplink ports from SWA-0 and from SWA-1 are combined as a second LAG connection (`ae1`) to SWD-1. LAG `ae1` is used for another VLAN.



NOTE: If the remote end of the LAG link is a security device, LACP might not be supported because security devices require a deterministic configuration. In this case, do not configure LACP. All links in the LAG are permanently operational unless the switch detects a link failure within the Ethernet physical layer or data link layers.

Figure 15: Topology for LAGs Connecting an EX4200 Virtual Chassis Access Switch to an EX4200 Virtual Chassis Distribution Switch

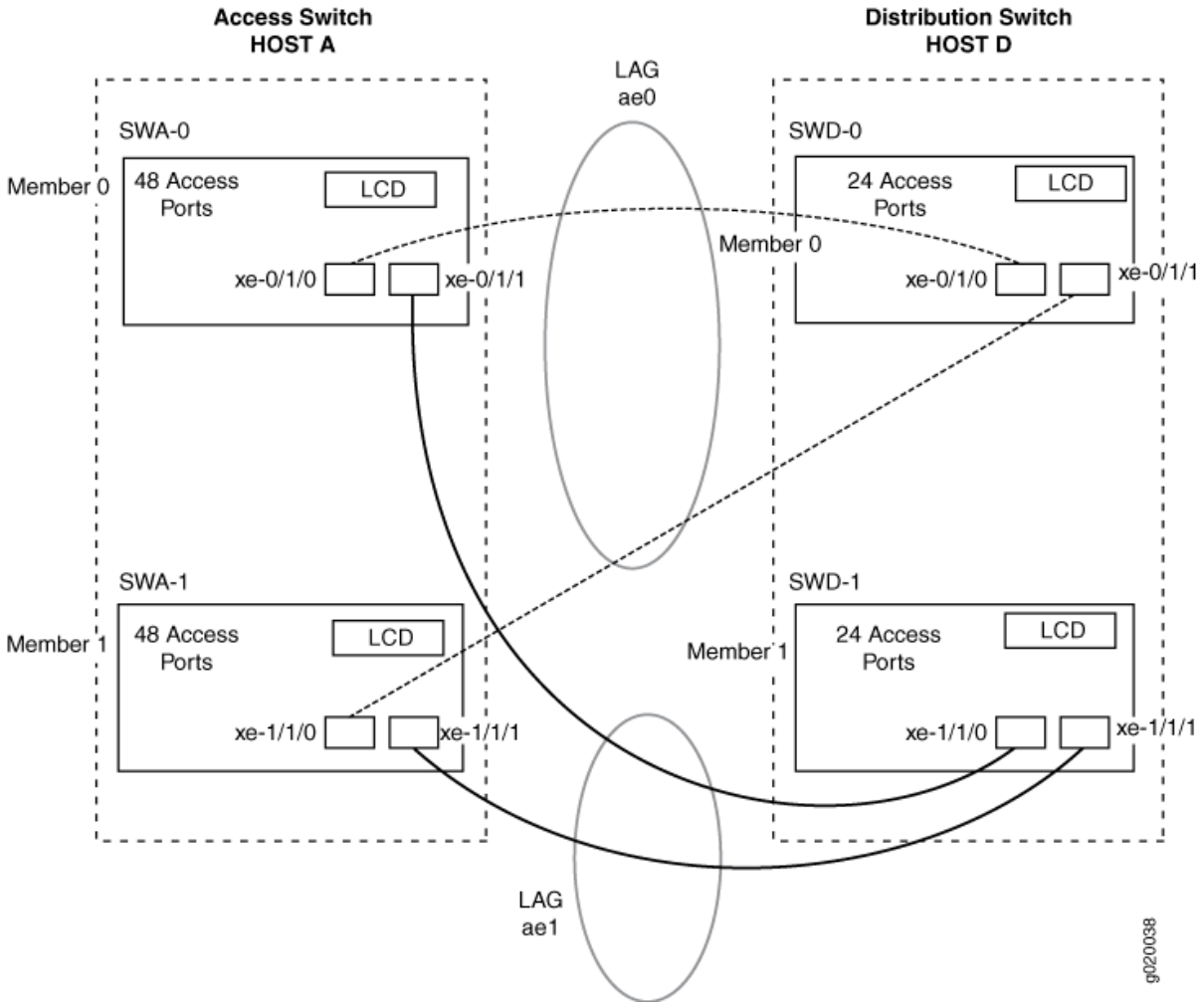


Table 73 on page 298 details the topology used in this configuration example.

Table 73: Components of the Topology for Connecting a Virtual Chassis Access Switch to a Virtual Chassis Distribution Switch

Switch	Hostname and VCID	Base Hardware	Uplink Module	Member ID	Trunk Port
SWA-0	Host-A Access switch VCID 1	EX4200-48P switch	One XFP uplink module	0	`xe-0/1/0` to SWD-0 `xe-0/1/1` to SWD-1

Table 73: Components of the Topology for Connecting a Virtual Chassis Access Switch to a Virtual Chassis Distribution Switch *(Continued)*

Switch	Hostname and VCID	Base Hardware	Uplink Module	Member ID	Trunk Port
SWA-1	Host-A Access switch VCID 1	EX4200-48P switch	One XFP uplink module	1	xe-1/1/0 to SWD-0 xe-1/1/1 to SWD-1
SWD-0	Host-D Distribution switch VCID 4	EX4200 L-24F switch	One XFP uplink module	0	xe-0/1/0 to SWA-0 xe-0/1/1 to SWA-1
SWD-1	Host-D Distribution switch VCID 4	EX4200 L-24F switch	One XFP uplink module	1	xe-1/1/0 to SWA-0 xe-1/1/1 to SWA-1

Configuration

IN THIS SECTION

- Procedure | 300
- Results | 302

To configure two uplink LAGs from the Virtual Chassis access switch to the Virtual Chassis distribution switch.

Procedure

CLI Quick Configuration

To quickly configure aggregated Ethernet high-speed uplinks between a Virtual Chassis access switch and a Virtual Chassis distribution switch, copy the following commands and paste them into the switch terminal window:

```
[edit]
                        set chassis aggregated-devices ethernet device-count 2
                        set interfaces ae0 aggregated-ether-options minimum-links 1
                        set interfaces ae0 aggregated-ether-options link-speed 10g
                        set interfaces ae1 aggregated-ether-options minimum-links 1
                        set interfaces ae1 aggregated-ether-options link-speed 10g
                        set interfaces ae0 unit 0 family inet address 192.0.2.0/25
                        set interfaces ae1 unit 0 family inet address 192.0.2.128/25
                        set interfaces xe-0/1/0 ether-options 802.3ad ae0
                        set interfaces xe-1/1/0 ether-options 802.3ad ae0
                        set interfaces xe-0/1/1 ether-options 802.3ad ae1
                        set interfaces xe-1/1/1 ether-options 802.3ad ae1
```

Step-by-Step Procedure

To configure aggregated Ethernet high-speed uplinks between a Virtual Chassis access switch and a Virtual Chassis distribution switch:

1. Specify the number of LAGs to be created on the chassis:

```
[edit chassis]
user@Host-A# set aggregated-devices ethernet device-count 2
```

2. Specify the number of links that need to be present for the `ae0` LAG interface to be `up`:

```
[edit interfaces]
user@Host-A# set ae0 aggregated-ether-options minimum-links 1
```

3. Specify the number of links that need to be present for the ae1 LAG interface to be up:

```
[edit interfaces]
user@Host-A# set ae1 aggregated-ether-options minimum-links 1
```

4. Specify the media speed of the ae0 link:

```
[edit interfaces]
user@Host-A# set ae0  aggregated-ether-options link-speed 10g
```

5. Specify the media speed of the ae1 link:

```
[edit interfaces]
user@Host-A# set ae1  aggregated-ether-options link-speed 10g
```

6. Specify the interface ID of the uplinks to be included in LAG ae0:

```
[edit interfaces]
user@Host-A# set xe-0/1/0 ether-options 802.3ad ae0
user@Host-A# set xe-1/1/0 ether-options 802.3ad ae0
```

7. Specify the interface ID of the uplinks to be included in LAG ae1:

```
[edit interfaces]
user@Host-A# set xe-0/1/1 ether-options 802.3ad ae1
user@Host-A# set xe-1/1/1 ether-options 802.3ad ae1
```

8. Specify that LAG ae0 belongs to the subnet for the employee broadcast domain:

```
[edit interfaces]
user@Host-A# set ae0 unit 0 family inet address 192.0.2.0/25
```

9. Specify that LAG ae1 belongs to the subnet for the guest broadcast domain:

```
[edit interfaces]
user@Host-A# set ae1 unit 0 family inet address 192.0.2.128/25
```

Results

Display the results of the configuration:

```
[edit]
chassis {
    aggregated-devices {
        ethernet {
            device-count 2;
        }
    }
}
interfaces {
    ae0 {
        aggregated-ether-options {
            link-speed 10g;
            minimum-links 1;
        }
        unit 0 {
            family inet {
                address 192.0.2.0/25;
            }
        }
    }
    ae1 {
        aggregated-ether-options {
            link-speed 10g;
            minimum-links 1;
        }
        unit 0 {
            family inet {
                address 192.0.2.128/25;
            }
        }
    xe-0/1/0 {
```

```
        ether-options {
            802.3ad ae0;
        }
    }
    xe-1/1/0 {
        ether-options {
            802.3ad ae0;
        }
    }
    xe-0/1/1 {
        ether-options {
            802.3ad ae1;
        }
    }
    xe-1/1/1 {
        ether-options {
            802.3ad ae1;
        }
    }
}
```

Verification

IN THIS SECTION

- Verifying That LAG ae0 Has Been Created | **303**
- Verifying That LAG ae1 Has Been Created | **304**

To verify that switching is operational and two LAGs have been created, perform these tasks:

Verifying That LAG ae0 Has Been Created

Purpose

Verify that LAG `ae0` has been created on the switch.

Action

```
show interfaces ae0 terse
```

```
Interface               Admin      Link Proto      Local            Remote
ae0                     up         up
ae0.0                   up         up     inet     192.0.2.0/25
```

Meaning

The output confirms that the `ae0` link is up and shows the `family` and IP address assigned to this link.

Verifying That LAG ae1 Has Been Created

Purpose

Verify that LAG `ae1` has been created on the switch

Action

```
show interfaces ae1 terse
```

```
Interface               Admin Link Proto     Local               Remote
ae1                     up    down
ae1.0                   up    down inet          192.0.2.128/25
```

Meaning

The output shows that the `ae1` link is down.

Troubleshooting

IN THIS SECTION

- Troubleshooting a LAG That Is Down | 305

Troubleshooting a LAG That Is Down

Problem

The `show interfaces terse` command shows that the LAG is `down`

Solution

Check the following:

- Verify that there is no configuration mismatch.
- Verify that all member ports are up.
- Verify that a LAG is part of family ethernet switching (Layer 2 LAG) or family inet (Layer 3 LAG).
- Verify that the LAG member is connected to the correct LAG at the other end.
- Verify that the LAG members belong to the same switch (or the same Virtual Chassis).

SEE ALSO

Example: Configuring an EX4200 Virtual Chassis with a Primary and Backup in a Single Wiring Closet

Example: Connecting an EX Series Access Switch to a Distribution Switch

Virtual Chassis Cabling Configuration Examples for EX4200 Switches

Installing an Uplink Module in an EX4200 Switch

Example: Configuring Link Aggregation Between a QFX Series Product and an Aggregation Switch

IN THIS SECTION

- Requirements | 306
- Overview and Topology | 306
- Configuration | 307
- Verification | 311

Troubleshooting | 312

A QFX Series product allows you to combine multiple Ethernet links into one logical interface for higher bandwidth and redundancy. The ports that are combined in this manner are referred to as a link aggregation group (LAG) or bundle. The number of Ethernet links you can combine into a LAG depends on your QFX Series product model. You can configure LAGs to connect a QFX Series product or an EX4600 switch to other switches, like aggregation switches, servers, or routers. This example describes how to configure LAGs to connect a QFX3500, QFX3600, EX4600, QFX5100, and QFX10002 switch to an aggregation switch.

Requirements

This example uses the following software and hardware components:

- Junos OS Release 11.1 or later for the QFX3500 and QFX3600 switches, Junos OS 13.2 or later for the QFX5100 and EX4600 switch, and Junos OS Release 15.1X53-D10 or later for QFX10002 switches.
- One QFX3500, QFX3600, EX4600, QFX5100, or QFX10002 switch.

Overview and Topology

In this example, the switch has one LAG comprising two 10-Gigabit Ethernet interfaces. This LAG is configured in port-mode trunk (or interface-mode trunk) so that the switch and the VLAN to which it has been assigned can send and receive traffic.

Configuring the Ethernet interfaces as LAGs has the following advantages:

- If one physical port is lost for any reason (a cable is unplugged or a switch port fails), the logical port transparently continues to function over the remaining physical port.
- Link Aggregation Control Protocol (LACP) can optionally be configured for link monitoring and automatic addition and deletion of individual links without user intervention.



NOTE: If the remote end of the LAG link is a security device, LACP might not be supported because security devices require a deterministic configuration. In this case, do not configure LACP. All links in the LAG are permanently operational unless the switch detects a link failure within the Ethernet physical layer or data link layers.

The topology used in this example consists of one switch with a LAG configured between two of its 10-Gigabit Ethernet interfaces. The switch is connected to an aggregation switch.

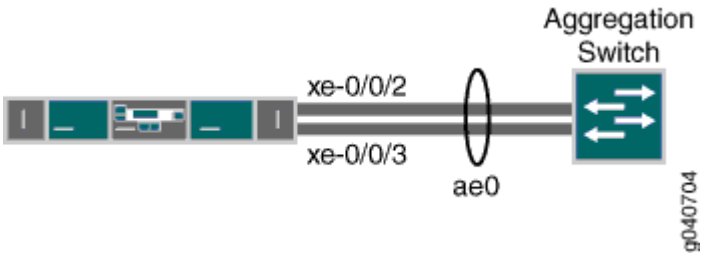


Table 74 on page 307 details the topology used in this configuration example.

Table 74: Components of the Topology for Configuring a LAG Between a Switch and an Aggregation Switch

Hostname	Base Hardware	Trunk Port
switch	QFX3500, QFX3600, EX4600, QFX5100, or QFX10002 switch	`ae0` is configured as a trunk port and combines the following two interfaces: `xe-0/0/2` and `xe-0/0/3` .

Configuration

IN THIS SECTION

- Procedure | 307
- Results | 310

To configure a LAG between two 10-Gigabit Ethernet interfaces.

Procedure

CLI Quick Configuration

To quickly configure a LAG between two 10-Gigabit Ethernet interfaces on a switch, copy the following commands and paste them into the switch terminal window:



NOTE: If you are configuring a LAG using Enhanced Layer 2 Software—for example, on the EX4600, QFX5100, or QFX10002 switch—use the `interface-mode` statement instead of the `port-mode` statement. For ELS details, see Using the Enhanced Layer 2 Software CLI.

```
[edit]
set chassis aggregated-devices ethernet device-count 1
set interfaces ae0 aggregated-ether-options minimum-links 1
set interfaces ae0 aggregated-ether-options link-speed 10g
set interfaces ae0 unit 0 family ethernet-switching vlan members green
set interfaces xe-0/0/2 ether-options 802.3ad ae0
set interfaces xe-0/0/3 ether-options 802.3ad ae0
set interfaces ae0 unit 0 family ethernet-switching port-mode trunk
set interfaces ae0 aggregated-ether-options lacp active
set interfaces ae0 aggregated-ether-options lacp periodic fast
```

Step-by-Step Procedure

To configure a LAG between a QFX Series switch and an aggregation switch:

1. Specify the number of LAGs to be created on the switch:

```
[edit chassis]
user@switch# set aggregated-devices ethernet device-count 1
```

2. Specify the number of links that need to be present for the `ae0` LAG interface to be `up`:

```
[edit interfaces]
user@switch# set ae0 aggregated-ether-options minimum-links 1
```

3. Specify the media speed of the `ae0` link:

```
[edit interfaces]
user@switch# set ae0 aggregated-ether-options link-speed 10g
```

4. Specify the members to be included within the aggregated Ethernet bundle:

```
[edit interfaces]
user@switch# set interfaces xe-0/0/2 ether-options 802.3ad ae0

[edit interfaces]
user@switch# set interfaces xe-0/0/3 ether-options 802.3ad ae0
```

5. Assign a port mode of trunk to the `ae0` link:

> **NOTE**: If you are configuring a LAG using Enhanced Layer 2 Software—for example, on the EX4600, QFX5100, or QFX10002 switch—use the `interface-mode` statement instead of the `port-mode` statement. For ELS details, see Using the Enhanced Layer 2 Software CLI.

```
[edit interfaces]
user@switch# set ae0 unit 0 family ethernet-switching port-mode trunk
```

or

```
[edit interfaces]
user@switch# set ae0 unit 0 family ethernet-switching interface-mode trunk
```

6. Assign the LAG to a VLAN:

```
[edit interfaces]
user@switch# set ae0 unit 0 family ethernet-switching vlan members green vlan-id 200
```

7. (Optional): Designate one side of the LAG as active for LACP:

```
[edit interfaces]
user@switch# set ae0 aggregated-ether-options lacp active
```

8. (Optional): Designate the interval and speed at which the interfaces send LACP packets:

```
[edit interfaces]
user@switch# set ae0 aggregated-ether-options lacp periodic fast
```

Results

Display the results of the configuration on a QFX3500 or QFX3600 switch:

```
[edit]
chassis {
    aggregated-devices {
        ethernet {
            device-count 1;
        }
    }
}
green {
    vlan-id 200;
    }
}
interfaces {
    ae0 {
        aggregated-ether-options {
            link-speed 10g;
            minimum-links 1;
        }
        unit 0 {
            family ethernet-switching {
                port-mode trunk;
                    vlan {
                    members green;
                }
        }
    }
    xe-0/0/2 {
        ether-options {
            802.3ad ae0;
        }
    }
    xe-0/0/3 {
```

```
        ether-options {
            802.3ad ae0;
        }
    }
}
```

Verification

IN THIS SECTION

- Verifying That LAG ae0.0 Has Been Created | **311**
- Verifying That LAG ae0 Has Been Created | **312**

To verify that switching is operational and one LAG has been created, perform these tasks:

Verifying That LAG ae0.0 Has Been Created

Purpose

Verify that LAG `ae0.0` has been created on the switch.

Action

`show interfaces ae0 terse`

```
Interface               Admin Link Proto    Local                 Remote
ae0                     up    up
ae0.0                   up    up eth-switch
```

Meaning

The output confirms that the `ae0.0` link is up and shows the `family` and IP address assigned to this link.

Verifying That LAG ae0 Has Been Created

Purpose

Verify that LAG `ae0` has been created on the switch

Action

```
show interfaces ae0 terse
```

```
Interface               Admin Link Proto    Local                 Remote
ae0                     up    down
ae0.0                   up    down eth-switch
```

Meaning

The output shows that the `ae0.0` link is down.

Troubleshooting

IN THIS SECTION

- Troubleshooting a LAG That Is Down | **312**

Troubleshooting a LAG That Is Down

Problem

The `show interfaces terse` command shows that the LAG is `down`.

Solution

Check the following:

- Verify that there is no configuration mismatch.
- Verify that all member ports are up.
- Verify that a LAG is part of family ethernet switching (Layer 2 LAG) or family inet (Layer 3 LAG).

- Verify that the LAG member is connected to the correct LAG at the other end.

SEE ALSO

> Verifying the Status of a LAG Interface
>
> *show lacp statistics interfaces (View)*

Configuring Aggregated Ethernet LACP

IN THIS SECTION

- Configuring the LACP Interval | **315**
- Configuring LACP Link Protection | **315**
- Configuring LACP System Priority | **317**
- Configuring LACP System Identifier | **317**
- Configuring LACP administrative Key | **317**
- Configuring LACP Port Priority | **318**
- Tracing LACP Operations | **318**
- LACP Limitations | **319**
- Example: Configuring Aggregated Ethernet LACP | **319**

For aggregated Ethernet interfaces, you can configure the Link Aggregation Control Protocol (LACP). LACP is one method of bundling several physical interfaces to form one logical interface. You can configure both VLAN-tagged and untagged aggregated Ethernet with or without LACP enabled.

For Multichassis Link Aggregation (MC-LAG), you must specify the `system-id` and `admin key`. MC-LAG peers use the same `system-id` while sending the LACP messages. The `system-id` can be configured on the MC-LAG network device and synchronized between peers for validation.

LACP exchanges are made between actors and partners. An actor is the local interface in an LACP exchange. A partner is the remote interface in an LACP exchange.

LACP is defined in IEEE 802.3ad, *Aggregation of Multiple Link Segments*.

LACP was designed to achieve the following:

- Automatic addition and deletion of individual links to the aggregate bundle without user intervention
- Link monitoring to check whether both ends of the bundle are connected to the correct group

The Junos OS implementation of LACP provides link monitoring but not automatic addition and deletion of links.

The LACP mode can be active or passive. If the actor and partner are both in passive mode, they do not exchange LACP packets, which results in the aggregated Ethernet links not coming up. If either the actor or partner is active, they do exchange LACP packets. By default, LACP is turned off on aggregated Ethernet interfaces. If LACP is configured, it is in passive mode by default. To initiate transmission of LACP packets and response to LACP packets, you must configure LACP in active mode.

To enable LACP active mode, include the `lacp` statement at the `[edit interfaces` *`interface-name`* `aggregated-ether-options]` hierarchy level, and specify the `active` option:

```
[edit interfaces interface-name aggregated-ether-options]
lacp {
    active;
}
```



NOTE: The LACP process exists in the system only if you configure the system in either active or passive LACP mode.

To restore the default behavior, include the `lacp` statement at the `[edit interfaces` *`interface-name`* `aggregated-ether-options]` hierarchy level, and specify the `passive` option:

```
[edit interfaces interface-name aggregated-ether-options]
lacp {
    passive;
}
```

Starting with Junos OS release 12.2, you can also configure LACP to override the IEEE 802.3ad standard and to allow the standby link always to receive traffic. Overriding the default behavior facilitates subsecond failover.

To override the IEEE 802.3ad standard and facilitate subsecond failover, include the `fast-failover` statement at the `[edit interfaces` *`interface-name`* `aggregated-ether-options lacp]` hierarchy level.

For more information, see the following sections:

Configuring the LACP Interval

By default, the actor and partner send LACP packets every second. You can configure the interval at which the interfaces send LACP packets by including the `periodic` statement at the `[edit interfaces` *`interface-name`* `aggregated-ether-options lacp]` hierarchy level:

```
[edit interfaces interface-name aggregated-ether-options lacp]
periodic interval;
```

The interval can be fast (every second) or slow (every 30 seconds). You can configure different periodic rates on active and passive interfaces. When you configure the active and passive interfaces at different rates, the transmitter honors the receiver's rate.



NOTE: Source address filtering does not work when LACP is enabled.

Percentage policers are not supported on aggregated Ethernet interfaces with the CCC protocol family configured. For more information about percentage policers, see the Routing Policies, Firewall Filters, and Traffic Policers User Guide.

Generally, LACP is supported on all untagged aggregated Ethernet interfaces. For more information, see Configuring Untagged Aggregated Ethernet Interfaces.

Configuring LACP Link Protection



NOTE: When using LACP link protection, you can configure only two member links to an aggregated Ethernet interface: one active and one standby.

To force active and standby links within an aggregated Ethernet, you can configure LACP link protection and system priority at the aggregated Ethernet interface level using the `link-protection` and `system-priority` statements. Configuring values at this level results in only the configured interfaces using the defined configuration. LACP interface configuration also enables you to override global (chassis) LACP settings.

LACP link protection also uses port priority. You can configure port priority at the Ethernet interface `[ether-options]` hierarchy level using the `port-priority` statement. If you choose not to configure port priority, LACP link protection uses the default value for port priority (127).



NOTE: LACP link protection supports per-unit scheduling configuration on aggregated Ethernet interfaces.

To enable LACP link protection for an aggregated Ethernet interfaces, use the `link-protection` statement at the `[edit interfaces aeX aggregated-ether-options lacp]` hierarchy level:

```
[edit interfaces aeX aggregated-ether-options lacp]
link-protection;
    disable;
    revertive;
    non-revertive;
}
```

By default, LACP link protection reverts to a higher-priority (lower-numbered) link when that higher-priority link becomes operational or a link is added to the aggregator that is determined to be higher in priority. However, you can suppress link calculation by adding the `non-revertive` statement to the LACP link protection configuration. In nonrevertive mode, once a link is active and collecting and distributing packets, the subsequent addition of a higher-priority (better) link does not result in a switch and the current link remains active.

If LACP link protection is configured to be nonrevertive at the global (`[edit chassis]` hierarchy) level, you can add the `revertive` statement to the LACP link protection configuration to override the nonrevertive setting for the interface. In revertive mode, the addition of a higher-priority link to the aggregator results in LACP performing a priority recalculation and switching from the current active link to the new active link.



CAUTION: If both ends of an aggregator have LACP link protection enabled, make sure to configure both ends of the aggregator to use the same mode. Mismatching LACP link protection modes can result in lost traffic.

We strongly recommend you to use LACP on both ends of the aggregator, when you connect an aggregated Ethernet interface with two member interfaces to any other vendor device. Otherwise, the vendor device (say a Layer 2 switch, or a router), will not be able to manage the traffic coming from the two link aggregated Ethernet bundle. As a result, you might observe the vendor device sending back the traffic to the backup member link of the aggregated Ethernet interface.

Currently, MX-MPC2-3D, MX-MPC2-3D-Q, MX-MPC2-3D-EQ, MX-MPC1-3D, MX-MPC1-3D-Q, and MPC-3D-16XGE-SFPP do not drop traffic coming back to the backup link, whereas DPCE-R-Q-20GE-2XGE, DPCE-R-Q-20GE-SFP, DPCE-R-Q-40GE-SFP, DPCE-R-Q-4XGE-XFP, DPCE-X-Q-40GE-SFP, and DPCE-X-Q-4XGE-XFP drop traffic coming to the backup link.

Configuring LACP System Priority

To configure LACP system priority for aggregated Ethernet interfaces on the interface, use the `system-priority` statement at the `[edit interfaces aeX aggregated-ether-options lacp]` hierarchy level:

```
[edit interfaces aeX aggregated-ether-options lacp]
system-priority;
```

The system priority is a 2-octet binary value that is part of the LACP system ID. The LACP system ID consists of the system priority as the two most-significant octets and the interface MAC address as the six least-significant octets. The system with the numerically lower value for system priority has the higher priority. By default, system priority is 127, with a range of 0 to 65,535.

Configuring LACP System Identifier

To configure the LACP system identifier for aggregated Ethernet interfaces, use the `system-id` statement at the `[edit interfaces aeX aggregated-ether-options lacp]` hierarchy level:

```
[edit interfaces aeX aggregated-ether-options lacp]
system-id system-id;
```

The user-defined system identifier in LACP enables two ports from two separate devices to act as though they were part of the same aggregate group.

The system identifier is a 48-bit (6-byte) globally unique field. It is used in combination with a 16-bit system-priority value, which results in a unique LACP system identifier.

Configuring LACP administrative Key

To configure an administrative key for LACP, include the `admin-key number` statement at the `edit interfaces aex aggregated-ether-options lacp]` hierarchy level:

```
[edit interfaces ae x aggregated-ether-options-lacp]
admin-key number;
```



NOTE: You must configure MC-LAG to configure the `admin-key` statement. For more information about MC-LAG, see Configuring Multichassis Link Aggregation on MX Series Routers .

Configuring LACP Port Priority

To configure LACP port priority for aggregated Ethernet interfaces, use the `port-priority` statement at the `[edit interfaces` *`interface-name`* `ether-options 802.3ad ae`*`X`* `lacp]` or `[edit interfaces` *`interface-name`* `ether-options 802.3ad ae`*`X`* `lacp]` hierarchy levels:

```
[edit interfaces interface-name ether-options 802.3ad aeX lacp]
port-priority priority;
```

The port priority is a 2-octet field that is part of the LACP port ID. The LACP port ID consists of the port priority as the two most-significant octets and the port number as the two least-significant octets. The system with the numerically lower value for port priority has the higher priority. By default, port priority is 127, with a range of 0 to 65,535.

Port aggregation selection is made by each system based on the highest port priority and are assigned by the system with the highest priority. Ports are selected and assigned starting with the highest priority port of the highest priority system and working down in priority from there.



NOTE: Port aggregation selection (discussed above) is performed for the active link when LACP link protection is enabled. Without LACP link protection, port priority is not used in port aggregation selection.

Tracing LACP Operations

To trace the operations of the LACP process, include the `traceoptions` statement at the `[edit protocols lacp]` hierarchy level:

```
[edit protocols lacp]
traceoptions {
    file <filename> <files number> <size size> <world-readable | no-world-readable>;
    flag flag;
    no-remote-trace;
}
```

You can specify the following flags in the `protocols lacp traceoptions` statement:

- `all`—All LACP tracing operations
- `configuration`—Configuration code
- `packet`—Packets sent and received
- `process`—LACP process events

- `protocol`—LACP protocol state machine
- `routing-socket`—Routing socket events
- `startup`—Process startup events

LACP Limitations

LACP can link together multiple different physical interfaces, but only features that are supported across all of the linked devices will be supported in the resulting link aggregation group (LAG) bundle. For example, different PICs can support a different number of forwarding classes. If you use link aggregation to link together the ports of a PIC that supports up to 16 forwarding classes with a PIC that supports up to 8 forwarding classes, the resulting LAG bundle will only support up to 8 forwarding classes. Similarly, linking together a PIC that supports WRED with a PIC that does not support it will result in a LAG bundle that does not support WRED.

Example: Configuring Aggregated Ethernet LACP

IN THIS SECTION

- Topology | 319

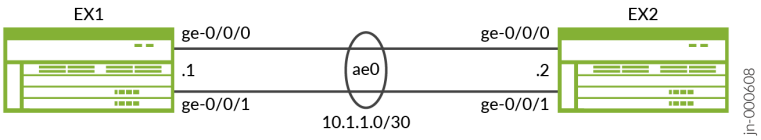
This example shows how to configure an aggregated ethernet interface with active LACP between two EX switches.

Topology

IN THIS SECTION

- Verification | 321

Two EX switches are connected together using two interfaces in an aggregated ethernet configuration.



Configure aggregated Ethernet LACP over an untagged interface:

NOTE: We are only showing the configuration for EX1 in this example. EX2 has the same configuration except for the IP address.

LACP with Untagged Aggregated Ethernet

The chassis configuration allows for 1 aggregated ethernet interface. The `802.3ad` configuration associates both interfaces `ge-0/0/0` and `ge-0/0/1` with interface `ae0`. The `ae0 aggregated-ether-options` configuration enables active mode LACP.

```
user@EX1# show
...
chassis {
    aggregated-devices {
        ethernet {
            device-count 1;
        }
    }
}
interfaces {
    ge-0/0/0 {
        ether-options {
            802.3ad ae0;
        }
    }
    ge-0/0/1 {
        ether-options {
            802.3ad ae0;
        }
    }
    ae0 {
        aggregated-ether-options {
            lacp {
                active;
```

```
            }
        }
        unit 0 {
            family inet {
                address 10.1.1.1/30;
            }
        }
    }
}
```

Verification

IN THIS SECTION

- Verifying the Aggregated Ethernet Interface | 321
- Verifying LACP is Active | 322
- Verify Reachability | 322

Verifying the Aggregated Ethernet Interface

Purpose

Verify the aggregated ethernet interface has been created and is up.

Action

Use the command `show interfaces terse | match ae` from operational mode.

```
user@EX1> show interfaces terse | match ae
ge-0/0/0.0              up    up   aenet    --> ae0.0
ge-0/0/1.0              up    up   aenet    --> ae0.0
ae0                     up    up
ae0.0                   up    up   inet     10.1.1.1/30
```

Meaning

The output shows that ge-0/0/0 and ge-0/0/1 are bundled together to create the aggregated ethernet interface `ae0` and the interface is up.

Verifying LACP is Active

Purpose

Verify which interfaces are participating in LACP and the current state.

Action

Use the command `show lacp interfaces` from operational mode.

```
user@EX1> show lacp interfaces
Aggregated interface: ae0
    LACP state:       Role   Exp   Def  Dist  Col  Syn  Aggr  Timeout  Activity
      ge-0/0/0       Actor    No    No   Yes  Yes  Yes   Yes     Fast    Active
      ge-0/0/0     Partner    No    No   Yes  Yes  Yes   Yes     Fast    Active
      ge-0/0/1       Actor    No    No   Yes  Yes  Yes   Yes     Fast    Active
      ge-0/0/1     Partner    No    No   Yes  Yes  Yes   Yes     Fast    Active
    LACP protocol:        Receive State  Transmit State          Mux State
      ge-0/0/0                  Current   Fast periodic Collecting distributing
      ge-0/0/1                  Current   Fast periodic Collecting distributing
```

Meaning

The output shows that the active mode LACP is enabled.

Verify Reachability

Purpose

Verify that ping works between the two EX switches.

Action

Use the `ping 10.1.1.2 count 2` operational mode command on EX1.

```
user@EX1> ping 10.1.1.2 count 2
PING 10.1.1.2 (10.1.1.2): 56 data bytes
64 bytes from 10.1.1.2: icmp_seq=0 ttl=64 time=2.249 ms
64 bytes from 10.1.1.2: icmp_seq=1 ttl=64 time=2.315 ms
```

```
--- 10.1.1.2 ping statistics ---
2 packets transmitted, 2 packets received, 0% packet loss
round-trip min/avg/max/stddev = 2.249/2.282/2.315/0.033 ms
```

Meaning

EX1 is able to ping EX2 across the aggregated ethernet interface.

RELATED DOCUMENTATION

- lacp
- link-protection
- traceoptions

Configuring LACP Link Protection of Aggregated Ethernet Interfaces for Switches

IN THIS SECTION

- Configuring LACP Link Protection for a Single Link at the Global Level | **325**
- Configuring LACP Link Protection for a Single Link at the Aggregated Interface Level | **325**
- Configuring Subgroup Bundles to Provide LACP Link Protection to Multiple Links in an Aggregated Ethernet Interface | **326**

You can configure LACP link protection and system priority at the global level on the switch or for a specific aggregated Ethernet interface. When using LACP link protection to protect a single link in the aggregated ethernet bundle, you configure only two member links for an aggregated Ethernet interface: one active and one standby. LACP link protection ensures that only one link—the link with the higher priority—is used for traffic. The other link is forced to stay in a *waiting* state.

Use the following command to verify the active and standby links.

```
user@host# run show interfaces redundancy
Interface  State           Last change  Primary      Secondary    Current status
```

```
ae0        On secondary     14:56:50     xe-0/0/1     xe-0/0/2     both up
```

When using LACP link protection to protect multiple links in an aggregated ethernet bundle, you configure links into primary and backup subgroups. A link protection subgroup is a collection of ethernet links within the aggregated ethernet bundle. When you use link protection subgroups, you configure a primary subgroup and a backup subgroup. The configuration process includes assigning member links to each subgroup. When the configuration process is complete, the primary subgroup is used to forward traffic until a switchover event, such as a link failure, occurs and causes the backup subgroup to assume control of traffic that was travelling on the links in the primary subgroup within the bundle.

By default LACP link protection reverts to a higher-priority (lower-numbered) link when the higher-priority link becomes operational or when a higher-priority link is added to the aggregated Ethernet bundle. For priority purposes, LACP link protection treats subgroups like links. You can suppress link calculation by adding the `non-revertive` statement to the link protection configuration. In nonrevertive mode, when a link is active in sending and receiving LACP packets, adding a higher-priority link to the bundle does not change the status of the currently active link. It remains active.

If LACP link configuration is specified to be nonrevertive at the global `[edit chassis]` hierarchy level, you can specify the `revertive` statement in the LACP link protection configuration at the aggregated Ethernet interface level to override the nonrevertive setting for the interface. In revertive mode, adding a higher-priority link to the aggregated Ethernet bundle results in LACP recalculating the priority and switching the status from the currently active link to the newly added, higher-priority link.



NOTE: When LACP link protection is enabled on both local and remote sides of the link, both sides must use the same mode (either revertive or nonrevertive).

Configuring LACP link configuration at the aggregated Ethernet level results in only the configured interfaces using the defined configuration. LACP interface configuration also enables you to override global (chassis) LACP settings.

Before you configure LACP link protection, be sure you have:

- Configured the aggregated Ethernet bundles—also known as link aggregation groups (LAGs). For EX Series, see Configuring Aggregated Ethernet Links (CLI Procedure).
- Configured LACP for the interface. For Ex Series, see "Configuring Aggregated Ethernet LACP (CLI Procedure)" on page 283.

You can configure LACP link protection for all aggregated Ethernet interfaces on the switch by enabling it at the global level on the switch or configure it for a specific aggregated Ethernet interface by enabling it on that interface.

Configuring LACP Link Protection for a Single Link at the Global Level

To configure LACP link protection for aggregated Ethernet interfaces at the global level:

1. Enable LACP link protection on the switch:

```
[edit chassis aggregated-devices ethernet lacp]
user@switch# set link-protection
```

2. (Optional) Configure the LACP link protection for the aggregated Ethernet interfaces to be in nonrevertive mode:



NOTE: LACP link protection is in revertive mode by default.

```
[edit chassis aggregated-devices ethernet lacp link-protection]
user@switch# set non-revertive
```

3. (Optional)To configure LACP system priority for the aggregated Ethernet interfaces:

```
[edit chassis aggregated-devices ethernet lacp]
user@switch# set system-priority
```

Configuring LACP Link Protection for a Single Link at the Aggregated Interface Level

To enable LACP link protection for a specific aggregated Ethernet interface:

1. Enable LACP link protection for the interface:

```
[edit interfaces aeX aggregated-ether-options lacp]
user@switch# set link-protection
```

2. (Optional) Configure the LACP link protection for the aggregated Ethernet interface to be in revertive or nonrevertive mode:

   - To specify revertive mode:

```
[edit interfaces aeX aggregated-ether-options lacp link-protection]
user@switch# set revertive
```

- To specify nonrevertive mode:

```
[edit interfaces aeX aggregated-ether-options lacp link-protection]
user@switch# set non-revertive
```

3. (Optional)To configure LACP system priority for an aggregated Ethernet interface:

```
[edit interfaces aeX aggregated-ether-options lacp link-protection]
user@switch# set system-priority
```

4. (Optional) To configure LACP port priority for an aggregated Ethernet interface:

```
[edit interfaces ge-fpc/pic/port ether-options 802.3ad lacp]
user@switch# set port-priority
```

Configuring Subgroup Bundles to Provide LACP Link Protection to Multiple Links in an Aggregated Ethernet Interface

You can configure link protection subgroup bundles to provide link protection for multiple links in an aggregated ethernet bundle.

Link protection subgroups allow you to provide link protection to a collection of Ethernet links within a LAG bundle, instead of providing protection to a single link in the aggregated ethernet bundle only. You can, for instance, configure a primary subgroup with three member links and a backup subgroup with three different member links and use the backup subgroup to provide link protection for the primary subgroup.

To configure link protection using subgroups:

1. Configure the primary link protection subgroup in the aggregated ethernet interface:

```
[edit interfaces aeX aggregated-ether-options]
user@switch# set link-protection-sub-group group-name primary
```

For instance, to create a primary link protection subgroup named **subgroup-primary** for interface **ae0**:

```
[edit interfaces ae0 aggregated-ether-options]
user@switch# set link-protection-sub-group subgroup-primary primary
```

2. Configure the backup link protection subgroup in the aggregated ethernet interface:

```
[edit interfaces aeX aggregated-ether-options]
user@switch# set link-protection-sub-group group-name backup
```

For instance, to create a backup link protection subgroup named **subgroup-backup** for interface **ae0**:

```
[edit interfaces ae0 aggregated-ether-options]
user@switch# set link-protection-sub-group subgroup-backup backup
```

NOTE: You can create one primary and one backup link protection subgroup per aggregated ethernet interface.

3. Attach interfaces to the link protection subgroups:

```
[edit interfaces interface-name ether-options 802.3ad]
user@switch# set link-protection-sub-group group-name
```

NOTE: The primary and backup link protection subgroups must contain the same number of interfaces. For instance, if the primary link protection subgroup contains three interfaces, the backup link protection subgroup must also contain three interfaces.

For instance, to configure interfaces **ge-0/0/0** and **ge-0/0/1** into link protection subgroup **subgroup-primary** and interfaces **ge-0/0/2** and **ge-0/0/3** into link protection subgroup **subgroup-backup**:

```
[edit interfaces ge-0/0/0 ether-options 802.3ad]
user@switch# set link-protection-sub-group subgroup-primary
[edit interfaces ge-0/0/1 ether-options 802.3ad]
user@switch# set link-protection-sub-group subgroup-primary
[edit interfaces ge-0/0/2 ether-options 802.3ad]
user@switch# set link-protection-sub-group subgroup-backup
[edit interfaces ge-0/0/3 ether-options 802.3ad]
user@switch# set link-protection-sub-group subgroup-backup
```

4. (Optional) Configure the port priority for link protection:

```
[edit interfaces interface-name ether-options 802.3ad]
user@switch# set port-priority priority
```

The port priority is used to select the active link.

5. Enable link protection

To enable link protection at the LAG level:

```
[edit interfaces aeX aggregated-ether-options]
user@switch# set link-protection
```

NOTE: ACX Series routers do not support static link protection.

To enable link protection at the LACP level:

```
[edit interfaces aeX aggregated-ether-options lacp]
user@switch# set link-protection
```

For instance, to enable link protection on **ae0** at the LAG level:

```
[edit interfaces ae0 aggregated-ether-options]
user@switch# set link-protection
```

For instance, to enable link protection on **ae0** at the LACP level:

```
[edit interfaces ae0 aggregated-ether-options lacp]
user@switch# set link-protection
```

NOTE: The LACP decides active and back up state of links. When configuring LACP, the state of the backup link should not be configured manually as down. The following command is not supported if LACP is configured:`set interfaces` **ae0** `aggregated-ether-options link-protection backup-state down`

RELATED DOCUMENTATION

lacp (Aggregated Ethernet)

Configuring LACP Hold-UP Timer to Prevent Link Flapping on LAG Interfaces

On link aggregation group (LAG) interfaces, when a member (child) link goes down, its state changes from current to expired. This link might flap from the current state to the expired state and back to current state when it receives intermittent LACP protocol data units (PDUs) and keepalive timeouts. Such flapping can adversely affect the traffic on the link.

To prevent excessive flapping of a LAG child link, you can configure a hold-up timer on the LAG interface that is applicable to all member links on that particular interface. To hold up, in networking terms, means to prevent the transitioning of an interface from down to up for a specified time interval.

When configured, the hold-up timer is triggered when an LACP state machine tries to move to the current state from the expired or default state when it receives an LACP PDU. The hold-up timer is triggered only if the LACP state machine had acquired the current state at least once earlier. The timer is not triggered if LACP attempts to transition to the current state for the first time. LACP monitors the PDUs received on the child link but prevents the link from transitioning to current state. If no flapping is observed when the link receives the PDUs, the hold-up timer expiries and triggers the member link to transition back to the current state. This transition is triggered as soon as the hold-up timer expires and not necessarily when the link receives a PDU.

To configure LACP hold-up timer for LAG interface, use the `hold-time up` statement at the `[edit interfaces aex aggregated-ether-options lacp]` hierarchy level.



NOTE:

- The hold-up timer keeps running even when the interface that receives the LACP PDU moves to the port disable state. The timer is then restarted if, before the timer expires, the interface comes up again and receives an LACP PDU from its neighbor. This ensures that the timer is maintained even during a quick physical port flap.
- When the following events occur, a hold-up timer is not triggered until the member link acquires the current state after the event:
  - LACP daemon restart

- Deactivation and reactivation of child or aggregated Ethernet interface
- Deletion and reconfiguration of child or aggregated Ethernet interface
- System reboot
- Routing Engine switchover

Verifying That LACP Is Configured Correctly and Bundle Members Are Exchanging LACP Protocol Packets

IN THIS SECTION

- Verifying the LACP Setup | 330
- Verifying That LACP Packets Are Being Exchanged | 331

Verify that LACP has been set up correctly and that the bundle members are transmitting LACP protocol packets.

Verifying the LACP Setup

IN THIS SECTION

- Purpose | 330
- Action | 331
- Meaning | 331

Purpose

Verify that the LACP has been set up correctly.

Action

To verify that LACP has been enabled as active on one end:

```
user@switch>show lacp interfaces xe-0/1/0
Aggregated interface: ae0
    LACP state:       Role   Exp   Def  Dist  Col  Syn  Aggr  Timeout  Activity
      xe-0/1/0       Actor    No    No   Yes  Yes  Yes   Yes     Fast    Active
      xe-0/1/0     Partner    No    No   Yes  Yes  Yes   Yes     Fast   Passive
    LACP protocol:        Receive State  Transmit State          Mux State
      xe-0/1/0              Current   Fast periodic Collecting distributing
```

Meaning

This example shows that LACP has been configured with one side as active and the other as passive. When LACP is enabled, one side must be set as active in order for the bundled link to be up.

Verifying That LACP Packets Are Being Exchanged

IN THIS SECTION

- Purpose | 331
- Action | 331
- Meaning | 332

Purpose

Verify that LACP packets are being exchanged between interfaces.

Action

Use the `show lacp statistics interfaces` *`interface-name`* command to display LACP BPDU exchange information.

```
show lacp statistics interfaces ae0
Aggregated interface: ae0
    LACP Statistics:       LACP Rx     LACP Tx   Unknown Rx   Illegal Rx
```

```
    xe-0/0/2              1352        2035          0          0
    xe-0/0/3              1352        2056          0          0
```

Meaning

The output here shows that the link is up and that PDUs are being exchanged.

RELATED DOCUMENTATION

Verifying the Status of a LAG Interface

show lacp statistics interfaces (View)

Example: Configuring Aggregated Ethernet High-Speed Uplinks with LACP Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch

IN THIS SECTION

- Requirements | **333**
- Overview and Topology | **333**
- Configuring LACP for the LAGs on the Virtual Chassis Access Switch | **334**
- Configuring LACP for the LAGs on the Virtual Chassis Distribution Switch | **335**
- Verification | **336**
- Troubleshooting | **339**

EX Series switches allow you to combine multiple Ethernet links into one logical interface for higher bandwidth and redundancy. The ports that are combined in this manner are referred to as a link aggregation group (LAG) or bundle. EX Series switches allow you to further enhance these links by configuring Link Aggregation Control Protocol (LACP).

This example describes how to overlay LACP on the LAG configurations that were created in *Example: Configuring Aggregated Ethernet High-Speed Uplinks Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch*:

Requirements

This example uses the following software and hardware components:

- Junos OS Release 9.0 or later for EX Series switches
- Two EX4200-48P switches
- Two EX4200-24F switches
- Four EX Series XFP uplink modules

Before you configure LACP, be sure you have:

- Set up the Virtual Chassis switches. See *Configuring an EX4200, EX4500, or EX4550 Virtual Chassis (CLI Procedure)*.
- Configured the uplink ports on the switches as trunk ports. See Configuring Gigabit Ethernet Interfaces (CLI Procedure).
- Configured the LAGs. See *Example: Configuring Aggregated Ethernet High-Speed Uplinks Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch*.

Overview and Topology

This example assumes that you are familiar with *Example: Configuring Aggregated Ethernet High-Speed Uplinks Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch*. The topology in this example is exactly the same as the topology in that other example. This example shows how to use LACP to enhance the LAG functionality.

LACP exchanges are made between *actors* (the transmitting link) and *partners* (the receiving link). The LACP mode can be either active or passive.



NOTE: If the actor and partner are both in passive mode, they do not exchange LACP packets, which results in the aggregated Ethernet links not coming up. By default, LACP is in passive mode. To initiate transmission of LACP packets and responses to LACP packets, you must enable LACP in active mode.

By default, the actor and partner send LACP packets every second.

The interval can be fast (every second) or slow (every 30 seconds).

Configuring LACP for the LAGs on the Virtual Chassis Access Switch

IN THIS SECTION

- Procedure | 334

To configure LACP for the access switch LAGs, perform these tasks.

Procedure

CLI Quick Configuration

To quickly configure LACP for the access switch LAGs, copy the following commands and paste them into the switch terminal window:

```
[edit]
set interfaces ae0 aggregated-ether-options lacp active periodic fast
                        set interfaces ae1 aggregated-ether-options lacp active periodic
fast
```

Step-by-Step Procedure

To configure LACP for Host-A LAGs `ae0` and `ae1`:

1. Specify the aggregated Ethernet options for both bundles:

```
[edit interfaces]
user@Host-A#set ae0 aggregated-ether-options lacp active periodic fast
user@Host-A#set ae1 aggregated-ether-options lacp active periodic fast
```

Results

Display the results of the configuration:

```
[edit interfaces]
user@Host-A# show
```

```
ae0 {
    aggregated-ether-options {
        lacp {
            active;
            periodic fast;
        }
    }
}
ae1 {
    aggregated-ether-options {
        lacp {
            active;
            periodic fast;
        }
    }
}
```

Configuring LACP for the LAGs on the Virtual Chassis Distribution Switch

IN THIS SECTION

- Procedure | 335

To configure LACP for the two uplink LAGs from the Virtual Chassis access switch to the Virtual Chassis distribution switch, perform these tasks.

Procedure

CLI Quick Configuration

To quickly configure LACP for the distribution switch LAGs, copy the following commands and paste them into the switch terminal window:

```
[edit interfaces]
                        set ae0 aggregated-ether-options lacp passive periodic fast
                        set ae1 aggregated-ether-options lacp passive periodic
fast
```

Step-by-Step Procedure

To configure LACP for Host D LAGs `ae0` and `ae1`:

1. Specify the aggregated Ethernet options for both bundles:

```
[edit interfaces]
user@Host-D#set ae0 aggregated-ether-options lacp passive periodic fast
user@Host-D#set ae1 aggregated-ether-options lacp passive periodic fast
```

Results

Display the results of the configuration:

```
[edit interfaces]
user@Host-D# show
ae0 {
    aggregated-ether-options {
        lacp {
            passive;
            periodic fast;
        }
    }
}
ae1 {
    aggregated-ether-options {
        lacp {
            passive
            periodic fast;
        }
    }
}
```

Verification

IN THIS SECTION

Verifying the LACP Settings | 337

- Verifying That the LACP Packets Are Being Exchanged | 338

To verify that LACP packets are being exchanged, perform these tasks:

Verifying the LACP Settings

Purpose

Verify that LACP has been set up correctly.

Action

Use the `show lacp interfaces` *`interface-name`* command to check that LACP has been enabled as active on one end.

```
user@Host-A> show lacp interfaces xe-0/1/0

Aggregated interface: ae0

    LACP state:       Role   Exp   Def  Dist  Col  Syn  Aggr  Timeout  Activity

      xe-0/1/0      Actor    No   Yes    No   No   No   Yes     Fast    Active

      xe-0/1/0    Partner    No   Yes    No   No   No   Yes     Fast   Passive

    LACP protocol:   Receive State    Transmit State          Mux State

      xe-0/1/0          Defaulted     Fast periodic           Detached
```

Meaning

The output indicates that LACP has been set up correctly and is active at one end.

Verifying That the LACP Packets Are Being Exchanged

Purpose

Verify that LACP packets are being exchanged.

Action

Use the `show interfaces ae`*x* `statistics` command to display LACP information.

```
user@Host-A> show interfaces ae0 statistics

Physical interface: ae0, Enabled, Physical link is Down
  Interface index: 153, SNMP ifIndex: 30
  Link-level type: Ethernet, MTU: 1514, Speed: Unspecified, Loopback: Disabled,
  Source filtering: Disabled, Flow control: Disabled, Minimum links needed: 1,
  Minimum bandwidth needed: 0
  Device flags   : Present Running
  Interface flags: Hardware-Down SNMP-Traps Internal: 0x0
  Current address: 02:19:e2:50:45:e0, Hardware address: 02:19:e2:50:45:e0
  Last flapped   : Never
  Statistics last cleared: Never
    Input packets : 0
    Output packets: 0
  Input errors: 0, Output errors: 0

  Logical interface ae0.0 (Index 71) (SNMP ifIndex 34)
    Flags: Hardware-Down Device-Down SNMP-Traps Encapsulation: ENET2
    Statistics        Packets        pps         Bytes          bps
    Bundle:
        Input :             0          0             0            0
        Output:             0          0             0            0
    Protocol inet
      Flags: None
      Addresses, Flags: Dest-route-down Is-Preferred Is-Primary
        Destination: 10.10.10/24, Local: 10.10.10.1, Broadcast: 10.10.10.255
```

Meaning

The output here shows that the link is down and that no protocol data units (PDUs) are being exchanged.

Troubleshooting

IN THIS SECTION

- Troubleshooting a Nonworking LACP Link | 339

To troubleshoot a nonworking LACP link, perform these tasks:

Troubleshooting a Nonworking LACP Link

Problem

The LACP link is not working.

Solution

Check the following:

- Remove the LACP configuration and verify whether the static LAG is up.
- Verify that LACP is configured at both ends.
- Verify that LACP is not passive at both ends.
- Verify whether LACP protocol data units (PDUs) are being exchanged by running the `monitor traffic-interface` *`lag-member`* `detail` command.

SEE ALSO

Example: Connecting an EX Series Access Switch to a Distribution Switch

Virtual Chassis Cabling Configuration Examples for EX4200 Switches

Installing an Uplink Module in an EX4200 Switch

Example: Configuring Link Aggregation with LACP Between a QFX Series Product and an Aggregation Switch

IN THIS SECTION

- Requirements | 340
- Overview and Topology | 341
- Configuring LACP for the LAG on the QFX Series | 341
- Verification | 342
- Troubleshooting | 345

QFX Series products allow you to combine multiple Ethernet links into one logical interface for higher bandwidth and redundancy. The ports that are combined in this manner are referred to as a link aggregation group (LAG) or bundle. The number of Ethernet links you can combine into a LAG depends on your QFX Series product model. On a standalone switch, you can group up to 32 Ethernet interfaces to form a LAG. On a QFabric system, you can group up to 8 Ethernet interfaces to form a LAG. QFX Series products allow you to further enhance these links by configuring Link Aggregation Control Protocol (LACP).

This example describes how to overlay LACP on the LAG configurations that were created in "Example: Configuring Link Aggregation Between a QFX Series Product and an Aggregation Switch" on page 305:

Requirements

This example uses the following software and hardware components:

- Junos OS Release 11.1 or later for the QFX3500 switch, Junos OS Release 12.1 or later for the QFX3600 switch, Junos OS Release 13.2 or later for the QFX5100 switch, and Junos OS Release 15.1X53-D10 or later for the QFX10002 switch.
- One QFX3500, QFX3600, QFX5100, QFX10002 switch.

Before you configure LACP, be sure you have:

- Configured the ports on the switches as trunk ports.
- Configured the LAG.

Overview and Topology

The topology in this example is exactly the same as the topology used in the Configuring a LAG Between a QFX Switch and an Aggregation Switch example. This example shows how to use LACP to enhance the LAG functionality.

LACP exchanges are made between *actors* (the transmitting link) and *partners* (the receiving link). The LACP mode can be either active or passive.



NOTE: If the actor and partner are both in passive mode, they do not exchange LACP packets, which results in the aggregated Ethernet links not coming up. By default, LACP is in passive mode. To initiate transmission of LACP packets and responses to LACP packets, you must enable LACP in active mode.

By default, the actor and partner send LACP packets every second. You can configure the interval at which the interfaces send LACP packets by including the `periodic` statement at the `[edit interfaces` *`interface-name`* `aggregated-ether-options lacp]` hierarchy level.

The interval can be fast (every second) or slow (every 30 seconds).

Configuring LACP for the LAG on the QFX Series

IN THIS SECTION

- Procedure | 341

To configure LACP for a QFX Series LAG, perform these tasks.

Procedure

CLI Quick Configuration

To quickly configure LACP for the access switch LAGs, copy the following commands and paste them into the switch terminal window:

```
[edit]
set interfaces ae0 aggregated-ether-options lacp active periodic fast
```

Step-by-Step Procedure

To configure LACP for LAG `ae0` :

1. Specify the aggregated Ethernet options for the LAG:

```
[edit interfaces]
user@switch# set ae0 aggregated-ether-options lacp active periodic fast
```

Results

Display the results of the configuration:

```
[edit interfaces]
user@switch# show
ae0 {
    aggregated-ether-options {
        lacp {
            active;
            periodic fast;
        }
    }
}
```

Verification

IN THIS SECTION

- Verifying the LACP Settings | 343
- Verifying That the LACP Packets Are Being Exchanged | 343

To verify that LACP packets are being exchanged, perform the following tasks:

Verifying the LACP Settings

Purpose

Verify that LACP has been set up correctly.

Action

Use the `show lacp interfaces` *`interface-name`* command to check that LACP has been enabled as active on one end.

```
user@switch> show lacp interfaces xe-0/02

Aggregated interface: ae0

    LACP state:       Role   Exp   Def  Dist  Col  Syn  Aggr  Timeout  Activity

      xe-0/0/2      Actor    No   Yes    No   No   No   Yes     Fast    Active

      xe-0/0/2    Partner    No   Yes    No   No   No   Yes     Fast   Passive

    LACP protocol:   Receive State    Transmit State          Mux State

      xe-0/0/2          Defaulted     Fast periodic           Detached
```

Meaning

The output indicates that LACP has been set up correctly and is active at one end.

Verifying That the LACP Packets Are Being Exchanged

Purpose

Verify that LACP packets are being exchanged.

Action

Use the `show interfaces ae`*x* `statistics` command to display LACP information.

```
user@switch> show interfaces ae0 statistics

Physical interface: ae0, Enabled, Physical link is Down
  Interface index: 153, SNMP ifIndex: 30
  Link-level type: Ethernet, MTU: 1514, Speed: Unspecified, Loopback: Disabled,
  Source filtering: Disabled, Flow control: Disabled, Minimum links needed: 1,
  Minimum bandwidth needed: 0
  Device flags   : Present Running
  Interface flags: Hardware-Down SNMP-Traps Internal: 0x0
  Current address: 02:19:e2:50:45:e0, Hardware address: 02:19:e2:50:45:e0
  Last flapped   : Never
  Statistics last cleared: Never
    Input packets : 0
    Output packets: 0
  Input errors: 0, Output errors: 0

  Logical interface ae0.0 (Index 71) (SNMP ifIndex 34)
    Flags: Hardware-Down Device-Down SNMP-Traps Encapsulation: ENET2
    Statistics        Packets        pps         Bytes          bps
    Bundle:
        Input :             0          0             0            0
        Output:             0          0             0            0
    Protocol inet
      Flags: None
      Addresses, Flags: Dest-route-down Is-Preferred Is-Primary
        Destination: 10.10.10/8, Local: 10.10.10.1, Broadcast: 10.10.10.255
```

Meaning

The output here shows that the link is down and that no PDUs are being exchanged.

Troubleshooting

IN THIS SECTION

- Troubleshooting a Nonworking LACP Link | 345

To troubleshoot a nonworking LACP link, perform these tasks:

Troubleshooting a Nonworking LACP Link

Problem

The LACP link is not working.

Solution

Check the following:

- Remove the LACP configuration and verify whether the static LAG is up.
- Verify that LACP is configured at both ends.
- Verify that LACP is not passive at both ends.
- Verify whether LACP protocol data units (PDUs) are being exchanged by running the `monitor traffic-interface` *`lag-member`* `detail` command.

SEE ALSO

Verifying the Status of a LAG Interface

Example: Configuring Link Aggregation Between a QFX Series Product and an Aggregation Switch | 305

Example: Configuring an FCoE LAG on a Redundant Server Node Group

show lacp statistics interfaces (View)

Understanding Independent Micro BFD Sessions for LAG

IN THIS SECTION

- Configuration Guidelines for Micro-BFD Sessions | 347

The Bidirectional Forwarding Detection (BFD) protocol is a simple detection protocol that quickly detects failures in the forwarding paths. To enable failure detection for aggregated Ethernet interfaces in a LAG, you can configure an independent, asynchronous-mode BFD session on every LAG member link in a LAG bundle. Instead of a single BFD session monitoring the status of the UDP port, independent micro-BFD sessions monitor the status of individual member links.

When you configure micro-BFD sessions on every member link in a LAG bundle, each individual session determines the Layer 2 and Layer 3 connectivity of each member link in a LAG.

After the individual session is established on a particular link, member links are attached to the LAG and then load balanced by either one of the following:

- Static configuration—The device control process acts as the client to the micro-BFD session.
- Link Aggregation Control Protocol (LACP)—LACP acts as the client to the micro-BFD session.

When the micro-BFD session is up, a LAG link is established and data is transmitted over that LAG link. If the micro-BFD session on a member link is down, that particular member link is removed from the load balancer, and the LAG managers stop directing traffic to that link. These micro-BFD sessions are independent of each other despite having a single client that manages the LAG interface.

Micro-BFD sessions run in the following modes:

- Distribution mode—In this mode, the Packet Forwarding Engine (PFE) sends and receives the packets at Layer 3. By default, micro-BFD sessions are distributed at Layer 3.
- Non-distribution mode—In this mode, the Routing Engine sends and receives the packets at Layer 2. You can configure the BFD session to run in this mode by including the `no-delegate-processing` statement under periodic packet management (PPM).

A pair of routing devices in a LAG exchange BFD packets at a specified, regular interval. The routing device detects a neighbor failure when it stops receiving a reply after a specified interval. This allows the quick verification of member link connectivity with or without LACP. A UDP port distinguishes BFD over LAG packets from BFD over single-hop IP packets. The Internet Assigned Numbers Authority (IANA) has allocated 6784 as the UDP destination port for micro-BFD.

Benefits

- Failure detection for LAG—Enables failure detection between devices that are in point-to-point connections.
- Multiple BFD sessions—Enables you to configure multiple micro-BFD sessions for each member link instead of a single BFD session for the entire bundle.

Configuration Guidelines for Micro-BFD Sessions

Consider the following guidelines as you configure individual micro-BFD sessions on an aggregated Ethernet bundle.

- This feature works only when both the devices support BFD. If BFD is configured at one end of the LAG, this feature does not work.
- Starting with Junos OS Release 13.3, IANA has allocated 01-00-5E-90-00-01 as the dedicated MAC address for micro BFD. Dedicated MAC mode is used by default for micro BFD sessions.
- In Junos OS, micro-BFD control packets are always untagged by default. For Layer 2 aggregated interfaces, the configuration must include `vlan-tagging` or `flexible-vlan-tagging` options when you configure Aggregated Ethernet with BFD. Otherwise, the system will throw an error while committing the configuration.
- When you enable micro-BFD on an aggregated Ethernet interface, the aggregated interface can receive micro-BFD packets. In Junos OS Release 19.3 and later, for MPC10E and MPC11E MPCs, you cannot apply firewall filters on the micro-BFD packets received on the aggregated Ethernet interface. For MPC1E through MPC9E, you can apply firewall filters on the micro-BFD packets received on the aggregated Ethernet interface only if the aggregated Ethernet interface is configured as an untagged interface.
- Starting with Junos OS Release 14.1, specify the neighbor in a BFD session. In releases before Junos OS Release 16.1, you must configure the loopback address of the remote destination as the neighbor address. Beginning with Junos OS Release 16.1, you can also configure this feature on MX Series routers with aggregated Ethernet interface address of the remote destination as the neighbor address.
- Beginning with Release 16.1R2, Junos OS checks and validates the configured micro-BFD `local-address` against the interface or loopback IP address before the configuration commit. Junos OS performs this check on both IPv4 and IPv6 micro-BFD address configurations, and if they do not match, the commit fails. The configured micro-BFD local address should match with the micro-BFD neighbour address that you have configured on the peer router.
- For the IPv6 address family, disable duplicate address detection before configuring this feature with aggregated Ethernet interface addresses. To disable duplicate address detection, include the `dad-disable` statement at the `[edit interface` *`aex`* `unit` *`y`* `family inet6]` hierarchy level.

- Starting in Junos OS 21.4R1, LACP minimum link with sync reset and microBFD configuration is supported on PTX10001-36MR, PTX10003, PTX10004, PTX10008, and PTX10016 routers.

> **CAUTION**: Deactivate `bfd-liveness-detection` at the `[edit interfaces aex aggregated-ether-options]` hierarchy level or deactivate the aggregated Ethernet interface before changing the neighbor address from the loopback IP address to the aggregated Ethernet interface IP address. Modifying the local and neighbor address without deactivating `bfd-liveness-detection` or the aggregated Ethernet interface first might cause micro-BFD sessions failure.

SEE ALSO

authentication

bfd-liveness-detection

detection-time

transmit-interval

Configuring Micro BFD Sessions for LAG

The Bidirectional Forwarding Detection (BFD) protocol is a simple detection protocol that quickly detects failures in the forwarding paths. A link aggregation group (LAG) combines multiple links between devices that are in point-to-point connections, thereby increasing bandwidth, providing reliability, and allowing load balancing. To run a BFD session on LAG interfaces, configure an independent, asynchronous mode BFD session on every LAG member link in a LAG bundle. Instead of a single BFD session monitoring the status of the UDP port, independent micro BFD sessions monitor the status of individual member links.

> **NOTE**: Starting in Junos OS Evolved Release 20.1R1, independent micro Bidirectional Forwarding Detection (BFD) sessions are enabled on a per member link basis of a Link Aggregation Group (LAG) bundle.

To enable failure detection for aggregated Ethernet interfaces:

1. Include the following statement in the configuration at the `[edit interfaces aex aggregated-ether-options]` hierarchy level:

```
bfd-liveness-detection
```

2. Configure the authentication criteria of the BFD session for LAG.

   To specify the authentication criteria, include the `authentication` statement:

```
bfd-liveness-detection {
    authentication {
        algorithm algorithm-name;
        key-chain key-chain-name;
        loose-check;
    }
}
```

   - Specify the algorithm to be used to authenticate the BFD session. You can use one of the following algorithms for authentication:
     - keyed-md5
     - keyed-sha-1
     - meticulous-keyed-md5
     - meticulous-keyed-sha-1
     - simple-password
   - To configure the key chain, specify the name that is associated with the security key for the BFD session. The name you specify must match one of the key chains configured in the `authentication-key-chains key-chain` statement at the `[edit security]` hierarchy level.
   - Configure loose authentication checking on the BFD session. Use only for transitional periods when authentication might not be configured at both ends of the BFD session.

3. Configure BFD timers for aggregated Ethernet interfaces.

   To specify the BFD timers, include the `detection-time` statement:

```
bfd-liveness-detection {
    detection-time {
        threshold  milliseconds;
```

```
    }
}
```

Specify the threshold value. This is the maximum time interval for detecting a BFD neighbor. If the transmit interval is greater than this value, the device triggers a trap.

4. Configure a hold-down interval value to set the minimum time that the BFD session must remain up before a state change notification is sent to the other members in the LAG network.

   To specify the hold-down interval, include the `holddown-interval` statement:

```
bfd-liveness-detection {
    holddown-interval milliseconds;
}
```

   You can configure a number in the range from 0 through 255,000 milliseconds, and the default is 0. If the BFD session goes down and then comes back up during the hold-down interval, the timer is restarted.

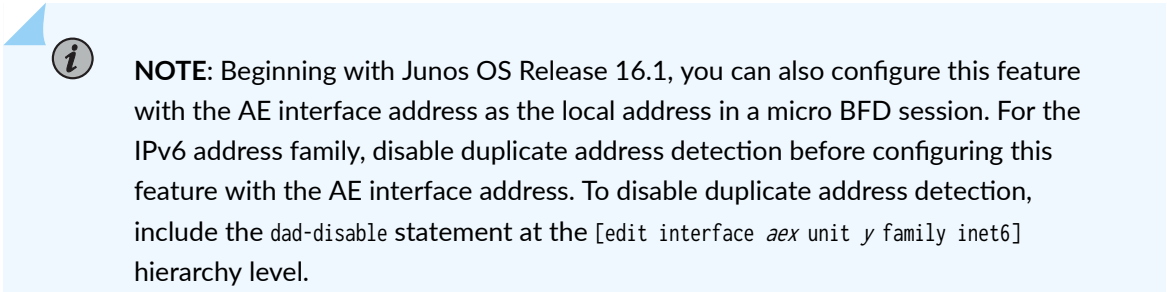
   This value represents the minimum interval at which the local routing device transmits BFD packets, as well as the minimum interval in which the routing device expects to receive a reply from a neighbor with which it has established a BFD session. You can configure a number in the range from 1 through 255,000 milliseconds. You can also specify the minimum transmit and receive intervals separately.

5. Configure the source address for the BFD session.

   To specify a local address, include the `local-address` statement:

```
bfd-liveness-detection {
    local-address bfd-local-address;
}
```

   The BFD local address is the loopback address of the source of the BFD session.



NOTE: Beginning with Junos OS Release 16.1, you can also configure this feature with the AE interface address as the local address in a micro BFD session. For the IPv6 address family, disable duplicate address detection before configuring this feature with the AE interface address. To disable duplicate address detection, include the `dad-disable` statement at the `[edit interface aex unit y family inet6]` hierarchy level.

Beginning with Release 16.1R2, Junos OS checks and validates the configured micro BFD `local-address` against the interface or loopback IP address before the configuration commit. Junos OS performs this check on both IPv4 and IPv6 micro BFD address configurations, and if they do not match, the commit fails. The configured micro-BFD `local-address` should match with the micro-BFD `neighbour-address` configured on the peer router.

6. Specify the minimum interval that indicates the time interval for transmitting and receiving data.

   This value represents the minimum interval at which the local routing device transmits BFD packets, as well as the minimum interval in which the routing device expects to receive a reply from a neighbor with which it has established a BFD session. You can configure a number in the range from 1 through 255,000 milliseconds. You can also specify the minimum transmit and receive intervals separately.

   To specify the minimum transmit and receive intervals for failure detection, include the `minimum-interval` statement:

```
bfd-liveness-detection {
    minimum-interval milliseconds;
}
```



NOTE: BFD is an intensive protocol that consumes system resources. Specifying a minimum interval for BFD less than 100 ms for Routing Engine-based sessions and 10 ms for distributed BFD sessions can cause undesired BFD flapping.

Depending on your network environment, these additional recommendations might apply:

- For large-scale network deployments with a large number of BFD sessions, specify a minimum interval of 300 ms for Routing Engine-based sessions and 100 ms for distributed BFD sessions.
- For very large-scale network deployments with a large number of BFD sessions, contact Juniper Networks customer support for more information.
- For BFD sessions to remain up during a Routing Engine switchover event when nonstop active routing is configured, specify a minimum interval of 2500 ms for Routing Engine-based sessions. For distributed BFD sessions with nonstop active routing configured, the minimum interval recommendations are unchanged and depend only on your network deployment.

7. Specify only the minimum receive interval for failure detection by including the `minimum-receive-interval` statement:

```
bfd-liveness-detection {
    minimum-receive-interval milliseconds;
}
```

This value represents the minimum interval in which the local routing device expects to receive a reply from a neighbor with which it has established a BFD session. You can configure a number in the range from 1 through 255,000 milliseconds.

8. Specify the number of BFD packets that were not received by the neighbor that causes the originating interface to be declared down by including the `multiplier` statement:

```
bfd-liveness-detection {
    multiplier number;
}
```

The default value is 3. You can configure a number in the range from 1 through 255.

9. Configure the neighbor in a BFD session.

The neighbor address can be either an IPv4 or an IPv6 address.

To specify the next hop of the BFD session, include the `neighbor` statement:

```
bfd-liveness-detection {
    neighbor bfd-neighbor-address;
}
```

The BFD neighbor address is the loopback address of the remote destination of the BFD session.



NOTE: Beginning with Junos OS Release 16.1, you can also configure the AE interface address of the remote destination as the BFD neighbor address in a micro BFD session.

10. (Optional) Configure BFD sessions not to adapt to changing network conditions.

To disable BFD adaptation, include the `no-adaptation` statement:

```
bfd-liveness-detection {
    no-adaptation;
}
```



NOTE: We recommend that you do not disable BFD adaptation unless it is preferable not to have BFD adaptation in your network.

11. Specify a threshold for detecting the adaptation of the detection time by including the `threshold` statement:

```
bfd-liveness-detection {
    detection-time {
        threshold milliseconds;
    }
}
```

When the BFD session detection time adapts to a value equal to or greater than the threshold, a single trap and a system log message are sent. The detection time is based on the multiplier of the minimum-interval or the minimum-receive-interval value. The threshold must be a higher value than the multiplier for either of these configured values. For example, if the minimum-receive-interval is 300 ms and the multiplier is 3, the total detection time is 900 ms. Therefore, the detection time threshold must have a value greater than 900.

12. Specify only the minimum transmit interval for failure detection by including the `transmit-interval minimum-interval` statement:

```
bfd-liveness-detection {
    transmit-interval {
        minimum-interval milliseconds;
    }
}
```

This value represents the minimum interval at which the local routing device transmits BFD packets to the neighbor with which it has established a BFD session. You can configure a value in the range from 1 through 255,000 milliseconds.

13. Specify the transmit threshold for detecting the adaptation of the transmit interval by including the `transmit-interval threshold` statement:

```
bfd-liveness-detection {
    transmit-interval {
        threshold milliseconds;
    }
}
```

The threshold value must be greater than the transmit interval. When the BFD session detection time adapts to a value greater than the threshold, a single trap and a system log message are sent. The detection time is based on the multiplier of the minimum-interval or the minimum-receive-interval value. The threshold must be a higher value than the multiplier for either of these configured values.

14. Specify the BFD version by including the `version` statement:

```
bfd-liveness-detection {
    version (1 | automatic);
}
```

The default is to have the version detected automatically.



NOTE:

- The `version` option is not supported on the QFX Series. Starting in Junos OS Release 17.2R1, a warning will appear if you attempt to use this command.
- This feature works when both the devices support BFD. If BFD is configured at only one end of the LAG, this feature does not work.

SEE ALSO

authentication

bfd-liveness-detection

detection-time

Example: Configuring Independent Micro BFD Sessions for LAG

Understanding the Algorithm Used to Hash LAG Bundle and Egress Next-Hop ECMP Traffic

IN THIS SECTION

- Understanding the Hashing Algorithm | 355
- IP (IPv4 and IPv6) | 356
- MPLS | 359
- MAC-in-MAC Packet Hashing | 361
- Layer 2 Header Hashing | 362
- Hashing Parameters | 363

Juniper Networks EX Series and QFX Series use a hashing algorithm to determine how to forward traffic over a link aggregation group (LAG) bundle or to the next-hop device when equal-cost multipath (ECMP) is enabled.

The hashing algorithm makes hashing decisions based on values in various packet fields, as well as on some internal values like source port ID and source device ID. You can configure some of the fields that are used by the hashing algorithm.



NOTE: Platform support depends on the Junos OS release in your installation.

This topic contains the following sections:

Understanding the Hashing Algorithm

The hashing algorithm is used to make traffic-forwarding decisions for traffic entering a LAG bundle or for traffic exiting a switch when ECMP is enabled.

For LAG bundles, the hashing algorithm determines how traffic entering a LAG bundle is placed onto the bundle's member links. The hashing algorithm tries to manage bandwidth by evenly load-balancing all incoming traffic across the member links in the bundle.

For ECMP, the hashing algorithm determines how incoming traffic is forwarded to the next-hop device.

The hashing algorithm makes hashing decisions based on values in various packet fields, as well as on some internal values like source port ID and source device ID. The packet fields used by the hashing

algorithm varies by the packet's EtherType and, in some instances, by the configuration on the switch. The hashing algorithm recognizes the following EtherTypes:

- IP (IPv4 and IPv6)
- MPLS
- MAC-in-MAC

Traffic that is not recognized as belonging to any of these EtherTypes is hashed based on the Layer 2 header. IP and MPLS traffic are also hashed based on the Layer 2 header when a user configures the hash mode as Layer 2 header.

You can configure some fields that are used by the hashing algorithm to make traffic forwarding decisions. You cannot, however, configure how certain values within a header are used by the hashing algorithm.

Note the following points regarding the hashing algorithm:

- The fields selected for hashing are based on the packet type only. The fields are not based on any other parameters, including forwarding decision (bridged or routed) or egress LAG bundle configuration (Layer 2 or Layer 3).
- The same fields are used for hashing unicast and multicast packets. Unicast and multicast packets are, however, hashed differently.
- The same fields are used by the hashing algorithm to hash ECMP and LAG traffic, but the hashing algorithm hashes ECMP and LAG traffic differently. LAG traffic uses a trunk hash while ECMP uses ECMP hashing. Both LAG and ECMP use the same RTAG7 seed but use different offsets of that 128B seed to avoid polarization. The initial config of the HASH function to use the trunk and ECMP offset are set at the PFE Init time. The different hashing ensures that traffic is not polarized when a LAG bundle is part of the ECMP next-hop path.
- The same fields are used for hashing regardless of whether the switch is or is not participating in a mixed or non-mixed Virtual Chassis or Virtual Chassis Fabric (VCF).

The fields used for hashing by each EtherType as well as the fields used by the Layer 2 header are discussed in the following sections.

IP (IPv4 and IPv6)

Payload fields in IPv4 and IPv6 packets are used by the hashing algorithm when IPv4 or IPv6 packets need to be placed onto a member link in a LAG bundle or sent to the next-hop device when ECMP is enabled.

The hash mode is set to Layer 2 payload field, by default. IPv4 and IPv6 payload fields are used for hashing when the hash mode is set to Layer 2 payload.

If the hash mode is configured to Layer 2 header, IPv4, IPv6, and MPLS packets are hashed using the Layer 2 header fields. If you want incoming IPv4, IPv6, and MPLS packets hashed by the source MAC address, destination MAC address, or EtherType fields, you must set the hash mode to Layer 2 header.

Table 75 on page 357 displays the IPv4 and IPv6 payload fields that are used by the hashing algorithm, by default.

- ✓—Field is used by the hashing algorithm, by default.
- X—Field is not used by the hashing algorithm, by default.
- (configurable)—Field can be configured to be used or not used by the hashing algorithm.

On EX2300 switches, following payload fields in IPv4 and IPv6 packets are used by the hashing algorithm when IPv4 or IPv6 packets need to be placed onto a member link in a LAG bundle or sent to the next-hop device when ECMP is enabled:

- For unicast traffic on LAG - SIP, DIP, L4SP, L4DP
- For known multicast traffic on LAG - Source IP, Destination IP, Ingress Mod Id, and Ingress Port Id
- For broadcast, unknown unicast, and unknown multicast traffic on LAG - Source MAC, Destination MAC, Ingress Mod Id, and Ingress Port Id
- ECMP load balancing: Destination IP, Layer 4 Source Port, and Layer 4 Destination Port

Table 75: IPv4 and IPv6 Hashing Fields

Fields	EX3400		EX4300		QFX5100		QFX5110 and QFX5120		QFX5200	
	LAG	ECMP	LAG	ECMP	LAG	ECMP	LAG	ECMP	LAG	ECMP
Source MAC	X	X	X	X	X	X	X	X	X	X
Destination MAC	X	X	X	X	X	X	X	X	X	X
EtherType	X	X	X	X	X	X	X	X	X	X

Table 75: IPv4 and IPv6 Hashing Fields *(Continued)*

Fields	EX3400		EX4300		QFX5100		QFX5110 and QFX5120		QFX5200	
VLAN ID	X (configurable)	X (configurable)	X (configurable)	X (configurable)	X (configurable)	X (configurable)	X (configurable)	X (configurable)	X (configurable)	X (configurable)
Source IP or IPv6	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)
Destination IP or IPv6	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)
Protocol (IPv4 only)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)
Next header (IPv6 only)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)
Layer 4 Source Port	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)

Table 75: IPv4 and IPv6 Hashing Fields *(Continued)*

Fields	EX3400		EX4300		QFX5100		QFX5110 and QFX5120		QFX5200	
Layer 4 Destination Port	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)
IPv6 Flow label (IPv6 only)	X	X	X	X	X	X	X	X	X	X
Ingress Mod Id	✓ (configurable)	X	X	X	X	X	X	X	X	X
Ingress Port Id	✓ (configurable)	X	X	X	X	X	X	X	X	X

MPLS

The hashing algorithm hashes MPLS packets using the source IP, destination IP, MPLS label 0, MPLS label 1, MPLS label 2, and MPLS 3 fields. On the QFX5110, QFX5120, and QFX5200 switches, LSR routers also support ECMP. ECMP uses these fields for hashing on an LSR router:

- Layer 3 VPN: MPLS Labels (top 3 labels), source IP, destination IP, and ingress port ID
- Layer 2 Circuit: MPLS Labels (top 3 labels) and ingress port ID

Table 76 on page 360 displays the MPLS payload fields that are used by the hashing algorithm, by default:

- ✓—Field is used by the hashing algorithm, by default.
- X—Field is not used by the hashing algorithm, by default.

The fields used by the hashing algorithm for MPLS packet hashing are not user-configurable.

The source IP and destination IP fields are not always used for hashing. For non-terminated MPLS packets, the payload is checked if the bottom of stack (BoS) flag is seen in the packet. If the payload is IPv4 or IPv6, then the IP source address and IP destination address fields are used for hashing along with the MPLS labels. If the BoS flag is not seen in the packet, only the MPLS labels are used for hashing.

Table 76: MPLS Hashing Fields

Field	EX3400	EX4300	QFX5100	QFX5110 and QFX5120	QFX5200
Source MAC	X	X	X	X	X
Destination MAC	X	X	X	X	X
EtherType	X	X	X	X	X
VLAN ID	X	X	X	X	X
Source IP	✓	✓	✓	✓	✓
Destination IP	✓	✓	✓	✓	✓
Protocol (for IPv4 packets)	X	X	X	X	X
Next header (for IPv6 packets)	X	X	X	X	X
Layer 4 Source Port	X	X	X	X	X
Layer 4 Destination Port	X	X	X	X	X
IPv6 Flow lab	X	X	X	X	X

Table 76: MPLS Hashing Fields *(Continued)*

Field	EX3400	EX4300		QFX5100	QFX5110 and QFX5120	QFX5200
MPLS label 0	X	✓		✓	✓	✓
MPLS label 1	✓	✓		✓	✓	✓
MPLS label 2	✓	✓		✓	✓	✓
MPLS label 3	✓	X		X	X	X
Ingress Port ID	✓ (LSR and L2Circuit)	X	X	X	✓ (LSR and L2Circuit)	✓ (LSR and L2Circuit)

MAC-in-MAC Packet Hashing

Packets using the MAC-in-MAC EtherType are hashed by the hashing algorithm using the Layer 2 payload source MAC, Layer 2 payload destination MAC, and Layer 2 payload EtherType fields. See Table 77 on page 362.

Hashing using the fields in the MAC-in-MAC EtherType packet is first supported on EX4300 switches in Release 13.2X51-D20. Hashing using the fields in the MAC-in-MAC EtherType is not supported on earlier releases.

The fields used by the hashing algorithm for MAC-in-MAC hashing are not user-configurable.

- ✓—Field is used by the hashing algorithm, by default.
- X—Field is not used by the hashing algorithm, by default.

Table 77: MAC-in-MAC Hashing Fields

Field	EX3400	EX4300	QFX5100	QFX5110 and QFX5120	QFX5200
Layer 2 Payload Source MAC	✓	✓	✓	✓	✓
Layer 2 Payload Destination MAC	✓	✓	✓	✓	✓
Layer 2 Payload EtherType	✓	✓	✓	✓	✓
Layer 2 Payload Outer VLAN	✓	X	X	X	X

Layer 2 Header Hashing

Layer 2 header fields are used by the hashing algorithm when a packet's EtherType is not recognized as IP (IPv4 or IPv6), MPLS, or MAC-in-MAC. The Layer 2 header fields are also used for hashing IPv4, IPv6, and MPLS traffic instead of the payload fields when the hash mode is set to Layer 2 header.

- ✓—Field is used by the hashing algorithm, by default.
- X—Field is not used by the hashing algorithm, by default.
- (configurable)—Field can be configured to be used or not used by the hashing algorithm.

Table 78: Layer 2 Header Hashing Fields

Field	EX3400	EX4300	QFX5100	QFX5110 and QFX5120	QFX5200
Source MAC	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)

Table 78: Layer 2 Header Hashing Fields *(Continued)*

Field	EX3400	EX4300	QFX5100	QFX5110 and QFX5120	QFX5200
Destination MAC	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)
EtherType	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)
VLAN ID	X (configurable)	X (configurable)	X (configurable)	✓ (configurable)	✓ (configurable)

Hashing Parameters

Starting in Junos OS Release 19.1R1, on the QFX5000 line of switches, you can change hashing parameters for the existing algorithms implemented. You can change the threshold of shared buffer pools for both ingress and egress buffer partitions and you can make changes to the hash function selection, hash algorithm, and other additional parameters. See Configuring Other Hashing Parameters later in this document.

Configuring the Fields in the Algorithm Used To Hash LAG Bundle and ECMP Traffic (CLI Procedure)

IN THIS SECTION

Configuring the Hashing Algorithm to Use Fields in the Layer 2 Header for Hashing | **364**

- Configuring the Hashing Algorithm to Use Fields in the IP Payload for Hashing | 365
- Configuring the Hashing Algorithm to Use Fields in the IPv6 Payload for Hashing | 365
- Configuring Other Hashing Parameters | 366

Juniper Networks EX Series and QFX Series switches use a hashing algorithm to determine how to forward traffic over a Link Aggregation group (LAG) bundle or to the next-hop device when equal-cost multipath (ECMP) is enabled.

The hashing algorithm makes hashing decisions based on values in various packet fields.. You can configure some of the fields that are used by the hashing algorithm.

Configuring the fields used by the hashing algorithm is useful in scenarios where most of the traffic entering the bundle is similar and the traffic needs to be managed in the LAG bundle. For instance, if the only difference in the IP packets for all incoming traffic is the source and destination IP address, you can tune the hashing algorithm to make hashing decisions more efficiently by configuring the algorithm to make hashing decisions using only those fields.



NOTE: Configuring the hash mode is not supported on QFX10002 and QFX10008 switches.

Configuring the Hashing Algorithm to Use Fields in the Layer 2 Header for Hashing

To configure the hashing algorithm to use fields in the Layer 2 header for hashing:

1. Configure the hash mode to Layer 2 header:

```
[edit forwarding-options enhanced-hash-key]
user@switch# set hash-mode layer2-header
```

   The default hash mode is Layer 2 payload. Therefore, this step must be performed if you have not previously configured the hash mode.

2. Configure the fields in the Layer 2 header that the hashing algorithm uses for hashing:

```
[edit forwarding-options enhanced-hash-key]
user@switch# set layer2 {no-destination-mac-address | no-ether-type | no-source-mac-address |
vlan-id}
```

By default, the hashing algorithm uses the values in the destination MAC address, Ethertype, and source MAC address fields in the header to hash traffic on the LAG. You can configure the hashing algorithm to not use the values in these fields by configuring `no-destination-mac-address`, `no-ether-type`, or `no-source-mac-address`.

You can also configure the hashing algorithm to include the VLAN ID field in the header by configuring the `vlan-id` option.

If you want the hashing algorithm to not use the Ethertype field for hashing:

```
[edit forwarding-options enhanced-hash-key]
user@switch# set layer2 no-ether-type
```

Configuring the Hashing Algorithm to Use Fields in the IP Payload for Hashing

To configure the hashing algorithm to use fields in the IP payload for hashing:

1. Configure the hash mode to Layer 2 payload:

   ```
   [edit forwarding-options enhanced-hash-key]
   user@switch# set hash-mode layer2-payload
   ```

   The IP payload is not checked by the hashing algorithm unless the hash mode is set to Layer 2 payload. The default hash mode is Layer 2 payload.
2. Configure the fields in the IP payload that the hashing algorithm uses for hashing:

   ```
   [edit forwarding-options enhanced-hash-key]
   user@switch# set inet {no-ipv4-destination-address | no-ipv4-source-address | no-l4-
   destination-port | no-l4-source-port | no-protocol | vlan-id}
   ```

   For instance, if you want the hashing algorithm to ignore the Layer 4 destination port, Layer 4 source port, and protocol fields and instead hash traffic based only on the IPv4 source and destination addresses:

   ```
   [edit forwarding-options enhanced-hash-key]
   user@switch# set inet no-l4-destination-port no-l4-source-port no-protocol
   ```

Configuring the Hashing Algorithm to Use Fields in the IPv6 Payload for Hashing

To configure the hashing algorithm to use fields in the IPv6 payload for hashing:

1. Configure the hash mode to Layer 2 payload:

```
[edit forwarding-options enhanced-hash-key]
user@switch# set hash-mode layer2-payload
```

   The IPv6 payload is not checked by the hashing algorithm unless the hash mode is set to Layer 2 payload. The default hash mode is Layer 2 payload.

2. Configure the fields in the IPv6 payload that the hashing algorithm uses for hashing:

```
[edit forwarding-options enhanced-hash-key]
user@switch# set inet6 {no-ipv6-destination-address | no-ipv6-source-address | no-l4-
destination-port | no-l4-source-port | no-next-header | vlan-id}
```

   For instance, if you want the hashing algorithm to ignore the Layer 4 destination port, Layer 4 source port, and the Next Header fields and instead hash traffic based only on the IPv6 source and IPv6 destination address fields only:

```
[edit forwarding-options enhanced-hash-key]
user@switch# set inet6 no-l4-destination-port no-l4-source-port no-next-header
```

Configuring Other Hashing Parameters

To configure hashing parameters for either ECMP or LAG traffic:

1. Configure the preprocess parameter:

```
[edit forwarding-options enhanced-hash-key]
user@switch# set hash-parameters (ecmp | lag) preprocess
```

2. Configure the function parameter:

```
[edit forwarding-options enhanced-hash-key]
user@switch# set hash-parameters (ecmp | lag) function  (crc16-bisync | crc16-ccitt | crc32-
hi | crc32-lo)
```

3. Configure the offset value:

```
[edit forwarding-options enhanced-hash-key]
user@switch# set hash-parameters (ecmp | lag) offset offset-value
```

RELATED DOCUMENTATION

Understanding the Algorithm Used to Hash LAG Bundle and Egress Next-Hop ECMP Traffic | 355

Understanding the Algorithm Used to Hash LAG Bundle and Egress Next-Hop ECMP Traffic (QFX 10002 and QFX 10008 Switches)

Change History Table

Feature support is determined by the platform and release you are using. Use Feature Explorer to determine if a feature is supported on your platform.

Release	Description
19.3	Starting with Junos OS Release 19.3 and later, for MPC10E and MPC11E MPCs, you cannot apply firewall filters on the MicroBFD packets received on the aggregated Ethernet Interface. For MPC1E through MPC9E, you can apply firewall filters on the MicroBFD packets received on the aggregated Ethernet Interface only if the aggregated Ethernet Interface is configured as an untagged Interface.
19.1R1	on the QFX5000 line of switches, you can change hashing parameters for the existing algorithms implemented.
16.1	Beginning with Junos OS Release 16.1, you can also configure this feature on MX series routers with aggregated Ethernet interface address of the remote destination as the neighbor address.
16.1	Beginning with Release 16.1R2, Junos OS checks and validates the configured micro BFD `local-address` against the interface or loopback IP address before the configuration commit.
14.1X53-D25	Starting in Junos OS Release 14.1X53-D25, local link bias can be enabled globally for all LAG bundles in a Virtual Chassis or VCF, or individually per LAG bundle in a Virtual Chassis.
14.1	Starting with Junos OS Release 14.1, specify the neighbor in a BFD session. In releases prior to Junos OS Release 16.1, you must configure the loopback address of the remote destination as the neighbor address.
13.3	Starting with Junos OS Release 13.3, IANA has allocated 01-00-5E-90-00-01 as the dedicated MAC address for micro BFD.

Load Balancing for Aggregated Ethernet Interfaces

IN THIS SECTION

- Load Balancing and Ethernet Link Aggregation Overview | **368**
- Configuring Load Balancing Based on MAC Addresses | **369**
- Configuring Load Balancing on a LAG Link | **371**
- Example: Configuring Load Balancing on a LAG Link | **371**
- Understanding Multicast Load Balancing on Aggregated 10-Gigabit Links for Routed Multicast Traffic on EX8200 Switches | **372**
- Example: Configuring Multicast Load Balancing for Use with Aggregated 10-Gigabit Ethernet Interfaces on EX8200 Switches | **377**

Load balancing is done on Layer 2 across the member links making the configuration better without congestion and maintaining redundancy. The below topics discuss the overview of load balancing, configuring load balancing based on MAC addresses and on LAG link, understanding the consistency through resilient hashing.

Load Balancing and Ethernet Link Aggregation Overview

You can create a link aggregation group (LAG) for a group of Ethernet ports. Layer 2 bridging traffic is load balanced across the member links of this group, making the configuration attractive for congestion concerns as well as for redundancy. Each LAG bundle contains up to 16 links. (Platform support depends on the Junos OS release in your installation.)

For LAG bundles, the hashing algorithm determines how traffic entering a LAG bundle is placed onto the bundle's member links. The hashing algorithm tries to manage bandwidth by evenly load-balancing all incoming traffic across the member links in the bundle. The hash-mode of the hashing algorithm is set to Layer 2 payload by default. When the hash-mode is set to Layer 2 payload, the hashing algorithm uses the IPv4 and IPv6 payload fields for hashing. You can also configure the load balancing hash key for Layer 2 traffic to use fields in the Layer 3 and Layer 4 headers using the `payload` statement. However, note that the load-balancing behavior is platform-specific and based on appropriate hash-key configurations.

For more information, see "Configuring Load Balancing on a LAG Link" on page 371.In a Layer 2 switch, one link is overutilized and other links are underutilized.

Configuring Load Balancing Based on MAC Addresses

IN THIS SECTION

- Platform-Specific MAC Address Based Load Balancing Behavior | 370

The hash key mechanism for load-balancing uses Layer 2 media access control (MAC) information such as frame source and destination address. To load-balance traffic based on Layer 2 MAC information, include the `multiservice` statement at the `[edit forwarding-options hash-key]` or `[edit chassis fpc` *`slot number`* `pic` *`PIC number`* `hash-key]` hierarchy level:

```
multiservice {
    source-mac;
    destination-mac;
    payload {
        ip {
            layer3-only;
            layer-3 (source-ip-only | destination-ip-only);
            layer-4;
            inner-vlan-id;
            outer-vlan-id;
        }
    }
}
```

Use Feature Expolorer to confirm platform and release support for specific features.

Review the "Platform-Specific MAC Address Based Load Balancing Behavior" on page 370 section for notes related to your platform.

To include the destination-address MAC information in the hash key, include the `destination-mac` option. To include the source-address MAC information in the hash key, include the `source-mac` option.

NOTE:

- Any packets that have the same source and destination address will be sent over the same path.
- You can configure per-packet load balancing to optimize EVPN traffic flows across multiple paths.
- Aggregated Ethernet member links will now use the physical MAC address as the source MAC address in 802.3ah OAM packets.

Platform-Specific MAC Address Based Load Balancing Behavior

Platform	Difference
ACX Series	• ACX7000 Series routers that support MAC address-based load balancing use symmetric hashing. For example, you need to configure both `source-mac` and `destination-mac` under "multiservice" options. You cannot use `source-mac` and `destination-mac` separately. Note the following about hashing on ACX7000 Series routers: • Do not support any default hashing. Load balancing does not happen if you do not configure "hash-key" option. Use the `[set forwarding-options hash-key family]` hierarchy. • Load balancing might or might not be symmetrical. Some links might carry more traffic than others. This traffic difference is based on the traffic profile. • Do not support weighted hashing.

SEE ALSO

multiservice

Configuring Load Balancing on a LAG Link

You can configure the load balancing hash key for Layer 2 traffic to use fields in the Layer 3 and Layer 4 headers inside the frame payload for load-balancing purposes using the `payload` statement. You can configure the statement to look at **layer-3** (and **source-ip-only** or **destination-ip-only** packet header fields) or **layer-4** fields. You configure this statement at the `[edit forwarding-options hash-key family multiservice]` hierarchy level.

You can configure Layer 3 or Layer 4 options, or both. The **source-ip-only** or **destination-ip-only** options are mutually exclusive. The `layer-3-only` statement is not available on MX Series routers.

By default, Junos implementation of 802.3ad balances traffic across the member links within an aggregated Ethernet bundle based on the Layer 3 information carried in the packet.

For more information about link aggregation group (LAG) configuration, see the Junos OS Network Interfaces Library for Routing Devices.

Example: Configuring Load Balancing on a LAG Link

This example configures the load-balancing hash key to use the source Layer 3 IP address option and Layer 4 header fields as well as the source and destination MAC addresses for load balancing on a link aggregation group (LAG) link:

```
[edit]
forwarding-options {
    hash-key {
        family multiservice {
            source-mac;
            destination-mac;
            payload {
                ip {
                    layer-3 {
                        source-ip-only;
                    }
                    layer-4;
                }
            }
        }
    }
}
```

NOTE: Any change in the hash key configuration requires a reboot of the FPC for the changes to take effect.

Understanding Multicast Load Balancing on Aggregated 10-Gigabit Links for Routed Multicast Traffic on EX8200 Switches

IN THIS SECTION

- Create LAGs for Multicasting in Increments of 10 Gigabits | **373**
- When Should I Use Multicast Load Balancing? | **375**
- How Does Multicast Load Balancing Work? | **375**
- How Do I Implement Multicast Load Balancing on an EX8200 Switch? | **376**

Streaming video technology was introduced in 1997. Multicast protocols were subsequently developed to reduce data replication and network overloads. With multicasting, servers can send a single stream to a group of recipients instead of sending multiple unicast streams. While the use of streaming video technology was previously limited to occasional company presentations, multicasting has provided a boost to the technology resulting in a constant stream of movies, real-time data, news clips, and amateur videos flowing nonstop to computers, TVs, tablets, and phones. However, all of these streams quickly overwhelmed the capacity of network hardware and increased bandwidth demands leading to unacceptable blips and stutters in transmission.

To satisfy the growing bandwidth demands, multiple links were virtually aggregated to form bigger logical point-to-point link channels for the flow of data. These virtual link combinations are called multicast interfaces, also known as link aggregation groups (LAGs).

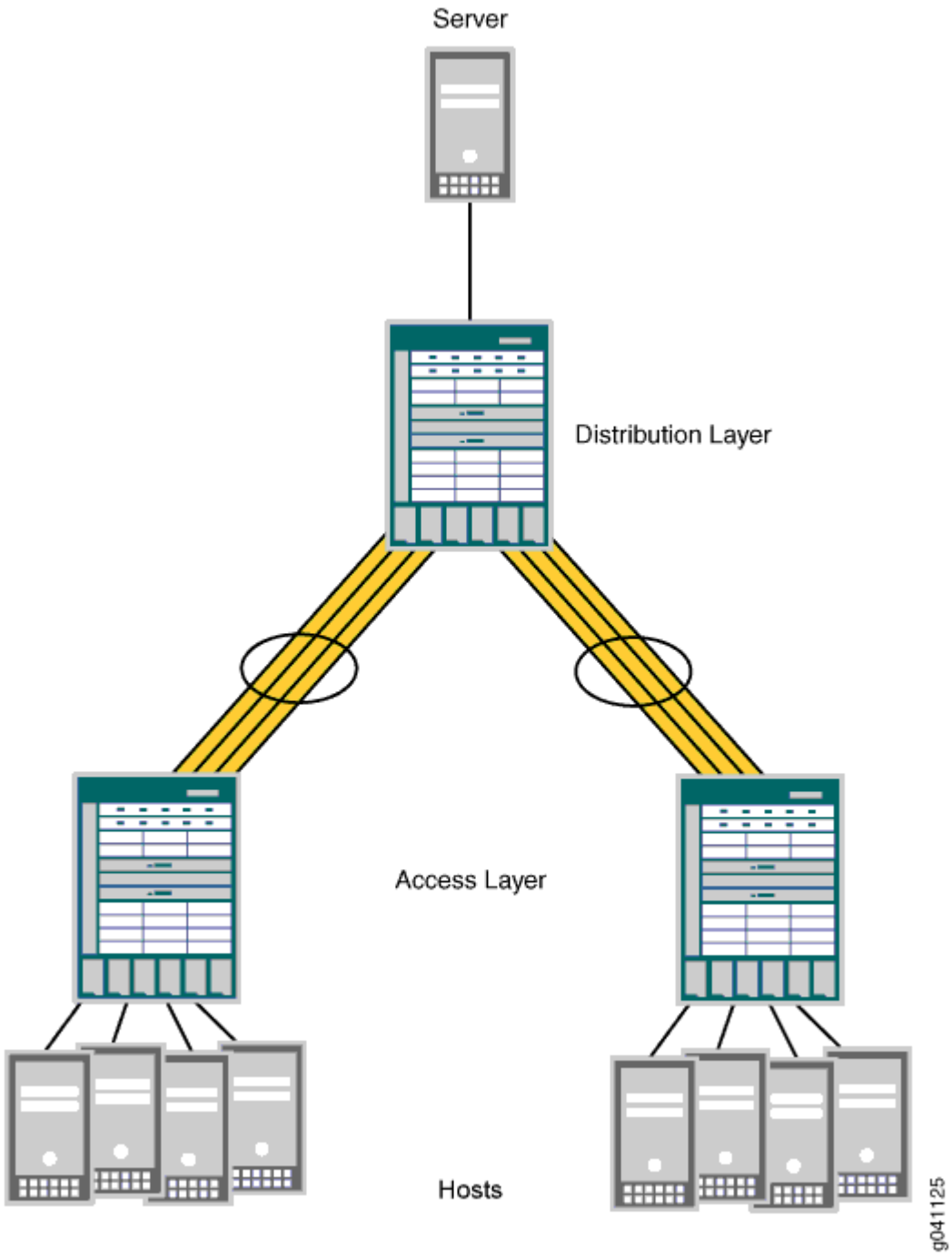
Multicast load balancing involves managing the individual links in each LAG to ensure that each link is used efficiently. Hashing algorithms continually evaluate the data stream, adjusting stream distribution over the links in the LAG, so that no link is underutilized or overutilized. Multicast load balancing is enabled by default on Juniper Networks EX8200 Ethernet Switches.

This topic includes:

Create LAGs for Multicasting in Increments of 10 Gigabits

The maximum link size on an EX8200 switch is 10 gigabits. If you need a larger link on an EX8200 switch, you can combine up to twelve 10-gigabit links. In the sample topology shown in Figure 16 on page 374, four 10-gigabit links have been aggregated to form each 40-gigabit link.

Figure 16: 40-Gigabit LAGs on EX8200 Switches



When Should I Use Multicast Load Balancing?

Use a LAG with multicast load balancing when you need a downstream link greater than 10 gigabits. This need frequently arises when you act as a service provider or when you multicast video to a large audience.

To use multicast load balancing, you need the following:

- An EX8200 switch—Standalone switches support multicast load balancing, while *Virtual Chassis* does not.
- A Layer 3 routed multicast setup—For information about configuring multicasting, see *Junos OS Routing Protocols Configuration Guide*.
- Aggregated 10-gigabit links in a LAG—For information about configuring LAGs with multicast load balancing , see Configuring Multicast Load Balancing for Use with Aggregated 10-Gigabit Ethernet Links on EX8200 Switches (CLI Procedure).

How Does Multicast Load Balancing Work?

When traffic can use multiple member links, traffic that is part of the same stream must always be on the same link.

Multicast load balancing uses one of seven available hashing algorithms and a technique called queue shuffling (alternating between two queues) to distribute and balance the data, directing streams over all available aggregated links. You can select one of the seven algorithms when you configure multicast load balancing, or you can use the default algorithm, crc-sgip, which uses a cyclic redundancy check (CRC) algorithm on the multicast packets' group IP address. We recommend that you start with the crc-sgip default and try other options if this algorithm does not evenly distribute the Layer 3 routed multicast traffic. Six of the algorithms are based on the hashed value of IP addresses (IPv4 or IPv6) and will produce the same result each time they are used. Only the balanced mode option produces results that vary depending on the order in which streams are added. See Table 79 on page 375 for more information.

Table 79: Hashing Algorithms Used by Multicast Load Balancing

Hashing Algorithms	Based On	Best Use
crc-sgip	Cyclic redundancy check of multicast packets' source and group IP address	Default—high-performance management of IP traffic on 10-Gigabit Ethernet network. Predictable assignment to the same link each time. This mode is complex but yields a good distributed hash.

Table 79: Hashing Algorithms Used by Multicast Load Balancing *(Continued)*

Hashing Algorithms	Based On	Best Use
crc-gip	Cyclic redundancy check of multicast packets' group IP address	Predictable assignment to the same link each time. Try this mode when crc-sgip does not evenly distribute the Layer 3 routed multicast traffic and the group IP addresses vary.
crc-sip	Cyclic redundancy check of multicast packets' source IP address	Predictable assignment to the same link each time. Try this mode when crc-sgip does not evenly distribute the Layer 3 routed multicast traffic and the stream sources vary.
simple-sgip	XOR calculation of multicast packets' source and group IP address	Predictable assignment to the same link each time. This is a simple hashing method that might not yield as even a distribution as crc-sgip yields. Try this mode when crc-sgip does not evenly distribute the Layer 3 routed multicast traffic.
simple-gip	XOR calculation of multicast packets' group IP address	Predictable assignment to the same link each time. This is a simple hashing method that might not yield as even a distribution as crc-gip yields. Try this when crc-gip does not evenly distribute the Layer 3 routed multicast traffic and the group IP addresses vary.
simple-sip	XOR calculation of multicast packets' source IP address	Predictable assignment to the same link each time. This is a simple hashing method that might not yield as even a distribution as crc-sip yields. Try this mode when crc-sip does not evenly distribute the Layer 3 routed multicast traffic and stream sources vary.
balanced	Round-robin calculation method used to identify multicast links with the least amount of traffic	Best balance is achieved, but you cannot predict which link will be consistently used because that depends on the order in which streams come online. Use when consistent assignment is not needed after every reboot.

How Do I Implement Multicast Load Balancing on an EX8200 Switch?

To implement multicast load balancing with an optimized level of throughput on an EX8200 switch, follow these recommendations:

- Allow 25 percent unused bandwidth in the aggregated link to accommodate any dynamic imbalances due to link changes caused by sharing multicast interfaces.
- For downstream links, use multicast interfaces of the same size whenever possible. Also, for downstream aggregated links, throughput is optimized when members of the aggregated link belong to the same devices.
- For upstream aggregated links, use a Layer 3 link whenever possible. Also, for upstream aggregated links, throughput is optimized when the members of the aggregated link belong to different devices.

SEE ALSO

Configuring Multicast Load Balancing for Use with Aggregated 10-Gigabit Ethernet Links on EX8200 Switches (CLI Procedure)

Example: Configuring Multicast Load Balancing for Use with Aggregated 10-Gigabit Ethernet Interfaces on EX8200 Switches

IN THIS SECTION

- Requirements | 378
- Overview and Topology | 378
- Configuration | 380
- Verification | 383

EX8200 switches support multicast load balancing on link aggregation groups (LAGs). Multicast load balancing evenly distributes Layer 3 routed multicast traffic over the LAGs, You can aggregate up to twelve 10-gigabit Ethernet links to form a 120-gigabit virtual link or LAG. The MAC client can treat this virtual link as if it were a single link to increase bandwidth, provide graceful degradation as link failures occur, and increase availability. On EX8200 switches, multicast load balancing is enabled by default. However, if it is explicitly disabled, you can reenable it. .



NOTE: An interface with an already configured IP address cannot form part of the LAG.

NOTE: Only EX8200 standalone switches with 10-gigabit links support multicast load balancing. Virtual Chassis does not support multicast load balancing.

This example shows how to configure a LAG and reenable multicast load balancing:

Requirements

This example uses the following hardware and software components:

- Two EX8200 switches, one used as the access switch and one used as the distribution switch
- Junos OS Release 12.2 or later for EX Series switches

Before you begin:

- Configure four 10-gigabit interfaces on the EX8200 distribution switch: xe-0/1/0, xe-1/1/0, xe-2/1/0, and xe-3/1/0. See Configuring Gigabit Ethernet Interfaces (CLI Procedure).

Overview and Topology

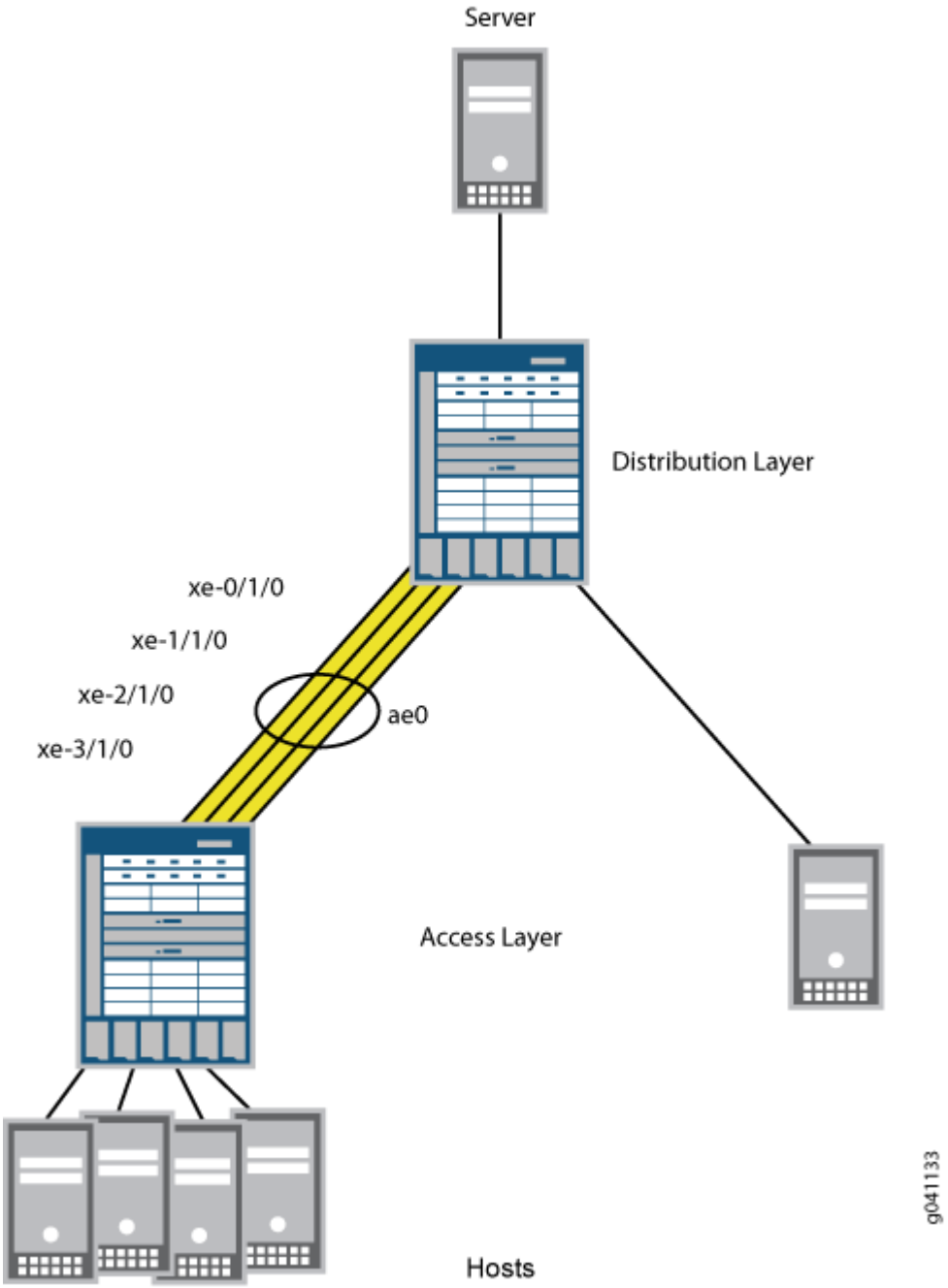
Multicast load balancing uses one of seven hashing algorithms to balance traffic between the individual 10-gigabit links in the LAG. For a description of the hashing algorithms, see *multicast-loadbalance*. The default hashing algorithm is crc-sgip. You can experiment with the different hashing algorithms until you determine the one that best balances your Layer 3 routed multicast traffic.

When a link larger than 10 gigabits is needed on an EX8200 switch, you can combine up to twelve 10-gigabit links to create more bandwidth. This example uses the link aggregation feature to combine four 10-gigabit links into a 40-gigabit link on the distribution switch. In addition, multicast load balancing is enabled to ensure even distribution of Layer 3 routed multicast traffic on the 40-gigabit link. In the sample topology illustrated in Figure 17 on page 379, an EX8200 switch in the distribution layer is connected to an EX8200 switch in the access layer.

NOTE: Link speed is automatically determined based on the size of the LAG configured. For example, if a LAG is composed of four 10-gigabit links, the link speed is 40 gigabits per second).

NOTE: The default hashing algorithm, crc-sgip, involves a cyclic redundancy check of both the multicast packet source and group IP addresses.

Figure 17: 40-Gigabit LAG Composed of Four 10-Gigabit Links



You will configure a LAG on each switch and reenable multicast load balancing. When reenabled, multicast load balancing will automatically take effect on the LAG, and the speed is set to 10 gigabits per second for each link in the LAG. Link speed for the 40-gigabit LAG is automatically set to 40 gigabits per second.

Configuration

IN THIS SECTION

- Procedure | 380

Procedure

CLI Quick Configuration

To quickly configure this example, copy the following commands, paste them into a text file, remove any line breaks, change any details necessary to match your network configuration, and then copy and paste the commands into the CLI at the `[edit]` hierarchy level.

```
set chassis aggregated-devices ethernet device-count 1
set interfaces ae0 aggregated-ether-options minimum-links 1
set interfaces xe-0/1/0 ether-options 802.3ad ae0
set interfaces xe-1/1/0 ether-options 802.3ad ae0
set interfaces xe-2/1/0 ether-options 802.3ad ae0
set interfaces xe-3/1/0 ether-options 802.3ad ae0
set chassis multicast-loadbalance hash-mode crc-gip
```

Step-by-Step Procedure

To configure a LAG and reenable multicast load balancing:

1. Specify the number of aggregated Ethernet interfaces to be created:

```
[edit chassis]
user@switch#  set aggregated-devices ethernet device-count 1
```

2. Specify the minimum number of links for the aggregated Ethernet interface (ae*x*), that is, the LAG, to be labeled `up`:



NOTE: By default, only one link needs to be up for the LAG to be labeled `up`.

```
[edit interfaces]
user@switch# set ae0 aggregated-ether-options minimum-links 1
```

3. Specify the four members to be included within the LAG:

```
[edit interfaces]
user@switch# set xe-0/1/0 ether-options 802.3ad ae0
user@switch# set xe-1/1/0 ether-options 802.3ad ae0
user@switch# set xe-2/1/0 ether-options 802.3ad ae0
user@switch# set xe-3/1/0 ether-options 802.3ad ae0
```

4. Reenable multicast load balancing:

```
[edit chassis]
user@switch#  set  multicast-loadbalance
```



NOTE: You do not need to set link speed the way you do for LAGs that do not use multicast load balancing. Link speed is automatically set to 40 gigabits per second on a 40-gigabit LAG.

5. You can optionally change the value of the `hash-mode` option in the **multicast-loadbalance** statement to try different algorithms until you find the one that best distributes your Layer 3 routed multicast traffic.

   If you change the hashing algorithm when multicast load balancing is disabled, the new algorithm takes effect after you reenable multicast load balancing.

Results

Check the results of the configuration:

```
user@switch>  show configuration
chassis
```

```
    aggregated-devices {
        ethernet {
             device-count 1;
        }
    }
    multicast-loadbalance {
        hash-mode crc-gip;
    }

interfaces
    xe-0/1/0 {
    ether-options {
            802.3ad ae0;
        }
    }
    xe-1/1/0 {
        ether-options {
            802.3ad ae0;
        }
    }
    xe-2/1/0 {
        ether-options {
            802.3ad ae0;
        }
    }
    xe-3/1/0 {
        ether-options {
            802.3ad ae0;
        }
    }
    ae0 {
        aggregated-ether-options {
            minimum-links 1;
        }
    }
}
```

Verification

IN THIS SECTION

- Verifying the Status of a LAG Interface | 383
- Verifying Multicast Load Balancing | 384

To confirm that the configuration is working properly, perform these tasks:

Verifying the Status of a LAG Interface

Purpose

Verify that a link aggregation group (LAG) (**ae0**) has been created on the switch.

Action

Verify that the **ae0** LAG has been created:

```
user@switch> show interfaces ae0 terse
```

```
Interface               Admin      Link     Proto      Local            Remote

ae0                      up       up

ae0.0                    up       up     inet     10.10.10.2/24
```

Meaning

The interface name ae*x* indicates that this is a LAG. *A* stands for aggregated, and *E* stands for Ethernet. The number differentiates the various LAGs.

Verifying Multicast Load Balancing

Purpose

Check that traffic is load-balanced equally across paths.

Action

Verify load balancing across the four interfaces:

```
user@switch> monitor interface traffic
```

```
Bytes=b, Clear=c, Delta=d, Packets=p, Quit=q or ESC, Rate=r, Up=^U, Down=^D
ibmoem02-re1                      Seconds: 3                  Time: 16:06:14

Interface    Link  Input packets        (pps)     Output packets        (pps)
 xe-0/1/0      Up        2058834         (10)          7345862         (19)
 xe-1/1/0      Up        2509289          (9)          6740592         (21)
 xe-2/1/0      Up        8625688         (90)         10558315         (20)
 xe-3/1/0      Up        2374154         (23)         71494375          (9)
```

Meaning

The interfaces should be carrying approximately the same amount of traffic.

SEE ALSO

> Configuring Multicast Load Balancing for Use with Aggregated 10-Gigabit Ethernet Links on EX8200 Switches (CLI Procedure)

Change History Table

Feature support is determined by the platform and release you are using. Use Feature Explorer to determine if a feature is supported on your platform.

Release	Description
10.1	Starting with Junos OS Release 10.1, you can also configure the load balancing hash key for Layer 2 traffic to use fields in the Layer 3 and Layer 4 headers using the `payload` statement.

Dynamic Load Balancing (DLB)

SUMMARY

Dynamic load balancing (DLB) is an improvement on static load balancing (SLB).

IN THIS SECTION

- Dynamic Load Balancing Overview | **385**
- Configuring Dynamic Load Balancing | **388**
- Example: Configure Dynamic Load Balancing | **390**
- Selective Dynamic Load Balancing (DLB) | **398**
- Customize Egress Port Link Quality Metrics for DLB | **404**
- Configure Flowset Table Size in DLB Flowlet Mode | **406**
- Reactive Path Rebalancing | **408**

Dynamic Load Balancing Overview

IN THIS SECTION

- Benefits | **387**
- DLB Modes | **387**
- Platform-Specific DLB Behavior | **388**

Load balancing is used to ensure that network traffic is distributed as evenly as possible across members in a given equal-cost multipath (ECMP) routing group or link aggregation group (LAG). In general, load balancing is classified as either static or dynamic. Static load balancing (SLB) computes hashing solely based on the packet contents (for example, source IP, destination IP, and so on). The biggest advantage of SLB is that packet ordering is guaranteed as all packets of a given flow take the same path. However,

because the SLB mechanism does not consider the path or link load, the network often experiences the following problems:

- Poor link bandwidth utilization
- Elephant flow on a single link completely dropping mice flows on it.

Dynamic load balancing (DLB) is an improvement on top of SLB.

For ECMP, you can configure DLB globally, whereas for LAG, you configure it for each aggregated Ethernet interface. You can apply DLB on selected *ether-type (Dynamic Load Balancing)* (IPv4, IPv6, and MPLS) based on configuration. If you don't configure any *ether-type (Dynamic Load Balancing)*, then DLB is applied to all EtherTypes. Note that you must explicitly configure the DLB mode because there is no default mode.

Use Feature Explorer to confirm platform and release support for specific features.



NOTE:

- You cannot configure both DLB and resilient hashing at the same time. Otherwise, a commit error will be thrown.
- DLB is applicable only for unicast traffic.
- DLB is not supported when the LAG is one of the egress ECMP members.
- DLB is not supported for remote LAG members.
- DLB is not supported on Virtual Chassis and Virtual Chassis Fabric (VCF).
- DLB on LAG and HiGig-trunk are not supported at the same time.

Here are some of the important behaviors of DLB:

- DLB is applicable for incoming EtherTypes only.
- From a DLB perspective, both Layer 2 and Layer 3 link aggregation group (LAG) bundles are considered the same.
- The link utilisation will not be optimal if you use dynamic load balancing in asymmetric bundles—that is, on ECMP links with different member capacities.
- With DLB, no reassignment of flow happens when a new link is added in per packet and assigned flow modes. This can cause suboptimal usage in link flap scenarios where a utilized link may not be utilized after it undergoes a flap if no new flow or flowlets are seen after the flap.

Benefits

- DLB considers member bandwidth utilization along with packet content for member selection. As a result, we achieve better link utilization based on real-time link loads.
- DLB ensures that links hogged by elephant flows are not used by mice flows. Thus, by using DLB, we avoid hash collision drops that occur with SLB. That is, with DLB the links are spread across, and thus the collision and the consequent drop of packets are avoided.

DLB Modes

You can use the following DLB modes to load-balance traffic:

- *Per packet mode*

  In this mode, DLB is initiated for each packet in the flow. This mode makes sure that the packet always gets assigned to the best-quality member port. However, in this mode, DLB may experience packet reordering problems that can arise due to latency skews.

- *Flowlet mode*

  This mode relies on assigning links based on *flowlets* instead of flows. Real-world application traffic relies on flow control mechanisms of upper-layer transport protocols such as TCP, which throttle the transmission rate. As a result, flowlets are created. You can consider flowlets as multiple bursts of the same flow separated by a period of inactivity between these bursts—this period of inactivity is referred to as the inactivity interval. The inactivity interval serves as the demarcation criteria for identifying new flowlets and is offered as a user-configurable statement under the DLB configuration. In this mode, DLB is initiated per flowlet—that is, for the new flow as well as for the existing flow that has been inactive for a sufficiently long period of time (configured `inactivity-interval`). The reordering problem of per packet mode is addressed in this mode as all the packets in a flowlet take the same link. If the `inactivity-interval` value is configured to be higher than the maximum latency skew across all ECMP paths, then you can avoid packet reordering across flowlets while increasing link utilization of all available ECMP links.

- *Assigned flow mode*

  You can use assigned flow mode to selectively disable rebalancing for a period of time to isolate problem sources. You cannot use this mode for real-time DLB or predict the egress ports that will be selected using this mode because assigned flow mode does not consider port load and queue size.

Platform-Specific DLB Behavior

NOTE: QFX5220, QFX5230-64CD, and QFX5240 switches do not support DLB on LAG.

Table 80: Platforms That Support Dynamic Load Balancing for ECMP/LAG

Platform	DLB Support for ECMP	DLB Support for LAG
QFX5120-32C	Yes	Yes
QFX5120-48Y	Yes	Yes
QFX5220	Yes	No
QFX5230-64CD	Yes	No
QFX5240	Yes	No

Configuring Dynamic Load Balancing

This topic describes how to configure dynamic load balancing (DLB) in flowlet mode.

Starting in Junos OS Release 19.4R1, QFX5120-32C and QFX5120-48Y switches support dynamic load balancing for both ECMP and LAG. For LAG, DLB must be configured on per aggregated ethernet interface basis.

Starting in Junos OS evolved Release 19.4R2, QFX5220 switches support dynamic load balancing (DLB) for ECMP. For ECMP, DLB must be configured globally.

Configuring DLB for ECMP (Flowlet mode)

To configure dynamic load balancing for ECMP with flowlet mode (QFX5120-32C, QFX5120-48Y, and QFX5220 switches):

1. Enable dynamic load balancing with flowlet mode:

```
[edit forwarding-options enhanced-hash-key]
user@router# set ecmp-dlb flowlet
```

2. (Optional) Configure the *inactivity-interval* value - minimum inactivity interval (in micro seconds) for link re-assignment:

```
[edit forwarding-options enhanced-hash-key]
user@router# set ecmp-dlb flowlet inactivity-interval (micro seconds)
```

3. (Optional) Configure dynamic load balancing with `ether-type`:

```
[edit forwarding-options enhanced-hash-key]
user@router# set ecmp-dlb ether-type mpls
```

4. (Optional) You can view the options configured for dynamic load balancing on ECMP using `show forwarding-options enhanced-hash-key` command.

Similarly, you can configure DLB for ECMP with *Per packet* or *Assigned flow* mode.

Configuring DLB for LAG (Flowlet mode)

Before you begin, create an aggregated ethernet (AE) bundle by configuring a set of router interfaces as aggregated Ethernet and with a specific aggregated ethernet (AE) group identifier.

To configure dynamic load balancing for LAG with flowlet mode (QFX5120-32C and QFX5120-48Y):

1. Enable dynamic load balancing with flowlet mode:

```
[edit interfaces ae-x aggregated-ether-options]
user@router# set dlb flowlet
```

2. (Optional) Configure the *inactivity-interval* value - minimum inactivity interval (in micro seconds) for link re-assignment:

```
[edit interfaces ae-x aggregated-ether-options]
user@router# set dlb flowlet inactivity-interval (micro seconds)
```

3. (Optional) Configure dynamic load balancing with `ether-type`:

```
[edit forwarding-options enhanced-hash-key]
user@router# set lag-dlb ether-type mpls
```

4. (Optional) You can view the options configured for dynamic load balancing on LAG using `show forwarding-options enhanced-hash-key` command.

Similarly, you can configure DLB for LAG with *Per packet* or *Assigned flow* mode.

SEE ALSO

> *dlb*
>
> *show forwarding-options enhanced-hash-key*

Example: Configure Dynamic Load Balancing

IN THIS SECTION

- Requirements | 390
- Overview | 391
- Configuration | 392
- Verification | 396

This example shows how to configure dynamic load balancing.

Requirements

This example uses the following hardware and software components:

- Two QFX5120-32C or QFX5120-48Y switches
- Junos OS Release 19.4R1 or later running on all devices

Overview

IN THIS SECTION

- Topology | 391

Dynamic load balancing (DLB) is an improvement on top of SLB.

For ECMP, you can configure DLB globally, whereas for LAG, you configure it for each aggregated Ethernet interface. You can apply DLB on selected *ether-type (Dynamic Load Balancing)* such as IPv4, IPv6, and MPLS based on configuration. If you don't configure any *ether-type (Dynamic Load Balancing)*, then DLB is applied to all EtherTypes. Note that you must explicitly configure the DLB mode because there is no default mode.

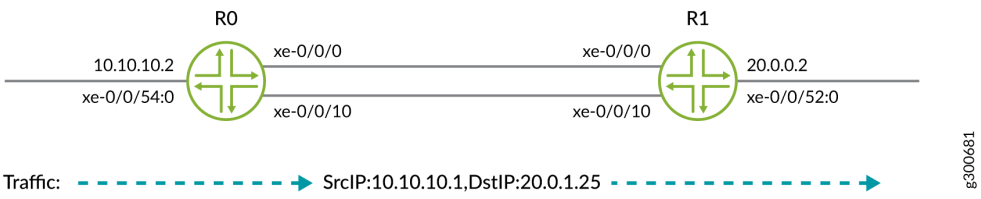
NOTE:

- Starting in Junos OS Release 19.4R1, QFX5120-32C and QFX5120-48Y switches support dynamic load balancing on both ECMP and LAG.
- You cannot configure both DLB and Resilient Hashing at the same time. Otherwise, commit error will be thrown.

Topology

In this topology, both R0 and R1 are connected.

Figure 18: Dynamic Load Balancing



NOTE: This example shows static configuration. You can also add configuration with dynamic protocols.

Configuration

IN THIS SECTION

- Verification | 395

CLI Quick Configuration

To quickly configure this example, copy the following commands, paste them into a text file, remove any line breaks, change any details necessary to match your network configuration, and then copy and paste the commands into the CLI at the `[edit]` hierarchy level.

R0

```
set interfaces xe-0/0/0 unit 0 family inet address 10.1.0.2/24
set interfaces xe-0/0/10 unit 0 family inet address 10.1.1.2/24
set interfaces xe-0/0/54:0 unit 0 family inet address 10.10.10.2/24
set forwarding-options enhanced-hash-key ecmp-dlb per-packet
set policy-options policy-statement loadbal then load-balance per-packet
set routing-options static route 20.0.1.0/24 next-hop 10.1.0.3
set routing-options static route 20.0.1.0/24 next-hop 10.1.1.3
set routing-options forwarding-table export loadbal
```

R1

```
set interfaces xe-0/0/0 unit 0 family inet address 10.1.0.3/24
set interfaces xe-0/0/10 unit 0 family inet address 10.1.1.3/24
set interfaces xe-0/0/52:0 unit 0 family inet address 20.0.0.2/16
```

Configure Dynamic Load Balancing for LAG (QFX5120-32C and QFX5120-48Y)

Step-by-Step Procedure

The following example requires you to navigate various levels in the configuration hierarchy. For information about navigating the CLI, see *Using the CLI Editor in Configuration Mode*.

To configure the R0 router:



NOTE: Repeat this procedure for the other routers, after modifying the appropriate interface names, addresses, and any other parameters for each router.

1. Configure Link Aggregation Group (LAG).

```
[edit interfaces]
user@R0# set interfaces xe-0/0/0 ether-options 802.3ad ae0
user@R0# set interfaces xe-0/0/10 ether-options 802.3ad ae0
user@R0# set interfaces ae0 aggregated-ether-options lacp active
user@R0# set interfaces ae0 unit 0 family inet address 10.1.0.2/24
user@R0# set routing-options static route 20.0.1.0/24 next-hop 10.1.0.3
```

   After configuring LAG, in the verification section, execute the steps in the *Verifying Traffic Load before configuring Dynamic Load Balancing Feature on LAG* section, to check the configuration or the traffic load before configuring DLB.

2. Configure Dynamic Load Balancing with per-packet mode for LAG.

```
[edit]
user@R0# set interfaces ae0 aggregated-ether-options dlb per-packet
```

   After configuring the DLB, in the verification section, execute the steps in the *Verifying Traffic Load after configuring Dynamic Load Balancing Feature on LAG* section, to check the configuration or the traffic load before configuring DLB.

Configure Dynamic Load Balancing for ECMP (QFX5120-32C, QFX5120-48Y, and QFX5220 switches)

Step-by-Step Procedure

The following example requires you to navigate various levels in the configuration hierarchy. For information about navigating the CLI, see *Using the CLI Editor in Configuration Mode*.

To configure the R0 router:



NOTE: Repeat this procedure for the other routers, after modifying the appropriate interface names, addresses, and any other parameters for each router.

1. Configure the Gigabit Ethernet interface link connecting from R0 to R1.

```
[edit interfaces]
user@R0# set interfaces xe-0/0/0 unit 0 family inet address 10.1.0.2/24
user@R0# set interfaces xe-0/0/10 unit 0 family inet address 10.1.1.2/24
user@R0# set interfaces xe-0/0/54:0 unit 0 family inet address 10.10.10.2/24
```

2. Create the static routes:

```
[edit interfaces]
user@R0# set routing-options static route 20.0.1.0/24 next-hop 10.1.0.3
user@R0# set routing-options static route 20.0.1.0/24 next-hop 10.1.1.3
```

3. Apply the load-balancing policy. The dynamic load balancing feature requires the multiple ECMP next hops to be present in the forwarding table.

```
[edit interfaces]
user@R0# set policy-options policy-statement loadbal then load-balance per-packet
user@R0# set routing-options forwarding-table export loadbal
```

4. Configure Dynamic Load Balancing with per-packet mode for ECMP.

```
[edit interfaces]
user@R0# set forwarding-options enhanced-hash-key ecmp-dlb per-packet
```

5. On R1, configure the Gigabit Ethernet interface link.

```
[edit interfaces]
user@R2# set interfaces xe-0/0/0 unit 0 family inet address 10.1.0.3/24
user@R2# set interfaces xe-0/0/10 unit 0 family inet address 10.1.1.3/24
user@R2# set interfaces xe-0/0/52:0 unit 0 family inet address 20.0.0.2/16
```

Verification

IN THIS SECTION

- Verify Traffic Load Before Configuring Dynamic Load Balancing Feature on LAG | 395
- Verify Traffic Load After Configuring Dynamic Load Balancing Feature on LAG | 395

Confirm that the configuration is working properly.

Verify Traffic Load Before Configuring Dynamic Load Balancing Feature on LAG

Purpose

Verify before the DLB feature is configured on the Link Aggregation Group.

Action

From operational mode, run the `show interfaces interface-name | match pps` command.

```
user@R0>show interfaces xe-0/0/0 | match pps
  Input rate     : 1240 bps (1 pps)
  Output rate    : 1024616 bps (1000 pps) ## all traffic in one link.
user@R0>show interfaces xe-0/0/10 | match pps
  Input rate     : 616 bps (0 pps)
  Output rate    : 1240 bps (1 pps)<<  Output rate    : 1240 bps (1 pps) ## no traffic
```

Verify Traffic Load After Configuring Dynamic Load Balancing Feature on LAG

Purpose

Verify that packets received on the R0 are load-balanced.

Action

From operational mode, run the `show interfaces` *`interface-name`* command.

```
user@R0>show interfaces xe-0/0/0 | match pps
  Input rate     : 616 bps (0 pps)
  Output rate    : 519096 bps (506 pps)<<  Output rate    : 519096 bps (506 pps) ## load equally
shared
user@R0>show interfaces xe-0/0/10 | match pps
  Input rate     : 1232 bps (1 pps)
  Output rate    : 512616 bps (500 pps)<<  Output rate    : 512616 bps (500 pps) ## load equally
shared
```

Meaning

Dynamic Load balancing with per-packet mode successfully working. After applying dynamic load balancing feature on LAG, the load is equally shared in the network.

Verification

IN THIS SECTION

- Verify Dynamic Load Balancing on R0 | **396**
- Verify Load Balancing on R1 | **397**

Confirm that the configuration is working properly at R0.

Verify Dynamic Load Balancing on R0

Purpose

Verify that packets received on the R0 are load-balanced.

Action

From operational mode, run the `run show route forwarding-table destination` *destination-address* command.

```
user@R0>show route forwarding-table destination 20.0.1.0/24
inet.0: 178 destinations, 178 routes (178 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both

20.0.1.0/24        *[Static/5] 1d 03:35:12
                    >  to 10.1.0.3 via xe-0/0/0.0
                       to 10.1.1.3 via xe-0/0/10.0
user@R0>show route 20.0.1.0/24
inet.0: 178 destinations, 178 routes (178 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both

20.0.1.0/24        *[Static/5] 1d 03:35:12
                    >  to 10.1.0.3 via xe-0/0/0.0
                       to 10.1.1.3 via xe-0/0/10.0
```

Meaning

The packets received on the R0 are load-balanced.

Verify Load Balancing on R1

Purpose

Confirm that the configuration is working properly at R1.

Action

From operational mode, run the `show route` command.

```
user@R1>show route 20.0.1.25
inet.0: 146 destinations, 146 routes (146 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both

20.0.0.0/16        *[Direct/0] 1d 03:37:11
                    >  via xe-0/0/52:0.0
```

Meaning

Dynamic Load balancing with per-packet mode successfully working. After applying dynamic load balancing feature on ECMP, the load is equally shared in the network.

SEE ALSO

dlb

show forwarding-options enhanced-hash-key

Selective Dynamic Load Balancing (DLB)

IN THIS SECTION

- Selective DLB Overview | **398**
- Selective DLB in AI-ML Data Centers | **399**
- Configuration | **399**
- Example: Selectively Enable DLB with a Firewall Filter Match Condition | **403**

Selective DLB Overview

IN THIS SECTION

- Benefits | **399**

With *selective DLB*, you no longer have to choose between DLB and SLB for all traffic traversing your device. You can configure your preferred DLB mode at the global level, configure a default type of load balancing, and then selectively enable or disable DLB for certain kinds of traffic.

Selective DLB is also useful when very large data flow, also called an elephant flow, encounters links that are too small for the entire data flow. In this scenario, selective DLB can calculate the optimal use of the links' available bandwidth in the data center fabric. When you enable selective per-packet DLB for

the elephant flow, the algorithm directs the packets to the best-quality link first. As the link quality changes, the algorithm directs subsequent packets to the next best-quality link.

Use Feature Explorer to confirm platform and release support for specific features.

Benefits

- Improve your network handling of large data flows.
- Use per-packet and per-flow load balancing in the same traffic stream to improve performance.
- Customize load balancing based on any firewall filter match condition.

Selective DLB in AI-ML Data Centers

In AI-ML workloads, the majority of the application traffic uses Remote Direct Memory Access (RDMA) over Converged Ethernet version 2 (RoCEv2) for transport. Dynamic load balancing (DLB) is ideal for achieving efficient load balancing and preventing congestion in RoCEv2 networks. However, static load balancing (SLB) can be more effective for some types of traffic. Selective DLB solves this problem.

You can enable load balancing in two ways: per flow or per packet. Per-flow load balancing has been the most widely used because it handles the largest number of packets at a time. The device classifies packets that have the same 5-tuple packet headers as a single flow. The device gives all packets in the flow the same load balancing treatment. Flow-based load balancing works well for general TCP and UDP traffic because the traffic utilizes all links fairly equally. However, per-packet load balancing can reorder some packets, which can impact performance.

Many AI clusters connect the application to the network through smart network interface cards (SmartNICs) that can handle out-of-order packets. To improve performance, enable per-packet DLB on your network. Then enable DLB for only those endpoint servers that are capable of handling out-of-order packets. Your device looks at the RDMA operation codes (opcodes) in the BTH+ headers of these packets in real time. Using any firewall filter match condition, you can selectively enable or disable DLB based on these opcodes. Other flows continue to use default hash-based load balancing, also known as SLB.

Configuration

IN THIS SECTION

- Configuration Overview | **400**
- Topology | **400**
- Disable DLB Globally and Selectively Enable DLB | **401**

- Enable DLB Globally and Selectively Disable DLB | **402**

Configuration Overview

You can selectively enable DLB in two ways: disable DLB by default and selectively enable DLB on certain flows, or enable DLB globally and selectively disable DLB. In either case, you'll need to first configure DLB in *per-packet mode*. Per-packet is the DLB mode used wherever DLB is enabled. You cannot configure DLB in per-flow and per-packet mode on the same device at the same time.

This feature is compatible with flowlet mode. You can optionally enable this feature when DLB is configured in flowlet mode.

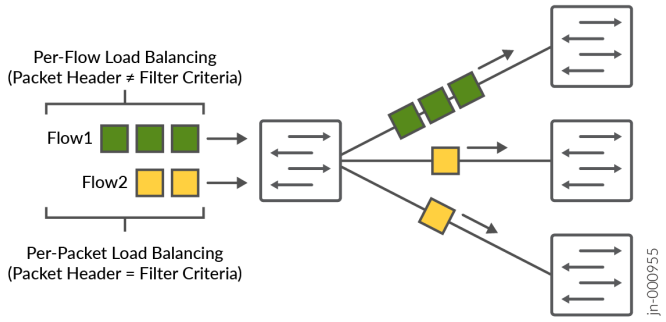
Topology

In the topology shown in Figure 2, DLB is disabled by default. We have enabled DLB selectively on Flow2 in per-packet mode. Table 2 summarizes the load balancing configuration on the two flows shown and the results of the load balancing applied on the flows:

Table 81: Flow Behaviors

Flow	DLB Enabled?	Result
Flow1	No	The device uses the default load balancing configuration, which is per-flow mode. The flow is directed to a single device.
Flow2	Yes	The device uses the DLB configuration, which is per-packet mode. The device splits this flow into packets. DLB assigns each packet to a path that is based on the RDMA opcode in the packet header and the corresponding filter.

Figure 19: Per-Flow and Per-Packet Load Balancing



Disable DLB Globally and Selectively Enable DLB

In cases where very few packets will require DLB, you can disable DLB at the global level and selectively enable it per flow.

1. Enable DLB per-packet mode. Whenever DLB is enabled on a flow, DLB uses this mode to direct traffic.

```
set system packet-forwarding-options firewall profiles <inet | inet6 | ethernet-switching>
udf-profile-name
set forwarding-options enhanced-hash-key ecmp-dlb per-packet
```

2. Disable DLB globally by turning it off for all Ethernet types. By default, all packets will get hash-based load balancing (SLB).

```
set forwarding-options enhanced-hash-key ecmp-dlb ether-type none
```

3. Configure a firewall filter to match a specific RDMA opcode within the BTH+ header.

   This example matches based on `rdma-opcode` 10.

```
set firewall family inet filter filter-name term term-name from rdma-opcode 10
```

4. Enable per-packet DLB within that firewall filter to only apply DLB to those packets with the chosen RDMA opcode in the BTH+ header.

```
set firewall family inet filter filter-name term term-name then dynamic-load-balance enable
```

5. Other packets get the default load balancing method, which is SLB.

```
set firewall family inet filter filter-name term default then accept
```

Enable DLB Globally and Selectively Disable DLB

In cases where most packets will benefit from DLB, enable DLB at the global level for all packets and selectively disable it per packet.

1. Configure DLB at the global level in per-packet mode for all flows.

```
set system packet-forwarding-options firewall profiles <inet | inet6 | ethernet-switching>
udf-profile-name
set forwarding-options enhanced-hash-key ecmp-dlb per-packet
```

2. Configure a firewall filter to match a specific RDMA opcode within the BTH+ header.

   This example matches based on `rdma-opcode` 10.

```
set firewall family inet filter filter-name term term-name from rdma-opcode 10
```

3. Disable per-packet DLB within that firewall filter for packets with the chosen RDMA opcode in the BTH+ header.

```
set firewall family inet filter filter-name term term-name then dynamic-load-balance disable
```

4. Other packets get the default load balancing method, which is DLB.

```
set firewall family inet filter filter-name term default then accept
```

5. Verify DLB is enabled as you expected using the following commands:

```
show forwarding-options enhanced-hash-key
```

```
show pfe filter hw profile-info
```

Example: Selectively Enable DLB with a Firewall Filter Match Condition

One of the benefits of selective DLB is that you can customize load balancing based on any firewall filter match condition. This example shows how to enable DLB based on a firewall filter that matches with RDMA queue pairs. Use this example to enable per-packet DLB only for those flows terminating on a network interface card (NIC) that supports packet reordering.

In a network that uses RoCEv2 for application traffic transport, an RDMA connection sends traffic on a send queue and receives traffic on a receive queue. These queues form the RDMA connection. Together, the send queue and receive queue are referred to as a queue pair. Each queue pair has an identifiable prefix. In this example, we use queue pair prefixes to control when DLB is enabled.

This example is configured on a QFX5240-64QD switch.

1. Create a user-defined field in a firewall for matching packets that is destined for a specific RDMA destination queue pair. Select a queue pair you know terminates on an NIC that is capable of reordering packets.

   We named our firewall filter `sDLB`. The term `QP-match` matches on incoming packets with a destination queue pair with the following characteristics.

   ```
   set firewall family inet filter sDLB term QP-match from flexible-match-range match-start
   layer-4
   set firewall family inet filter sDLB term QP-match from flexible-match-range byte-offset 13
   set firewall family inet filter sDLB term QP-match from flexible-match-range bit-length 24
   set firewall family inet filter sDLB term QP-match from flexible-match-range range 0x64
   ```

2. Configure the firewall filter to enable per-packet DLB on the queue pairs that match the filter.

   If the queue pair is not a match, the device uses the default load balancing type of SLB for that packet.

   ```
   set firewall family inet filter sDLB term QP-match then dynamic-load-balance enable
   ```

3. Configure a counter that increments each time there is a match.

   The counter `QP-match-count` tracks how many packets were load balanced with DLB. You can use this information when troubleshooting.

   ```
   set firewall family inet filter sDLB term QP-match then count QP-match-count
   ```

4. Enable your firewall filter on the relevant interface.

   ```
   set interfaces et-0/0/5 unit 0 family inet filter input sDLB
   ```

5. Verify your firewall filter term is matching on packets coming through the device.

   The `QP-match-count` counter shows the number of bytes and packets that the firewall filter has redirected for load balancing with DLB.

```
user@device> show firewall

Filter: sDLB
Counters:
Name                                Bytes              Packets
QP-match-count                 176695488320           552173401
```

Customize Egress Port Link Quality Metrics for DLB

IN THIS SECTION

- Overview | 404
- Configuration | 405

Overview

IN THIS SECTION

- Benefits | 405

Dynamic load balancing (DLB) selects an optimal link based on the quality of the link so that traffic flows are evenly distributed across your network. You (the network administrator) can customize the way DLB assigns quality metrics of egress ports so that DLB selects the optimal link.

DLB assigns each egress port that is part of equal-cost multipath (ECMP) to a quality band. Quality bands are numbered from 0 through 7, where 0 is the lowest quality and 7 is the highest quality. DLB tracks two metrics on each of the ports, and it uses these metrics to compute the link quality:

- Port load metric: The amount of traffic recently transmitted over each ECMP link, measured in bytes.

- Port queue metric: The amount of traffic enqueued on each ECMP link for transmission, measured in number of cells.

Based on the member port load and queue size, DLB assigns one of the quality bands to the member port. The port-to-quality band mapping changes based on the instantaneous port load and queue size metrics.

By default, DLB weighs the port load metric and port queue metric equally when evaluating link quality. You can configure DLB to base the link quality more heavily on the port load than the port queue, or vice versa. Configure the amount of weight DLB places on the port load using the `rate-weightage` statement at the `[edit forwarding-options enhanced-hash-key ecmp-dlb egress-quantization]` hierarchy level. DLB assigns the remaining weight percentage to the port queue. For example, if you configure the `rate-weightage` value to be 80, DLB places 80% weight on the port load and 20% weight on the port queue when evaluating the quality of a link.

You can also configure port load thresholds that determine the upper and lower quality bands. The thresholds are percentages of the total port load that you configure using the `min` and `max` options. DLB assigns any egress port with a port load falling below this minimum to the highest quality band (7). Any port load larger than the maximum threshold falls into the lowest quality band (0). DLB divides the remaining port load quantities among quality bands 1 through 6.

For example, if you configure the minimum to be 10 and the maximum to be 70, DLB assigns any egress port with a port load that takes up less than 10 percent (%) of the total port load to quality band 7. DLB assigns any egress port with a port load taking up more than 70% of the total port load to quality band 0. DLB then assigns egress ports with port loads taking up 10% through 70% of the total port load to quality bands 1 through 6.

Use Feature Explorer to confirm platform and release support for specific features.

Benefits

- Optimize load balancing based on port activity that is determined by both port load size and queues.
- Configure link quality parameters that best suit your network needs.
- Allow DLB to flexibly assign ports to quality bands based on real-time metrics.

Configuration

Configure the egress port quality metric.

1. Configure how much weight DLB puts on the port load metric, or amount of traffic, when determining the link quality.

   Range of `rate-weightage`: 0 through 100, where 100 means that DLB bases link quality 100% on the port load.

When the rate weightage changes, the device repairs all ECMP DLB groups with the new egress quantization values for each of their egress links. During the transition between configurations, traffic can drop.

```
set forwarding-options enhanced-hash-key ecmp-dlb egress-quantization rate-weightage rate-weightage
```

2. Configure the minimum port load in percentage.

   DLB assigns any egress port with a port load falling below this minimum to the highest quality band (7). Range of `min`: 1 through 100 (percent).

```
set forwarding-options enhanced-hash-key ecmp-dlb egress-quantization min min
```

3. Configure the maximum port load in percentage.

   DLB assigns any egress port with a port load above this maximum to the lowest quality band (0). Range of `max`: 1 through 100 (percent).

```
set forwarding-options enhanced-hash-key ecmp-dlb egress-quantization max max
```

4. Verify the configuration was successful.

```
show forwarding-options enhanced-hash-key
```

Configure Flowset Table Size in DLB Flowlet Mode

IN THIS SECTION

- Overview | 407
- Configuration | 408

Overview

IN THIS SECTION

- Benefits | 407

Dynamic load balancing (DLB) is a load balancing technique that selects an optimal egress link based on link quality so that traffic flows are evenly distributed. You (the network administrator) can configure DLB in *flowlet mode*.

In flowlet mode, DLB tracks the flows by recording the last seen timestamp and the egress interface that DLB selected based on the optimal link quality. DLB records this information in the flowset table allocated to each ECMP group. The DLB algorithm maintains a given flow on a particular link until the last seen timestamp exceeds the inactivity timer. When the inactivity timer expires for a particular flow, DLB rechecks whether that link is still optimal for that flow. If the link is no longer optimal, DLB selects a new egress link and updates the flowset table with the new link and the last known timestamp of the flow. If the link continues to be optimal, the flowset table continues to use the same egress link.

You (the network administrator) can increase the flowset table size to change the distribution of the flowset table entries among the ECMP groups. The more entries an ECMP group has in the flowset table, the more flows the ECMP group can accommodate. In environments such as AI-ML data centers that must handle large numbers of flows, it is particularly useful for DLB to use a larger flowset table size. When each ECMP group can accommodate a large number of flows, DLB achieves better flow distribution across the ECMP member links.

The flowset table holds 32,768 total entries, and these entries are divided equally among the DLB ECMP groups. The flowset table size for each ECMP group ranges from 256 through 32,768. Use the following formula to calculate the number of ECMP groups:

```
32,768/(flowset size) = Number of ECMP groups
```

By default, the flowset size is 256 entries, so by default there are 128 ECMP groups.

Use Feature Explorer to confirm platform and release support for specific features.

Benefits

- Improve load distribution over egress links.
- Group flows to minimize how many calculations DLB has to make for each flow.

- Customize flowset table entry allocation for maximum efficiency.
- Increase the efficiency of flowlet mode.

Configuration

Be aware of the following when configuring the flowset table size:

- When you change the flowset size, the scale of ECMP DLB groups also changes. Allocating a flowset table size greater than 256 reduces the number of DLB-capable ECMP groups.
- When you commit this configuration, traffic can drop during the configuration change.
- DLB is not supported when a link aggregation group (LAG) is one of the egress members of ECMP.
- Only underlay fabrics support DLB.
- QFX5240 switch ports with a speed less than 50 Gbps do not support DLB.

1. Configure DLB in flowlet mode. See Configuring Dynamic Load Balancing.
2. Configure the flowset table size.

```
set forwarding-options enhanced-hash-key ecmp-dlb flowlet flowset-table-size value
```

3. Verify the configuration was successful.

```
show forwarding-options enhanced-hash-key
```

Reactive Path Rebalancing

IN THIS SECTION

- Overview | 409
- Configuration | 409

Overview

IN THIS SECTION

- Benefits | 409

Dynamic load balancing (DLB) is an important tool for handling the large data flows (also known as elephant flows) inherent in AI-ML data center fabrics. *Reactive path rebalancing* is an enhancement to existing DLB features.

In the flowlet mode of DLB, you (the network administrator) configure an inactivity interval. The traffic uses the assigned outgoing (egress) interface until the flow pauses for longer than the inactivity timer. If the outgoing link quality deteriorates gradually, the pause within the flow might not exceed the configured inactivity timer. In this case, classic flowlet mode does not reassign the traffic to a different link, so the traffic cannot utilize a better-quality link. Reactive path rebalancing addresses this limitation by enabling the user to move the traffic to a better-quality link even when flowlet mode is enabled.

The device assigns a quality band to each equal-cost multipath (ECMP) egress member link that is based on the traffic flowing through the link. The quality band depends on the port load and the queue buffer. The port load is the number of egress bytes transmitted. The queue buffer is the number of bytes waiting to be transmitted from the egress port. You can customize these attributes based on the traffic pattern flowing through the ECMP.

Use Feature Explorer to confirm platform and release support for specific features.

Benefits

- Scalable solution to link degradation
- Optimal use of bandwidth for large data flows
- Avoidance of load balancing inefficiencies due to long-lived flows

Configuration

IN THIS SECTION

- Configuration Overview | 410
- Topology | 410

Configure Reactive Path Rebalancing | 411

Configuration Overview

Quality bands are numbered from 0 through 7, where 0 is the lowest quality and 7 is the highest quality. Based on the member port load and queue size, DLB assigns a quality band value to the member port. The port-to-quality band mapping changes based on instantaneous port load and queue size.

When both of the following conditions are met, reactive path rebalancing reassigns a flow to a higher-quality member link:

- A better-quality member link is available whose quality band is equal to or greater than the current member's quality band plus the configured reassignment *quality delta* value. The quality delta is the difference between the two quality bands. Configure the quality delta value using the `quality-delta` statement.
- The packet random value that the system generates is lower than the reassignment *probability threshold* value. Configure the probability threshold value using the `prob-threshold` statement.

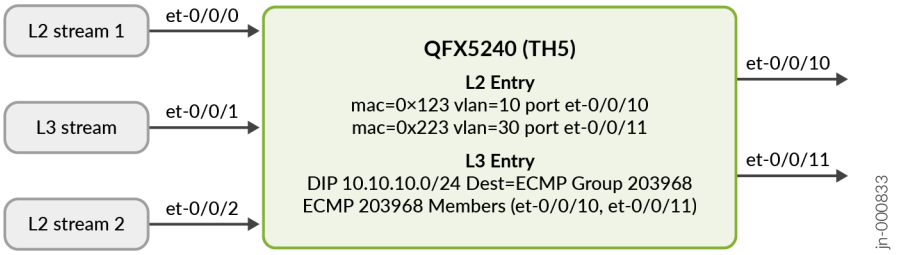
Be aware of the following when using this feature:

- Reactive path rebalancing is a global configuration and applies to all ECMP DLB configurations in the system.
- You can configure egress quantization in addition to reactive path rebalancing to control the flow reassignment.
- Packet reordering can occur when the flow moves from one port to another. Configuring reactive path rebalancing can cause momentary out-of-order issues when the flow is reassigned to the new link.

Topology

In this topology, the device has three ingress ports and two egress ports. Two of the ingress streams are Layer 2 (L2) traffic and one is Layer 3 (L3) traffic. The figure shows the table entries forwarding the traffic to each of the egress ports. All the ingress and egress ports are of the same speed.

Figure 20: Reactive Path Rebalancing



In this topology, reactive path rebalancing works as follows:

1. Quality delta of 2 is configured.
2. L2 stream 1 (`mac 0x123`) enters ingress port et-0/0/0 with a rate of 10 percent. It exits through et-0/0/10. The egress link utilization of et-0/0/10 is 10 percent and the quality band value is 6.
3. The L3 stream enters port et-0/0/1 with a rate of 50 percent. It exits through et-0/0/11 and selects the optimal link from the ECMP member list. The egress link utilization of et-0/0/11 is 50 percent with a quality band value of 5.
4. L2 stream 2 (`mac 0x223`) enters port et-0/0/2 with a rate of 40 percent. It also exits through et-0/0/11. This further degrades the et-0/0/11 link quality band value to 4. Now the difference in the quality band values of both ECMP member links is 2.
5. The reactive path balancing algorithm now becomes operational because the difference in quality band values for ports et-0/0/10 and et-0/0/11 is equal to or higher than the configured quality delta of 2. The algorithm moves the L3 stream from et-0/0/11 to a better-quality member link, which in this case is et-0/0/10.
6. After the L3 steam moves to et-0/0/10, the et-0/0/10 link utilization increases to 60 percent with a decrease in quality band value to 5. L2 stream 2 continues to exit through et-0/0/11. The et-0/0/11 link utilization remains at 40 percent with an increase in quality band value to 5.

Configure Reactive Path Rebalancing

1. Configure DLB in flowlet mode. See Configuring Dynamic Load Balancing.
2. Configure the required difference (delta) in quality between the current stream member and the member available for reassignment.

   Optimal selection of the quality delta is very important. An incorrect delta can result in continuous reassignment of flow from one link to another.

The range of the `quality-delta` statement is 0 through 8. Set it to 0 to disable reassignment of the flows.

```
set forwarding-options enhanced-hash-key ecmp-dlb flowlet reassignment quality-delta reassign-quality-delta
```

3. Set the probability threshold that reactive path rebalancing uses to reassign the existing flow to a better available member link.

   Note the following when configuring the probability threshold:

   - When `quality-delta` is configured, `prob-threshold` defaults to 100.
   - The range of `prob-threshold` is 0 through 255. Set it to 0 to disable reassignment of the flows.
   - A lower probability threshold value means that flows move to a higher-quality member link at a slower rate. For example, flows move to a higher-quality link more quickly with a probability threshold value of 200 than with a probability threshold value of 50.

   ```
   set forwarding-options enhanced-hash-key ecmp-dlb flowlet reassignment prob-threshold reassign-prob-threshold
   ```

4. Verify the configuration was successful.

   ```
   show forwarding-options enhanced-hash-key
   ```

Change History Table

Feature support is determined by the platform and release you are using. Use Feature Explorer to determine if a feature is supported on your platform.

Release	Description
22.3R1-EVO	Starting in Junos OS Release 22.3R1-EVO, QFX5130-32CD switches support dynamic load balancing for both ECMP and LAG.
19.4R2-EVO	Starting in Junos OS evolved Release 19.4R2, QFX5220 switches support dynamic load balancing (DLB) for ECMP. For ECMP, DLB must be configured globally.
19.4R1	Starting in Junos OS Release 19.4R1, QFX5120-32C and QFX5120-48Y switches support dynamic load balancing for both ECMP and LAG. For LAG, DLB must be configured on a per aggregated Ethernet interface basis.

RELATED DOCUMENTATION

- *dynamic-load-balance*
- *enhanced-hash-key*
- *rdma-opcode*
- *show forwarding-options enhanced-hash-key*

Global Load Balancing (GLB)

SUMMARY

GLB is a form of load balancing that takes into account the link utilization of remote links before deciding on the egress interface.

IN THIS SECTION

- GLB Overview | **413**
- GLB in AI-ML Data Centers | **414**
- Configuration | **414**

GLB Overview

IN THIS SECTION

- Benefits | **414**

Classic load balancing mechanisms use a hashing algorithm to decide the egress interface through which to send traffic. These algorithms operate the hash function on five tuples of the received packet. However, the algorithms do not consider the real-time utilization of the links through which they send packets. Even in DLB, the decision is completely local and the algorithm is unable to globally detect link utilization. If a node farther out is congested, that node might drop the packet. Global load balancing (GLB) is an enhancement to DLB that has visibility into congestion at the next-to-next-hop (NNH) level.

GLB takes into account the link utilization of remote links before deciding on the egress interface. Similarly to DLB, when one multipath leg experiences congestion, GLB can offload traffic to alternative

legs to mitigate the congestion. Unlike DLB, GLB can reroute traffic flows on leaf devices to avoid traffic congestion on the spine level.

Use Feature Explorer to confirm platform and release support for specific features.

Benefits

- Reduces packet loss due to congestion and remote link failures
- Effectively load-balances large data flows in Clos topologies end-to-end to avoid congestion
- Is particularly useful in deployments where large data flows increase the likelihood of traffic congestion

GLB in AI-ML Data Centers

AI-ML data centers have less entropy and larger data flows than other networks. Because hash-based load balancing does not always effectively load-balance large data flows of traffic with less entropy, dynamic load balancing (DLB) is often used instead. However, DLB takes into account only the local link bandwidth utilization. For this reason, DLB can effectively mitigate traffic congestion only on the immediate next hop. GLB more effectively load-balances large data flows by taking traffic congestion on remote links into account.

Configuration

IN THIS SECTION

- Considerations | **414**
- Configure GLB | **415**

Considerations

Keep the following in mind when configuring GLB:

- GLB is supported only in a 3-Clos (leaf-spine-leaf) topology.
- All the devices in the 3-Clos topology must support GLB before you can configure GLB.

- The 3-Clos topology can have a maximum of 64 leaf devices when it supports GLB.
- GLB supports only one link between the same pair of devices (for example, a spine device and leaf device).

GLB does not support the following features:

- Integrated routing and bridging (IRB) interfaces between top-of-rack (ToR) and spine devices
- Multihomed servers
- GLB for overlay routes (IPv4 or IPv6)
- GLB for BGP routes learned in routing instances

Configure GLB

1. Configure DLB.

   The DLB configuration on each device in the fabric must be identical. See Dynamic Load Balancing for how to configure DLB.
2. Configure a node ID for each node.

   Each node must have a node ID. Keep the following in mind when configuring the node ID:

   - Configure the node ID at one of these hierarchy levels:

     ```
     [edit routing-optons router-id router-id]
     [edit protocols bgp bgp-identifier bgp-identifier]
     ```

   - If you configure the `bgp-identifier` statement, you must configure it globally, not at a `group` or `neighbor` hierarchy level.
   - The BGP identifier for each node must be unique within the fabric.
3. On spine devices, configure GLB in `helper-only` mode.

   In `helper-only` mode, BGP sends the NNH node (NNHN) capability for the route it advertises. BGP instructs the GLB application to monitor the link qualities of all local links with EBGP sessions and flood that information to all direct neighbors. Configure this option on the spine devices in a 3-Clos architecture.

   ```
   set protocols bgp global-load-balancing helper-only
   set forwarding-options enhanced-hash-key ecmp-dlb <flowlet | per-packet>
   ```
4. On leaf devices, configure GLB in `load-balancer-only` mode.

In `load-balancer-only` mode, BGP does not send the NNHN capability for the route it advertises. The switch receives link qualities from neighboring nodes. It uses the combined link quality of next hops and NNHs to make load balancing decisions. Configure this option on the leaf devices of any Clos architecture.

```
set protocols bgp global-load-balancing load-balancer-only
set forwarding-options enhanced-hash-key ecmp-dlb <flowlet | per-packet>
```

5. Selectively disable GLB.

   After you globally configure GLB using the `global-load-balancing` statement, you can selectively disable it on a particular BGP group or peer. To selectively disable GLB, use the `no-global-load-balancing` statement at either of these hierarchy levels:

   ```
   [edit protocols bgp group group-name]
   ```

   ```
   [edit protocols bgp group group-name neighbor address]
   ```

   For example:

   ```
   set protocols bgp group group-name no-global-load-balancing
   ```

6. Verify the configuration was successful using the following commands:

   - **show bgp global-load-balancing**
   - **show bgp global-load-balancing path**
   - **show bgp global-load-balancing path-monitor**
   - **show bgp global-load-balancing profile**

RELATED DOCUMENTATION

enhanced-hash-key

Dynamic Load Balancing (DLB) | 385

4

CHAPTER

Flexible Ethernet Services Encapsulation

IN THIS CHAPTER

- Flexible Ethernet Services Encapsulation | **418**

Flexible Ethernet Services Encapsulation

IN THIS SECTION

- Understanding Flexible Ethernet Services Encapsulation on Switches | 418
- Configuring Flexible Ethernet Services Encapsulation to Support the Service Provider and Enterprise Styles of Configuration | 421
- Configure Flexible Ethernet Services Encapsulation to Include Layer 2 Interface Support with Other Encapsulations | 424
- Configure Flexible Ethernet Services Encapsulation to Support Multiple Logical Interfaces on the Same Physical Interface Mapped to the Same Bridge Domain | 426

Flexible Ethernet services is a type of encapsulation that enables a physical interface to support different types of Ethernet encapsulations at the logical interface level. You can configure the Flexible Ethernet services encapsulation to support the service provider and the enterprise-style configuration. The below topics discuss the overview of flexible Ethernet services encapsulation and its configuration details.

Understanding Flexible Ethernet Services Encapsulation on Switches

IN THIS SECTION

- Service Provider Style | 419
- Enterprise Style | 419
- Flexible Ethernet Services | 420

Junos OS supports two different styles of configuration for switch interfaces: the service provider style and the enterprise style. The service provider style requires more configuration but provides greater flexibility. The enterprise style is easier to configure but offers less functionality. Each configuration style requires a different Ethernet encapsulation type. You can configure a physical interface to support both styles of configuration using flexible Ethernet services.



NOTE: On EX4300, QFX5100 (running Junos OS 16.1R5 or earlier), and QFX5200, the service provider style and enterprise style interface configurations are handled differently within Junos OS. If the service provider style and enterprise style interface configurations are mixed, the egress VLAN translation within the hardware can be incorrectly programmed leading to forwarding issues across the configured ports. Use the service provider style configuration in a Q-in-Q scenario. For all other scenarios, use the enterprise style configuration.

Flexible Ethernet services is a type of encapsulation that enables a physical interface to support different types of Ethernet encapsulations at the logical interface level. Defining multiple per-unit Ethernet encapsulations makes it easier to customize Ethernet-based services to multiple hosts connected to the same physical interface.

Service Provider Style

The service provider style of configuration allows for customization of Ethernet-based services at the logical interface level. Service providers typically have multiple customers connected to the same physical interface. Using the service provider style, you can configure multiple logical interfaces on the physical interface, and associate each unit with a different VLAN. This provides the flexibility to configure different services for each customer, but also requires more configuration, because each feature must be explicitly configured on the logical interface.

When configuring a physical interface to support only the service provider style, the physical interface must be encapsulated with the `extended-vlan-bridge` option to support bridging features. VLAN tagging must also be configured on the physical interface so that it can operate in trunk mode and transmit Ethernet frames with VLAN tags for multiple VLANs. Each logical interface is bound to a unique VLAN ID.

Enterprise Style

The enterprise style of configuration is designed to provide basic bridging functionality for consumers of Ethernet-based services. The isolation of services for different customers on a single port is not required, because each port is typically connected to a host or is providing a trunk to another switch.

With the enterprise style of configuration, logical interfaces are placed into Layer 2 mode by specifying `ethernet-switching` as the interface family. Without using flexible Ethernet services, `ethernet-switching` can only be configured on a single logical unit, unit 0. You cannot bind a VLAN ID to unit 0, because these interfaces operate either in trunk mode, which supports traffic with various VLAN tags, or in access mode, which supports untagged traffic.

Flexible Ethernet Services

The flexible Ethernet services encapsulation type enables a physical interface to support both styles of configuration. To support the service provider style, flexible Ethernet services allows for encapsulations to be configured at the logical interface level instead of the physical interface. To support the enterprise style, flexible Ethernet services allows the `ethernet-switching` family to be configured on any logical interface unit number instead of only unit 0.

For example, the configuration below shows three logical interfaces configured on a physical interface, xe-0/0/51, that is encapsulated for flexible Ethernet services. Unit 100 and unit 200 are configured in the service provider style and unit 300 is configured in the enterprise style. The encapsulation type of `vlan-bridge` is used to enable bridging on unit 100 and unit 200, and `family ethernet-switching` enables bridging on unit 300.

```
interfaces {
    xe-0/0/51 {
        flexible-vlan-tagging;
        encapsulation flexible-ethernet-services;
        unit 100 {
            encapsulation vlan-bridge;
            vlan-id 100;
        }
        unit 200 {
            encapsulation vlan-bridge;
            vlan-id 200;
        }
        unit 300 {
            family ethernet-switching {
                interface-mode trunk;
                vlan members 300;
            }
        }
    }
}
```

Following are the guidelines to follow when you configure the flexible Ethernet services encapsulation:

- On the QFX10000 line of switches, configuring either `vlan-tagging` or `flexible-vlan-tagging` with `family ethernet-switching` on the same interface is not supported.
- Only on the QFX10000 and EX9200 line of switches, you can enable `vlan-ccc` encapsulation when `flexible-ethernet-services` is already enabled.

- On QFX5100 switches, you can combine encapsulations on the same physical interface for `vlan-bridge` and `family ethernet switching`. Starting with Junos OS Release 16.1R6, you can also combine encapsulations on the same physical interface for `family inet` and `family ethernet-switching`.
- It is not required that the unit number and VLAN ID match, but it is considered a best practice.

Configuring Flexible Ethernet Services Encapsulation to Support the Service Provider and Enterprise Styles of Configuration

Flexible Ethernet services is a type of *encapsulation* that enables a physical interface to specify Ethernet encapsulations at the logical interface level. Each logical interface can have a different Ethernet encapsulation. Defining multiple per-unit Ethernet encapsulations makes it easier to customize Ethernet-based services to multiple hosts connected to the same physical interface.

An Ethernet interface that is not encapsulated with flexible Ethernet services and is operating in Layer 2 mode is limited to a single logical interface unit (0). Bridging is enabled on the interface by configuring `ethernet-switching` as the interface family on unit 0. The `ethernet-switching` family can be configured only on logical interface unit 0, and no other logical units can be defined on that interface.

Some switching features, however, cannot be configured on logical interface unit 0. Features such as Q-in-Q tunneling require the logical interface to transmit VLAN-tagged frames. To enable a logical interface to receive and forward Ethernet frames tagged with a matching VLAN ID, you must bind the logical interface to that VLAN. These features must be configured on a logical interface unit other than 0, because you cannot bind a VLAN ID to unit 0.

When you encapsulate an interface by using flexible Ethernet services, you can configure a logical interface unit other than 0 with `family ethernet-switching`. You can also configure other logical interfaces on that same interface with different types of Ethernet encapsulations. This enables logical interfaces that are bound to a VLAN ID to coexist with logical interfaces configured with `family ethernet-switching`.

For example, if you configure PVLAN on the same physical interface on which you are configuring Q-in-Q tunneling, you can use flexible ethernet services to support the enterprise style of configuration for PVLAN, using `family ethernet-switching`, along with `vlan-bridge` encapsulation for Q-in-Q tunneling.



BEST PRACTICE: We recommend you configure the following statements using groups when configuring devices that function as hardware VTEPs:

- set interfaces *interface-name* flexible-vlan-tagging
- set interfaces *interface-name* encapsulation extended-vlan-bridge

- set interfaces *interface-name* native-vlan-id *vlan-id*

To configure the interface to support both the service provider and enterprise styles of configuration:

1. Enable flexible Ethernet services encapsulation on the interface. The `flexible-ethernet-services` statement allows configuration of both service-provider-style logical interfaces and enterprise-style logical interfaces:

```
[edit interfaces interface-name]
user@switch# set encapsulation flexible-ethernet-services
```

2. Enable the interface to transmit packets with 802.1Q VLAN single-tagged and dual-tagged frames:

```
[edit interfaces interface-name]
user@switch# set flexible-vlan-tagging
```

3. Configure a logical interface (unit) on the interface:

```
[edit interfaces interface-name]
user@switch# set unit unit-number
```



NOTE: Do not use logical interface unit 0. You must later bind a VLAN tag ID to the unit you specify in this step, and you cannot bind a VLAN tag ID to unit 0. It is a best practice to match the unit number to the VLAN ID to which the interface is bound.

4. Encapsulate the logical interface for service provider style bridging configuration—for example, use `vlan-bridge` encapsulation on an interface to be used for Q-in-Q tunneling. (If you were configuring the interface only for Q-in-Q tunneling, you would use `encapsulation extended-vlan-bridge` on the *physical* interface.)

```
[edit interfaces interface-name]
user@switch# set unit unit-number encapsulation vlan-bridge
```

5. Bind the logical interface from the preceding step to a VLAN ID:

```
[edit interfaces interface-name]
user@switch# set unit unit-number vlan-id vlan-id
```

6. Configure another logical interface. (If you were configuring just PVLAN, we would recommend that you configure a single logical interface for all PVLAN domains on an interface.)

```
[edit interfaces interface-name]
user@switch# set unit unit-number
```

7. Enable the logical interface in the preceding step for enterprise style bridging configuration:

```
[edit interfaces interface-name]
user@switch# set unit unit-number family ethernet-switching
```

8. Assign VLAN membership to the logical interface:

```
[edit interfaces interface-name]
user@switch# set unit unit-number family ethernet-switching vlan members vlan-id
```

9. Configure the interface as a trunk interface to transmit frames with 802.1Q VLAN tags:

```
[edit interfaces interface-name]
user@switch# set unit unit-number family ethernet-switching interface-mode trunk
```



NOTE: For EX4300 device, the service provider style configuration (encapsulation extended-vlan-bridge) is recommended only for QinQ scenarios. For other scenarios, use the enterprise style configuration.

SEE ALSO

Configuring Q-in-Q Tunneling on QFX Series, NFX Series, and EX4600 Switches with ELS Support

Creating a Private VLAN on a Single Switch with ELS Support (CLI Procedure)

Configure Flexible Ethernet Services Encapsulation to Include Layer 2 Interface Support with Other Encapsulations

SUMMARY

Flexible Ethernet services is a type of encapsulation that enables a physical interface to specify Ethernet encapsulations at the logical interface level. Perform the following steps to configure flexible Ethernet services to support a Layer 2 bridging interface while simultaneously supporting other encapsulation options on the same physical interface.



NOTE: On the QFX10000 line of Switches running Junos OS releases earlier than Release 21.2R1, you cannot configure `vlan-bridging` and any other encapsulations on an interface that has `flexible-ethernet-services` enabled.

Configure a physical or aggregated Ethernet interface to simultaneously support a VLAN based circuit cross-connect (CCC) connection, Layer 3 IP routing, and Layer 2 bridging:

1. Enable flexible Ethernet services encapsulation on the interface.

```
[edit interfaces interface-name]
user@switch# set encapsulation flexible-ethernet-services
```

2. Configure the interface to support 802.1Q VLAN single-tagged and dual-tagged frames.

```
[edit interfaces interface-name]
user@switch# set vlan-tagging
```

3. Define a logical interface to support Ethernet VLAN encapsulation for CCC:

```
[edit interfaces interface-name]
user@switch# set unit unit-number encapsulation vlan-ccc
```

4. Bind the L2 CCC logical interface from the preceding step to a VLAN ID. This step is needed for all logical interfaces because the physical interface is set for VLAN tagged traffic.

```
[edit interfaces interface-name]
user@switch# set unit unit-number vlan-id vlan-id
```

5. Configure a second logical interface as an L3 routed IP interface.

```
[edit interfaces interface-name]
user@switch# set unit unit-number family inet address ip-address/mask
```

6. Bind the L3 logical interface from the preceding step to a VLAN ID:

```
[edit interfaces interface-name]
user@switch# set unit unit-number vlan-id vlan-id
```

7. Configure a third logical interface to support VLAN based bridging by specifying `vlan-bridge` encapsulation on the logical unit.

```
[edit interfaces interface-name]
user@switch# set unit unit-number encapsulation vlan-bridge
```

8. Bind the logical interface from the preceding step to a VLAN ID.

```
[edit interfaces interface-name]
user@switch# set unit unit-number vlan-id vlan-id
```

9. Define a bridge domain and add the L2 logical interface.

```
[edit]
user@switch# set bridge-domains bridge-domain-name vlan-id vlan-id
```

```
[edit]
user@switch# set bridge-domains bridge-domain-name interface interface-id
```

Verify your configuration using the `show interfaces` *interface-name* command in the configuration mode.

```
user@switch> show interfaces xe-0/0/0
vlan-tagging;
encapsulation flexible-ethernet-services;
unit 1 {
    encapsulation vlan-ccc;
    vlan-id 103;
   }
unit 2 {
    vlan-id 102;
    family inet {
        address 10.0.0.1/30;
    }
unit 3 {
    encapsulation vlan-bridge;
    vlan-id 101;
   }

}
```

SEE ALSO

> *encapsulation (interfaces)*
>
> *encapsulation (Logical Interface)*
>
> *vlan-tagging*

Configure Flexible Ethernet Services Encapsulation to Support Multiple Logical Interfaces on the Same Physical Interface Mapped to the Same Bridge Domain

SUMMARY

Flexible Ethernet services is a type of encapsulation that enables a physical interface to specify Ethernet encapsulations at the logical interface level. Perform the following steps to configure multiple logical interfaces on the same physical interface mapped to the same bridge domain.

NOTE: The QFX10002-60C switches do not support this feature.

Configure a physical or aggregated Ethernet interface to simultaneously support multiple logical interfaces using the same bridge domain. You cannot configure an ESI interface as one of the logical interfaces over a physical interface when both are part of the same VLAN. ESI interfaces have a limitation.

NOTE: The combination of enterprise-style and service provider-style interfaces on the same physical interface is not supported when there are multiple service provider style logical interfaces attached to the same bridge domain.

1. Configure the interface to support 802.1Q VLAN single-tagged and dual-tagged frames.

```
[edit interfaces interface-name]
user@switch# set vlan-tagging
```

2. Enable flexible Ethernet services encapsulation on the interface.

```
[edit interfaces interface-name]
user@switch# set encapsulation flexible-ethernet-services
```

3. Configure a logical interface to support VLAN based bridging by specifying `vlan-bridge` encapsulation on the logical unit.

```
[edit interfaces interface-name]
user@switch# set unit unit-number encapsulation vlan-bridge
```

4. Bind the logical interface from the preceding step to a VLAN ID. This step is needed for all logical interfaces because the physical interface is set for VLAN tagged traffic.

```
[edit interfaces interface-name]
user@switch# set unit unit-number vlan-id vlan-id
```

5. Configure another logical interface to support VLAN based bridging by specifying `vlan-bridge` encapsulation on the logical unit.

```
[edit interfaces interface-name]
user@switch# set unit unit-number encapsulation vlan-bridge
```

6. Bind the logical interface from the preceding step to a VLAN ID.

```
[edit interfaces interface-name]
user@switch# set unit unit-number vlan-id vlan-id
```

7. Configure a bridge domain by specifying the VLAN name and assigning a VLAN ID:

```
[edit]
user@switch# set vlans vlan-name vlan-id vlan-id
```

8. Bind the first logical interface to the bridge domain:

```
[edit]
user@switch# set vlans vlan-name interfaces interface-name.unit
```

9. Bind the second logical interface to the bridge domain.

```
[edit]
user@switch# set vlans vlan-name  interfaces interface-name.unit
```

Verify your configuration using the `show interfaces interface-name` and `show vlans` command in the configuration mode.

```
user@switch> show interfaces xe-0/0/2:2
flexible-vlan-tagging;
encapsulation flexible-ethernet-services;
unit 1 {
    encapsulation vlan-bridge;
    vlan-id 1;
}
unit 2 {
    encapsulation vlan-bridge;
```

```
    vlan-id 2;
}
```

```
user@switch> show vlans
v100 {
    vlan-id 100;
    interface xe-0/0/2:2.1;
    interface xe-0/0/2:2.2;
}
```

SEE ALSO

- *encapsulation (interfaces)*
- *encapsulation (Logical Interface)*
- *vlan-tagging*
- *vlans*

Change History Table

Feature support is determined by the platform and release you are using. Use Feature Explorer to determine if a feature is supported on your platform.

Release	Description
16.1R3	Starting with Junos OS Release 16.1R6, you can also combine encapsulations on the same physical interface for `family inet` and `family ethernet-switching`.

5

CHAPTER

Monitoring and Troubleshooting Information

IN THIS CHAPTER

- Monitoring Interfaces | **431**
- Troubleshooting Interfaces | **440**

Monitoring Interfaces

IN THIS SECTION

- Monitoring Interface Status and Traffic | 431
- Monitoring System Process Information | 432
- Monitoring System Properties | 433
- Monitor Statistics for a Fast Ethernet or Gigabit Ethernet Interface | 436
- Trace Operations of the Interface Process | 438

The below topics discuss the monitoring of the status and traffic, system process information, system properties, statistics for a fast Ethernet and the tracing operations of the interface process.

Monitoring Interface Status and Traffic

IN THIS SECTION

- Purpose | 431
- Action | 431
- Meaning | 432

Purpose

View interface status to monitor interface bandwidth utilization and traffic statistics.

Action

- To view interface status for all the interfaces, enter **show interfaces xe**.
- To view status and statistics for a specific interface, enter **show interfaces xe** *interface-name*.

- To view status and traffic statistics for all interfaces, enter either **show interfaces xe detail** or **show interfaces xe extensive**.

Meaning

For details about output from the CLI commands, see *show interfaces xe*.

Monitoring System Process Information

IN THIS SECTION

- Purpose | 432
- Action | 432
- Meaning | 432

Purpose

View the processes running on the device.

Action

To view the software processes running on the device:

user@switch> **show system processes**

Meaning

Table 82 on page 433 summarizes the output fields in the system process information display.

The display includes the total CPU load and total memory utilization.

Table 82: Summary of System Process Information Output Fields

Field	Values
PID	Identifier of the process.
Name	Owner of the process.
State	Current state of the process.
CPU Load	Percentage of the CPU that is being used by the process.
Memory Utilization	Amount of memory that is being used by the process.
Start Time	Time of day when the process started.

SEE ALSO

show system uptime

Monitoring System Properties

IN THIS SECTION

- Purpose | 433
- Action | 434
- Meaning | 434

Purpose

View system properties such as the name, IP address, and resource usage.

Action

To monitor system properties in the CLI, enter the following commands:

- **show system uptime**
- **show system users**
- **show system storage**

Meaning

Table 83 on page 434 summarizes key output fields in the system properties display.

Table 83: Summary of Key System Properties Output Fields

Field	Values	Additional Information
General Information		
Serial Number	Serial number of device.	
Junos OS Version	Version of Junos OS active on the switch, including whether the software is for domestic or export use.	Export software is for use outside the USA and Canada.
Hostname	Name of the device.	
IP Address	IP address of the device.	
Loopback Address	Loopback address.	
Domain Name Server	Address of the domain name server.	
Time Zone	Time zone on the device.	
Time		

Table 83: Summary of Key System Properties Output Fields *(Continued)*

Field	Values	Additional Information
Current Time	Current system time, in Coordinated Universal Time (UTC).	
System Booted Time	Date and time when the device was last booted and how long it has been running.	
Protocol Started Time	Date and time when the protocols were last started and how long they have been running.	
Last Configured Time	Date and time when a configuration was last committed. This field also shows the name of the user who issued the last `commit` command.	
Load Average	CPU load average for 1, 5, and 15 minutes.	
Storage Media		
Internal Flash Memory	Usage details of internal flash memory.	
External Flash Memory	Usage details of external USB flash memory.	
Logged in Users Details		
User	Username of any user logged in to the switch.	
Terminal	Terminal through which the user is logged in.	
From	System from which the user has logged in. A hyphen indicates that the user is logged in through the console.	

Table 83: Summary of Key System Properties Output Fields *(Continued)*

Field	Values	Additional Information
Login Time	Time when the user logged in.	This is the **user@switch** field in `show system users` command output.
Idle Time	How long the user has been idle.	

SEE ALSO

show system processes

Monitor Statistics for a Fast Ethernet or Gigabit Ethernet Interface

IN THIS SECTION

- Purpose | 436
- Action | 436
- Meaning | 437

Purpose

To monitor statistics for a Fast Ethernet or Gigabit Ethernet interface, use the following Junos OS CLI operational mode command:

Action

```
user@host> monitor interface (fe-fpc/pic/port | ge-fpc/pic/port)
```

We recommend that you use the `monitor interface fe-`*fpc/pic/port* or `monitor interface ge-`*fpc/pic/port* command only for diagnostic purposes. Do not leave these commands on during normal router operations because real-time monitoring of traffic consumes additional CPU and memory resources.

Sample Output

The following sample output is for a Fast Ethernet interface:

```
user@host> monitor interface fe-2/1/0
Interface: fe-2/1/0, Enabled, Link is Up
Encapsulation: Ethernet, Speed: 100mbps
Traffic statistics:                                      Current Delta
  Input bytes:               282556864218 (14208 bps)          [40815]
  Output bytes:               42320313078 (384 bps)              [890]
  Input packets:                739373897 (11 pps)               [145]
  Output packets:               124798688 (1 pps)                 [14]
Error statistics:
  Input errors:                         0                         [0]
  Input drops:                          0                         [0]
  Input framing errors:                 0                         [0]
  Policed discards:               6625892                         [6]
  L3 incompletes:                      75                         [0]
  L2 channel errors:                    0                         [0]
  L2 mismatch timeouts:                 0                         [0]
  Carrier transitions:                  1                         [0]
  Output errors:                        0                         [0]
  Output drops:                         0                         [0]
  Aged packets:                         0                         [0]
Active alarms : None
Active defects: None
Input MAC/Filter statistics:
  Unicast packets               464751787                       [154]
  Packet error count                    0                         [0]
```

Meaning

Use the information from this command to help narrow down possible causes of an interface problem.

If you are accessing the router from the console connection, make sure you set the CLI terminal type using the `set cli terminal` command.

The second column shows cumulative statistics since the last time you cleared them using the `clear interfaces statistics interface-name` command. The third column shows cumulative statistics since you ran the `monitor interface interface-name` command. If input errors are increasing, follow these steps:

- Check the cabling to the router and ask the carrier to verify the line's integrity. Ensure you are using the correct cables for the interface port—single-mode fiber for a single-mode interface, and multimode fiber for a multimode interface.
- For fiber-optic connections, measure the received light level at the receiver end and ensure it meets the Ethernet interface's specification.
- Measure the transmit light level on the Tx port to confirm it is within the specified range.

Trace Operations of the Interface Process

To trace the operations of the router or switch interface process, dcd, perform the following steps:

1. In configuration mode, go to the `[edit interfaces]` hierarchy level:

```
[edit]
user@host# edit interfaces
```

2. Configure the `traceoptions` statement.

```
[edit interfaces]
user@host# edit traceoptions
```

3. Configure the `no-remote-trace` option to disable remote tracing.

```
[edit interfaces traceoptions]
user@host# set no-remote-trace
```

4. Configure the `file` *`filename`* option.

```
[edit interfaces traceoptions]
user@host# edit file
```

5. Configure the `files` *`number`* option, `match` *`regular-expression`* option, `size` *`size`* option, and `world-readable | no-world-readable` option.

```
[edit interfaces traceoptions file]
user@host# set files number
user@host# set match regular-expression
user@host# set size size
user@host# set word-readable | no-world-readable
```

6. Configure the tracing flag.

```
[edit interfaces traceoptions]
user@host# set flag flag-option
```

7. Configure the `disable` option in `flag` *`flag-option`* statement to disable the tracing operation. You can use this option to disable a single operation when you have defined a broad group of tracing operations, such as `all`.

```
[edit interfaces traceoptions]
user@host# set flag flag-option disable
```

You can specify the following flags in the `interfaces traceoptions` statement:

- `all`—Enable all configuration logging.
- `change-events`—Log changes that produce configuration events.
- `gres-events`—Log the events related to GRES.
- `resource-usage`—Log the resource usage for different states.
- `config-states`—Log the configuration state machine changes.
- `kernel`—Log configuration IPC messages to kernel.
- `kernel-detail`—Log details of configuration messages to kernel.
- `select-events`—Log the events on select state machine.

By default, interface process operations are placed in the file named dcd and three 1-MB files of tracing information are maintained.

For general information about tracing, see the tracing and logging information in the Junos OS Administration Library for Routing Devices.

SEE ALSO

Tracing Interface Operations Overview

Tracing Operations of an Individual Router Interface

traceoptions

Troubleshooting Interfaces

IN THIS SECTION

- Troubleshooting Network Interfaces | **440**
- Diagnosing a Faulty Twisted-Pair Cable (CLI Procedure) | **442**
- Troubleshooting Uplink Ports on EX2300 Switches | **446**

The below topics discuss the troubleshooting of network interfaces and diagnosing a faulty twisted-pair cable.

Troubleshooting Network Interfaces

IN THIS SECTION

- Statistics for logical interfaces on Layer 2 interfaces are not accurate | **441**
- The interface on the port in which an SFP or SFP+ transceiver is installed in an SFP or SFP+ module is down | **441**

Statistics for logical interfaces on Layer 2 interfaces are not accurate

IN THIS SECTION

- Problem | 441
- Solution | 441

Problem

Description

On QFX5000 switches, statistics for logical interfaces are not supported on Layer 2 interfaces or on any child member interfaces of Layer 2 aggregated Ethernet (AE) interfaces—that is, output for the `show interfaces` *`interface-name`* operational-mode command does not provide accurate I/O information for the logical interfaces.

Solution

If you need to see statistics for those logical interfaces, configure firewall filter rules to collect the information.

The interface on the port in which an SFP or SFP+ transceiver is installed in an SFP or SFP+ module is down

IN THIS SECTION

- Problem | 442
- Cause | 442
- Solution | 442

Problem

Description

The switch has an SFP or SFP+ module installed. The interface on the port in which an SFP or SFP+ transceiver is installed is down.

Symptoms

When you check the status with the CLI command `show interfaces` *`interface-name`*, the disabled port is not listed.

Cause

By default, the SFP or SFP+ module operates in the 10-Gigabit Ethernet mode and supports only SFP or SFP+ transceivers. The operating mode for the module is incorrectly set.

Solution

Only SFP or SFP+ transceivers can be installed in SFP or SFP+ modules. You must configure the operating mode of the SFP or SFP+ module to match the type of transceiver you want to use. For SFP+ transceivers, configure 10-Gigabit Ethernet operating mode.

Diagnosing a Faulty Twisted-Pair Cable (CLI Procedure)

IN THIS SECTION

- Problem | 442
- Solution | 443

Problem

Description

A 10/100/1000BASE-T Ethernet interface has connectivity problems that you suspect might be caused by a faulty cable.

Solution

Use the time domain reflectometry (TDR) test to determine whether a twisted-pair Ethernet cable is faulty.

The TDR test:

- Detects and reports faults for each twisted pair in an Ethernet cable. Faults detected include open circuits, short circuits, and impedance mismatches.
- Reports the distance to fault to within 1 meter.
- Detects and reports pair swaps, pair polarity reversals, and excessive pair skew.

The TDR test is supported on the following switches and interfaces:

- EX2200, EX2300, EX3200, EX3300, EX3400, EX4200, and EX4300 switches—RJ-45 network interfaces. The TDR test is not supported on management interfaces and SFP interfaces.
- EX6200 and EX8200 switches—RJ-45 network interfaces on line cards.



NOTE: We recommend running the TDR tests on an interface when there is no traffic on the interface.

Use the following guidelines while running the TDR test:

- Run a TDR test on a port that is operationally "down".
- When you perform a TDR test at an operationally "up" interface, the diagnostic test result can be inconclusive.
- Have sufficient time between repeated diagnostic tests on the same port.
- Run a TDR test on a port that is operationally "down". When you perform a TDR test on an operationally "up" interface.
  - The port flaps
  - The diagnostic test result can be inconclusive



NOTE: When you run a TDR test on a port that is in "up" state, you can expect a port flap.

To diagnose a cable problem by running the TDR test:

1. Run the `request diagnostics tdr` command.

```
user@switch> request diagnostics tdr start interface ge-0/0/10

Interface TDR detail:
Test status                    : Test successfully executed  ge-0/0/10
```

2. View the results of the TDR test with the `show diagnostics tdr` command.

```
user@switch> show diagnostics tdr interface ge-0/0/10

Interface TDR detail:
Interface name                 : ge-0/0/10
Test status                    : Passed
Link status                    : Down
MDI pair                       : 1-2
  Cable status                 : Normal
  Distance fault               : 0 Meters
  Polartiy swap                : N/A
  Skew time                    : N/A
MDI pair                       : 3-6
  Cable status                 : Normal
  Distance fault               : 0 Meters
  Polartiy swap                : N/A
  Skew time                    : N/A
MDI pair                       : 4-5
  Cable status                 : Open
  Distance fault               : 1 Meters
  Polartiy swap                : N/A
  Skew time                    : N/A
MDI pair                       : 7-8
  Cable status                 : Normal
  Distance fault               : 0 Meters
  Polartiy swap                : N/A
  Skew time                    : N/A
Channel pair                   : 1
  Pair swap                    : N/A
Channel pair                   : 2
  Pair swap                    : N/A
Downshift                      : N/A
```

3. Examine the **Cable status** field for the four MDI pairs to determine if the cable has a fault. In the preceding example, the twisted pair on pins 4 and 5 is broken or cut at approximately one meter from the **ge-0/0/10** port connection.

> **NOTE**: The **Test Status** field indicates the status of the TDR test, not the cable. The value **Passed** means the test completed—it does not mean that the cable has no faults.

The following is additional information about the TDR test:

- The TDR test can take some seconds to complete. If the test is still running when you execute the `show diagnostics tdr` command, the **Test status** field displays **Started**. For example:

```
user@switch> show diagnostics tdr interface ge-0/0/22

Interface TDR detail:
Interface name                 : ge-0/0/22
Test status                    : Started
```

- You can terminate a running TDR test before it completes by using the `request diagnostics tdr abort interface` *`interface-name`* command. The test terminates with no results, and the results from any previous test are cleared.

- You can display summary information about the last TDR test results for all interfaces on the switch that support the TDR test by not specifying an interface name with the `show diagnostics tdr` command. For example:

```
user@switch> show diagnostics tdr
Interface   Test status    Link status  Cable status  Max distance fault
 ge-0/0/0        Passed     UP           OK             0
 ge-0/0/1   Not Started    N/A          N/A            N/A
 ge-0/0/2        Passed     UP           OK             0
 ge-0/0/3   Not Started    N/A          N/A            N/A
 ge-0/0/4        Passed     UP           OK             0
 ge-0/0/5        Passed     UP           OK             0
 ge-0/0/6        Passed     UP           OK             0
 ge-0/0/7   Not Started    N/A          N/A            N/A
 ge-0/0/8        Passed    Down          OK             0
 ge-0/0/9   Not Started    N/A          N/A            N/A
ge-0/0/10        Passed    Down         Fault           1
ge-0/0/11        Passed     UP           OK             0
ge-0/0/12   Not Started    N/A          N/A            N/A
```

```
ge-0/0/13   Not Started    N/A          N/A          N/A
ge-0/0/14   Not Started    N/A          N/A          N/A
ge-0/0/15   Not Started    N/A          N/A          N/A
ge-0/0/16   Not Started    N/A          N/A          N/A
ge-0/0/17   Not Started    N/A          N/A          N/A
ge-0/0/18   Not Started    N/A          N/A          N/A
ge-0/0/19        Passed    Down          OK            0
ge-0/0/20   Not Started    N/A          N/A          N/A
ge-0/0/21   Not Started    N/A          N/A          N/A
ge-0/0/22        Passed     UP           OK            0
ge-0/0/23   Not Started    N/A          N/A          N/A
```

SEE ALSO

Troubleshooting Interface Configuration and Cable Faults

request diagnostics tdr

show diagnostics tdr

Troubleshooting Uplink Ports on EX2300 Switches

IN THIS SECTION

- Speeds 10-Mbps and 100-Mbps not supported on uplink ports 4 and 5 on EX2300-48MP switches | **447**

This topic provides troubleshooting information for specific problems related to interfaces on EX2300 switches.

Speeds 10-Mbps and 100-Mbps not supported on uplink ports 4 and 5 on EX2300-48MP switches

IN THIS SECTION

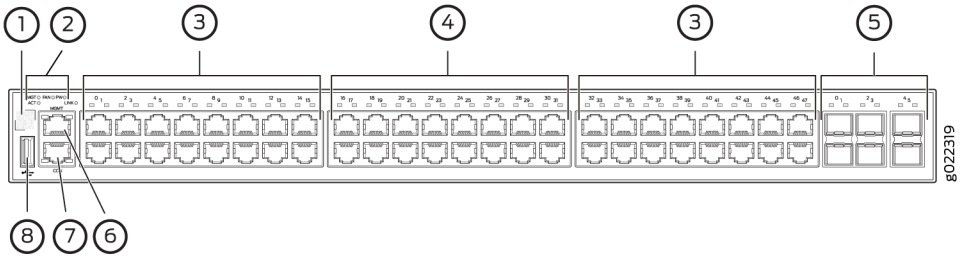
- Problem | 447
- Cause | 448
- Solution | 448

Problem

Description

The uplink ports 4 and 5 (see Figure 1) do not support the speeds 10-Mbps and 100-Mbps.

Figure 21: Front Panel of an EX2300-48MP Switch



1– QR code	5– 10-Gigabit Ethernet uplink ports
2– System LEDs	6– Management port
3– 10/100/1000 BASE-T Gigabit Ethernet ports with PoE/PoE+ capability	7– Console port
4– 100/1000/2500 BASE-T Gigabit Ethernet ports	8– USB port

Environment

A transceiver is installed in the uplink port 4 or 5 or both.

Symptoms

When you check the status with the CLI command `show interfaces ge` or with the J-Web user interface, the port is not listed.

Cause

EX2300-48MP switches do not support 10-Mbps and 100-Mbps speeds on uplink ports 4 and 5. This is an ASIC limitation.

Solution

Use the other ports if you need 10-Mbps and 100-Mbps speeds.

CHAPTER 6

Configuration Statements and Operational Commands

IN THIS CHAPTER

- Common Output Fields Description | **450**
- Junos CLI Reference Overview | **460**

Common Output Fields Description

IN THIS SECTION

- Damping Field | 450
- Destination Class Field | 451
- Enabled Field | 451
- Filters Field | 452
- Flags Fields | 452
- Label-Switched Interface Traffic Statistics Field | 457
- Policer Field | 458
- Protocol Field | 458
- RPF Failures Field | 459
- Source Class Field | 459

This chapter explains the content of the output fields, which appear in the output of most **show interfaces** commands.

Damping Field

For the physical interface, the Damping field shows the setting of the following damping parameters:

- `half-life`—Decay half-life. The number of seconds after which the accumulated interface penalty counter is reduced by half if the interface remains stable.
- `max-suppress`—Maximum hold-down time. The maximum number of seconds that an interface can be suppressed irrespective of how unstable the interface has been.
- `reuse`—Reuse threshold. When the accumulated interface penalty counter falls below this number, the interface is no longer suppressed.
- `suppress`—Cutoff (suppression) threshold. When the accumulated interface penalty counter exceeds this number, the interface is suppressed.

- `state`—Interface damping state. If damping is enabled on an interface, it is suppressed during interface flaps that match the configured damping parameters.

Destination Class Field

For the logical interface, the `Destination class` field provides the names of destination class usage (DCU) counters per family and per class for a particular interface. The counters display packets and bytes arriving from designated user-selected prefixes. For example:

```
                                   Packets                    Bytes
Destination class                  (packet-per-second)        (bits-per-second)

             gold                          1928095                 161959980
                                (              889)           (   597762)
            bronze                            0                       0
                                (              0)             (        0)
            silver                            0                       0
                                (              0)             (        0)
```

Enabled Field

For the physical interface, the `Enabled` field provides information about the state of the interface, displaying one or more of the following values:

- `Administratively down, Physical link is Down`—The interface is turned off, and the physical link is inoperable and cannot pass packets even when it is enabled.To change the interface state to `Enabled`, use the following command:

  ```
  user@host# set interfaces interface enable
  ```

  Manually verify the connections to bring the physical link up.

- `Administratively down, Physical link is Up`—The interface is turned off, but the physical link is operational and can pass packets when it is enabled.To change the interface state to `Enabled`, use the following command:

```
user@host# set interfaces interface enable
```

- `Enabled, Physical link is Down`—The interface is turned on, but the physical link is inoperable and cannot pass packets. Manually verify the connections to bring the physical link up.
- `Enabled, Physical link is Up`—The interface is turned on, and the physical link is operational and can pass packets.

Filters Field

For the logical interface, the `Filters` field provides the name of the firewall filters to be evaluated when packets are received or transmitted on the interface. The format is `Filters: Input: filter-name` and `Filters: Output: filter-name`. For example:

```
Filters: Input: sample-all
Filters: Output: cp-ftp
```

Flags Fields

The following sections provide information about flags that are specific to interfaces:

Addresses, Flags Field

The `Addresses, Flags` field provides information about the addresses configured for the protocol family on the logical interface and displays one or more of the following values:

- `Dest-route-down`—The routing process detected that the link was not operational and changed the interface routes to nonforwarding status
- `Is-Default`—The default address of the router used as the source address by SNMP, ping, traceroute, and other network utilities.

- `Is-Preferred`—The default local address for packets originating from the local router and sent to destinations on the subnet.
- `Is-Primary`—The default local address for broadcast and multicast packets originated locally and sent out the interface.
- `Preferred`—This address is a candidate to become the preferred address.
- `Primary`—This address is a candidate to become the primary address.
- `Trunk`—Interface is a trunk.
- `Trunk, Inter-Switch-Link`—Interface is a trunk, and InterSwitch Link protocol (ISL) is configured on the trunk port of the primary VLAN in order to connect the routers composing the PVLAN to each other.

Device Flags Field

The `Device flags` field provides information about the physical device and displays one or more of the following values:

- `ASIC Error`—Device is down because of ASIC wedging and due to which PFE is disabled.
- `Down`—Device has been administratively disabled.
- `Hear-Own-Xmit`—Device receives its own transmissions.
- `Link-Layer-Down`—The link-layer protocol has failed to connect with the remote endpoint.
- `Loopback`—Device is in physical loopback.
- `Loop-Detected`—The link layer has received frames that it sent, thereby detecting a physical loopback.
- `No-Carrier`—On media that support carrier recognition, no carrier is currently detected.
- `No-Multicast`—Device does not support multicast traffic.
- `Present`—Device is physically present and recognized.
- `Promiscuous`—Device is in promiscuous mode and recognizes frames addressed to all physical addresses on the media.
- `Quench`—Transmission on the device is quenched because the output buffer is overflowing
- `Recv-All-Multicasts`—Device is in multicast promiscuous mode and therefore provides no multicast filtering.
- `Running`—Device is active and enabled.

Family Flags Field

The `Family flags` field provides information about the protocol family on the logical interface and displays one or more of the following values:

- `DCU`—Destination class usage is enabled.
- `Dest-route-down`—The software detected that the link is down and has stopped forwarding the link's interface routes.
- `Down`—Protocol is inactive.
- `Is-Primary`—Interface is the primary one for the protocol.
- `Mac-Validate-Loose`—Interface is enabled with loose MAC address validation.
- `Mac-Validate-Strict`—Interface is enabled with strict MAC address validation.
- `Maximum labels`—Maximum number of MPLS labels configured for the MPLS protocol family on the logical interface.
- `MTU-Protocol-Adjusted`—The effective MTU is not the configured value in the software.
- `No-Redirects`—Protocol redirects are disabled.
- `Primary`—Interface can be considered for selection as the primary family address.
- `Protocol-Down`—Protocol failed to negotiate correctly.
- `SCU-in`—Interface is configured for source class usage input.
- `SCU-out`—Interface is configured for source class usage output.
- `send-bcast-packet-to-re`—Interface is configured to forward IPv4 broadcast packets to the Routing Engine.
- `targeted-broadcast`—Interface is configured to forward IPv4 broadcast packets to the LAN interface and the Routing Engine.
- `Unnumbered`—Protocol family is configured for unnumbered Ethernet. An unnumbered Ethernet interface borrows an IPv4 address from another interface, which is referred to as the donor interface.
- `Up`–Protocol is configured and operational.
- `uRPF`—Unicast Reverse Path Forwarding is enabled.

Interface Flags Field

The `Interface flags` field provides information about the physical interface and displays one or more of the following values:

- `Admin-Test`—Interface is in test mode and some sanity checking, such as loop detection, is disabled.
- `Disabled`—Interface is administratively disabled.
- `Down`—A hardware failure has occurred.
- `Hardware-Down`—Interface is nonfunctional or incorrectly connected.
- `Link-Layer-Down`—Interface keepalives have indicated that the link is incomplete.
- `No-Multicast`—Interface does not support multicast traffic.
- `No-receive No-transmit`—Passive monitor mode is configured on the interface.
- `OAM-On-SVLAN`—(MX Series routers with MPC/MIC interfaces only) Interface is configured to propagate the Ethernet OAM state of a static, single-tagged service VLAN (S-VLAN) on a Gigabit Ethernet, 10-Gigabit Ethernet, or aggregated Ethernet interface to a dynamic or static double-tagged customer VLAN (C-VLAN) that has the same S-VLAN (outer) tag as the S-VLAN.
- `Point-To-Point`—Interface is point-to-point.
- `Pop all MPLS labels from packets of depth`—MPLS labels are removed as packets arrive on an interface that has the `pop-all-labels` statement configured. The depth value can be one of the following:
  - `1`—Takes effect for incoming packets with one label only.
  - `2`—Takes effect for incoming packets with two labels only.
  - `[ 1 2 ]`—Takes effect for incoming packets with either one or two labels.
- `Promiscuous`—Interface is in promiscuous mode and recognizes frames addressed to all physical addresses.
- `Recv-All-Multicasts`—Interface is in multicast promiscuous mode and provides no multicast filtering.
- `SNMP-Traps`—SNMP trap notifications are enabled.
- `Up`—Interface is enabled and operational.

Link Flags Field

The `Link flags` field provides information about the physical link and displays one or more of the following values:

- ACFC—Address control field compression is configured. The Point-to-Point Protocol (PPP) session negotiates the ACFC option.
- Give-Up—Link protocol does not continue connection attempts after repeated failures.
- Loose-LCP—PPP does not use the Link Control Protocol (LCP) to indicate whether the link protocol is operational.
- Loose-LMI—Frame Relay does not use the Local Management Interface (LMI) to indicate whether the link protocol is operational.
- Loose-NCP—PPP does not use the Network Control Protocol (NCP) to indicate whether the device is operational.
- No-Keepalives—Link protocol keepalives are disabled.
- PFC—Protocol field compression is configured. The PPP session negotiates the PFC option.

Logical Interface Flags Field

The Logical interface flags field provides information about the logical interface and displays one or more of the following values:

- ACFC Encapsulation—Address control field Compression (ACFC) encapsulation is enabled (negotiated successfully with a peer).
- Device-down—Device has been administratively disabled.
- Disabled—Interface is administratively disabled.
- Down—A hardware failure has occurred.
- Clear-DF-Bit—GRE tunnel or IPsec tunnel is configured to clear the Don't Fragment (DF) bit.
- Hardware-Down—Interface protocol initialization failed to complete successfully.
- PFC—Protocol field compression is enabled for the PPP session.
- Point-To-Point—Interface is point-to-point.
- SNMP-Traps—SNMP trap notifications are enabled.
- Up—Interface is enabled and operational.

Label-Switched Interface Traffic Statistics Field

When you use the vrf-table-label statement to configure a VRF routing table, a label-switched interface (LSI) logical interface label is created and mapped to the VRF routing table.

Any routes present in a VRF routing table and configured with the vrf-table-label statement are advertised with the LSI logical interface label allocated for the VRF routing table. When packets for this VPN arrive on a core-facing interface, they are treated as if the enclosed IP packet arrived on the LSI interface and are then forwarded and filtered based on the correct table. For more information on the vrf-table-label statement, including a list of supported interfaces, see the *Junos VPNs Configuration Guide*.

If you configure the family mpls statement at the [edit interfaces *interface-name* unit *logical-unit-number*] hierarchy level and you also configure the vrf-table-label statement at the [edit routing-instances *routing-instance-name*] hierarchy level, the output for the show interface *interface-name* extensive command includes the following output fields about the LSI traffic statistics:

- Input bytes—Number of bytes entering the LSI and the current throughput rate in bits per second (bps).
- Input packets—Number of packets entering the LSI and the current throughput rate in packets per second (pps).



NOTE: If LSI interfaces are used with VPLS when no-tunnel-services is configured or L3VPN when vrf-table-label configuration is applied inside the routing-instance, the Input packets field associated with the core-facing interfaces may not display the correct value. Only the Input counter is affected because the LSI is used to receive traffic from the remote PEs. Traffic that arrives on an LSI interface might not be counted at both the Traffic Statistics and the Label-switched interface (LSI) traffic statistics levels.

This note applies to the following platforms:

- M Series routers with -E3 FPC model numbers or configured with an Enhanced CFEB (CFEB-E), and M120 routers
- MX Series routers with DPC or ADPC only

The following example shows the LSI traffic statistics that you might see as part of the output of the show interface *interface-name* extensive command:

```
Label-switched interface (LSI) traffic statistics:
   Input  bytes:                        0                0 bps
   Input  packets:                    0                0 pps
```

Policer Field

For the logical interface, the `Policer` field provides the policers that are to be evaluated when packets are received or transmitted on the interface. The format is `Policer: Input: type-fpc/picport-in-policer, Output: type-fpc/pic/port-out-policer`. For example:

```
Policer: Input: at-1/2/0-in-policer, Output: at-2/4/0-out-policer
```

Protocol Field

For the logical interface, the `Protocol` field indicates the protocol family or families that are configured on the interface, displaying one or more of the following values:

- `aenet`—Aggregated Ethernet. Displayed on Fast Ethernet interfaces that are part of an aggregated Ethernet bundle.
- `ccc`—Circuit cross-connect (CCC). Configured on the logical interface of CCC physical interfaces.
- `inet`—IP version 4 (IPv4). Configured on the logical interface for IPv4 protocol traffic, including Open Shortest Path First (OSPF), Border Gateway Protocol (BGP), Internet Control Message Protocol (ICMP), and Internet Protocol Control Protocol (IPCP).
- `inet6`—IP version 6 (IPv6). Configured on the logical interface for IPv6 protocol traffic, including Routing Information Protocol for IPv6 (RIPng), Intermediate System-to-Intermediate System (IS-IS), and BGP.
- `iso`—International Organization for Standardization (ISO). Configured on the logical interface for IS-IS traffic.
- `mlfr-uni-nni`—Multilink Frame Relay (MLFR) FRF.16 user-to-network network-to-network (UNI NNI). Configured on the logical interface for link services bundling.
- `mlfr-end-to-end`—Multilink Frame Relay end-to-end. Configured on the logical interface for multilink bundling.
- `mlppp`—Multilink Point-to-Point Protocol (MLPPP). Configured on the logical interface for multilink bundling.
- `mpls`—Multiprotocol Label Switching (MPLS). Configured on the logical interface for participation in an MPLS path.

- `pppoe`— Point-to-Point Protocol over Ethernet (PPPoE). Configured on Ethernet interfaces enabled to support multiple protocol families.
- `tcc`—Translational cross-connect (TCC). Configured on the logical interface of TCC physical interfaces.
- `tnp`—Trivial Network Protocol (TNP). Used to communicate between the Routing Engine and the router's packet forwarding components. The Junos OS automatically configures this protocol family on the router's internal interfaces only.
- `vpls`—Virtual private LAN service (VPLS). Configured on the logical interface on which you configure VPLS.

RPF Failures Field

For the logical interface, the `RPF Failures` field provides information about the amount of incoming traffic (in packets and bytes) that failed a unicast reverse path forwarding (RPF) check on a particular interface. The format is `RPF Failures: Packets: xx,Bytes: yy`. For example:

```
RPF Failures: Packets: 0, Bytes:0
```

Source Class Field

For the logical interface, the `Source class` field provides the names of source class usage (SCU) counters per family and per class for a particular interface. The counters display packets and bytes arriving from designated user-selected prefixes. For example:

```
                                                Packets
Bytes
Source class                    (packet-per-second)                (bits-per-second)


                                 gold             1928095                      161959980
                                           (               889)               (          5977
62)
                               bronze                       0
    0
                                           (                  0)                (
```

```
        0)
                                silver                          0
      0
                                          (                    0)                (
        0)
```

Junos CLI Reference Overview

We've consolidated all Junos CLI commands and configuration statements in one place. Learn about the syntax and options that make up the statements and commands and understand the contexts in which you'll use these CLI elements in your network configurations and operations.

- Junos CLI Reference

Click the links to access Junos OS and Junos OS Evolved configuration statement and command summary topics.

- Configuration Statements
- Operational Commands