

# Interfaces User Guide for Switches

Published  
2025-07-09

Juniper Networks, Inc.  
1133 Innovation Way  
Sunnyvale, California 94089  
USA  
408-745-2000  
[www.juniper.net](http://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 | xi

## 1

### Gigabit Ethernet Interfaces for Switches

#### Gigabit Ethernet Interfaces | 2

Speed on Gigabit Ethernet Interfaces | 2

Configure Interface Speed on Switches | 2

Configure Speed on QFX5100-48T Switches | 2

Autonegotiation on Gigabit Ethernet Interfaces | 4

Autonegotiation Support for EX4300-48MP Switches | 4

Autonegotiation Support for EX4400 Switches | 5

Autonegotiation Support for EX4100 Switches | 6

Autonegotiation Support for EX4600-40F, QFX5110-48S and QFX5100-48S with JNP-SFPP-10GE-T Transceiver | 6

Autonegotiation Support for QFX5120-48Y with JNP-SFPP-10GE-T Transceiver | 8

Autonegotiation on QFX5100-48T Switches | 9

#### Configure Gigabit and 10-Gigabit Ethernet Interfaces | 11

Configure Gigabit Ethernet Interfaces for EX Series Switches with ELS Support | 16

Configure VLAN Options and Interface Mode | 16

Configure the Link Settings | 17

Configure the IP Options | 20

## 2

### Aggregated Ethernet Interfaces for Switches

#### Aggregated Ethernet Interfaces Overview | 22

Overview | 22

Example: Configure Link Aggregation Between a QFX Series Switches and an Aggregation Switch | 27

Requirements | 27

Overview and Topology | 27

Configuration | 28

Verification | 32

Troubleshooting | 33

## **Configure Aggregated Ethernet Interfaces | 34**

### **Aggregated Ethernet LACP for Switches | 38**

Force LAG Links or Interfaces with Limited LACP Capability to Be Up | 39

Configure Aggregated Ethernet LACP (CLI Procedure) | 39

Configure LACP Link Protection of Aggregated Ethernet Interfaces for Switches | 41

Configure LACP Link Protection for a Single Link at the Global Level | 43

Configure LACP Link Protection for a Single Link at the Aggregated Interface Level | 44

Configure Subgroup Bundles to Provide LACP Link Protection to Multiple Links in an Aggregated Ethernet Interface | 44

Verify That LACP Is Configured Correctly and Bundle Members Are Exchanging LACP Protocol Packets | 47

Verify the LACP Setup | 47

Verify That LACP Packets Are Being Exchanged | 48

Example: Configuring Aggregated Ethernet High-Speed Uplinks with LACP Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch | 49

Requirements | 50

Overview and Topology | 50

Configuring LACP for the LAGs on the Virtual Chassis Access Switch | 51

Configuring LACP for the LAGs on the Virtual Chassis Distribution Switch | 52

Verification | 54

Troubleshooting | 56

Example: Configuring Link Aggregation with LACP Between a QFX Series Product and an Aggregation Switch | 57

Requirements | 57

Overview and Topology | 58

Configuring LACP for the LAG on the QFX Series | 58

Verification | 60

Troubleshooting | 62

### **Aggregated Ethernet Link Protection | 63**

Configuring Link Protection for Aggregated Ethernet Interfaces | 64

Configuring Primary and Backup Links for Link Aggregated Ethernet Interfaces | 64

Reverting Traffic to a Primary Link When Traffic is Passing Through a Backup Link | 65

Disabling Link Protection for Aggregated Ethernet Interfaces | 66

## Local Link Bias | 66

Local Link Bias Overview | 66

Configure Local Link Bias | 68

Local Minimum Links Overview | 69

## Load Balancing for Aggregated Ethernet Interfaces

### Load Balancing Overview | 74

Load Balancing and Ethernet Link Aggregation Overview | 74

Configure Load Balancing Based on MAC Addresses | 75

Configure Load Balancing on a LAG Link | 77

Example: Configuring Load Balancing on a LAG Link | 78

Understand Multicast Load Balancing on Aggregated 10-Gigabit Links for Routed Multicast Traffic on EX8200 Switches | 78

Example: Configure Multicast Load Balancing for Use with Aggregated 10-Gigabit Ethernet Interfaces on EX8200 Switches | 83

Requirements | 84

Overview and Topology | 84

Configuration | 86

Verification | 89

### Dynamic Load Balancing (DLB) | 91

Dynamic Load Balancing Overview | 91

Configuring Dynamic Load Balancing | 94

Configure DLB for ECMP (Flowlet mode) | 94

Configure DLB for LAG (Flowlet mode) | 95

Example: Configure Dynamic Load Balancing | 95

Requirements | 96

Overview | 96

Configuration | 97

Verification | 102

Selective Dynamic Load Balancing (DLB) | 104

Selective DLB Overview | 104

Selective DLB in AI-ML Data Centers | 105

Configuration | 105

Example: Selectively Enable DLB with a Firewall Filter Match Condition | 108

Customize Egress Port Link Quality Metrics for DLB | 110

Overview | 110

Configuration | 111

Configure Flowset Table Size in DLB Flowlet Mode | 112

Overview | 112

Configuration | 113

Reactive Path Rebalancing | 114

Overview | 114

Configuration | 115

## Hashing Algorithms for LAG and ECMP | 119

Understand the Algorithm Used to Hash LAG Bundle and Egress Next-Hop ECMP Traffic | 119

Configure the Fields in the Algorithm Used To Hash LAG Bundle and ECMP Traffic | 128

Configure the Hashing Algorithm to Use Fields in the Layer 2 Header for Hashing | 128

Configure the Hashing Algorithm to Use Fields in the IP Payload for Hashing | 129

Configure the Hashing Algorithm to Use Fields in the IPv6 Payload for Hashing | 130

Configure Other Hashing Parameters | 130

Example: Configure Link Aggregation Between a QFX Series Switches and an Aggregation Switch | 131

Requirements | 132

Overview and Topology | 132

Configuration | 133

Verification | 136

Troubleshooting | 138

Resilient Hashing on LAGs and ECMP groups | 139

Understand the Use of Resilient Hashing to Minimize Flow Remapping in LAGs/ECMP Groups | 139

Configure Resilient Hashing for LAGs/ECMP Groups | 142

## Global Load Balancing (GLB) | 143

GLB Overview | 144

GLB in AI-ML Data Centers | 145

Configure GLB | 145

Considerations | 145

Configure GLB | 146

## 4

### Energy Efficient Ethernet Interfaces for Switches

Energy Efficient Ethernet Interfaces | 149

Benefits of Energy Efficient Ethernet Interfaces | 149

Reduce Power Consumption on Interfaces Using Energy Efficient Ethernet | 149

Configure Energy Efficient Ethernet on Interfaces | 150

Enable EEE on an EEE-Capable Base-T Copper Ethernet Port | 151

Disable EEE on a Base-T Copper Ethernet Port | 151

Verify EEE-Enabled Ports | 151

## 5

### Switching Interface Features

Targeted Broadcast | 154

Overview | 154

Understand Targeted Broadcast | 155

Configure Targeted Broadcast | 157

Configure Targeted Broadcast | 157

Display Targeted Broadcast Configuration Options | 159

Configure Targeted Broadcast (CLI Procedure) | 161

Example: Configure Targeted Broadcast on a Switch | 162

Requirements | 163

Overview and Topology | 163

Verify IP Directed Broadcast Status | 165

Uplink Failure Detection | 165

Overview of Uplink Failure Detection | 166

Configure Interfaces for Uplink Failure Detection | 168

Example: Configure Interfaces for Uplink Failure Detection | 170

Requirements | 170

Overview and Topology | 170

Configure Uplink Failure Detection on both Switches | 172

Verification | 175

Verify That Uplink Failure Detection Is Working Correctly | 176

## Generic Routing Encapsulation (GRE) | 178

Understand GRE | 178

Configure Generic Routing Encapsulation (GRE) Tunneling | 182

Configure a GRE Tunnel | 183

Verify That Generic Routing Encapsulation Tunneling Is Working Correctly | 184

6

## Optical Transceivers for Switches

Optical Transceivers | 187

400ZR Optics Support on QFX5220-32CD and QFX5130 Switches | 190

7

## Port Speed for Switches

Port Speed Overview | 194

Configure Port Speed at Chassis Level and Interface Level | 196

Port Speed on EX Series Switches | 200

Channelizing Interfaces 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 EX4100-H-24MP and EX4100-H-24F Switches | 231

Port Speed on EX4000-12MP, EX4000-24MP, and EX4000-48MP Switches | 234

Operating Speed of Interfaces on EX Switches | 236

Port Speed on QFX Series Switches | 240

Port Speed on QFX5100-24Q Switches | 241



Port Speed on QFX5110-48S Switches	242
Port Speed on QFX5120-32C Switches	243
Port Speed on QFX5120-48T Switches	244
Port Speed on QFX5120-48Y Switches	245
Port Speed on QFX5120-48YM Switches	246
Port Speed on QFX5130-32CD Switches	247
Port Speed on QFX5130-48C/QFX5130-48CM Switches	249
Port Speed on QFX5200-32C Switches	256
Port Speed on QFX5210-64C Switches	256
Port Speed on QFX5230-64CD Switches	257
Port Speed on QFX5240 Switches	261
Port Speed on QFX5700 Switches	265

## 8

## Monitor Interfaces

Monitor Interface Status and Traffic	267
Monitor System Process Information	268
Monitor System Properties	269
Monitor Statistics for a Fast Ethernet or Gigabit Ethernet Interface	272
Trace Operations of the Interface Process	274

## 9

## Troubleshoot Interfaces

Troubleshoot Network Interfaces	278
Statistics for logical interfaces on Layer 2 interfaces are not accurate	278
The interface on the port in which an SFP or SFP+ transceiver is installed in an SFP or SFP+ module is down	279
Diagnose a Faulty Twisted-Pair Cable (CLI Procedure)	280
Platform-Specific Time Domain Reflectometry (TDR) Behavior	283
Troubleshoot Uplink Ports on EX2300 Switches	284

Speeds 10-Mbps and 100-Mbps Not Supported on Uplink Ports 4 and 5 on EX2300-48MP Switches | 284

## **Troubleshoot an Aggregated Ethernet Interface | 286**

Show Interfaces Command Shows the LAG Is Down | 286

Logical Interface Statistics Do Not Reflect all Traffic | 287

IPv6 Interface Traffic Statistics Are Not Supported | 287

SNMP Counters ifHCInBroadcastPkts and ifInBroadcastPkts Are Always 0 | 288

## **Configuration Statements and Operational Commands**

Common Output Fields Description | 290

Junos CLI Reference Overview | 300

# About This Guide

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

- **Switches Interface Basics:** Learn about [Switch Interfaces](#), [Physical Interface Properties](#), [Logical Interface Properties](#), [Damping Interfaces](#), and [Loopback Interfaces](#) in the Interfaces Fundamentals Guide for Junos OS.
- **Configure Ethernet Interfaces:** [Gigabit Ethernet Interfaces](#), [Aggregated Ethernet Interfaces Overview](#), and [Energy Efficient Ethernet Interfaces](#).
- **Switching Interface Features:** Find out how to optimize Switching Interfaces configuration and administration. See [Switching Interface Features](#).
- **Optical Transceivers for Switches:** See [Optical Transceivers for Switches](#).
- **Port Speed on a Switch or Line Card:** Understand channelization support and the port speed configuration. See [Port Speed Overview](#).
- **Learn how to Monitor and Troubleshoot Interfaces:** See [Monitoring and Troubleshooting Interfaces](#).

# 1

CHAPTER

## Gigabit Ethernet Interfaces for Switches

---

### IN THIS CHAPTER

- Gigabit Ethernet Interfaces | 2
  - Configure Gigabit and 10-Gigabit Ethernet Interfaces | 11
-

# Gigabit Ethernet Interfaces

## SUMMARY

Learn about speed and autonegotiation on Gigabit Ethernet interfaces, and how to configure speed and autonegotiation.

## IN THIS SECTION

- [Speed on Gigabit Ethernet Interfaces | 2](#)
- [Autonegotiation on Gigabit Ethernet Interfaces | 4](#)

## Speed on Gigabit Ethernet Interfaces

### IN THIS SECTION

- [Configure Interface Speed on Switches | 2](#)
- [Configure Speed on QFX5100-48T Switches | 2](#)

### Configure Interface Speed on Switches

On 1/10GbE capable SFP interfaces, the duplex is always full. Also, the speed matches the inserted optic. These interfaces support either 1GbE or 10GbE SFP optics. See [Configure Speed at Interfaces Level](#).

### Configure Speed on QFX5100-48T Switches

See [Configure Port Speed at Chassis Level and Interface Level](#) to configure speed.

For information about speed support, see *speed*.

[Table 1 on page 3](#) provides QFX5100-48T details and description.

**Table 1: QFX5100-48T Details and Description**

Detail	Description
Duplex Mode	Full duplex

Following are guidelines for configuring speed 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 have 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.

**Table 2: Configure Speed**

Configure Speed	Use Configuration
To configure a particular speed, mention the speed.	<p>For a port to only advertise a specific speed, start with a specific speed. It is mandatory that you:</p> <ul style="list-style-type: none"> <li>• Enable the autonegotiation option.</li> <li>• Configure the interface with a specific supported speed.</li> </ul> <pre>set interfaces xe-0/0/0 ether-options auto-negotiation set interfaces xe-0/0/0 speed <i>speed</i></pre> <p>For example to configure 1-Gbps speed, execute the following command:</p> <pre>set interfaces xe-0/0/0 ether-options auto-negotiation set interfaces xe-0/0/0 speed 1g</pre>

## Autonegotiation on Gigabit Ethernet Interfaces

### IN THIS SECTION

- [Autonegotiation Support for EX4300-48MP Switches | 4](#)
- [Autonegotiation Support for EX4400 Switches | 5](#)
- [Autonegotiation Support for EX4100 Switches | 6](#)
- [Autonegotiation Support for EX4600-40F, QFX5110-48S and QFX5100-48S with JNP-SFPP-10GE-T Transceiver | 6](#)
- [Autonegotiation Support for QFX5120-48Y with JNP-SFPP-10GE-T Transceiver | 8](#)
- [Autonegotiation on QFX5100-48T Switches | 9](#)

### 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 GbE uplink module port is solid green (instead of blinking green) because of a device limitation. However, device functionality does not change.

Table 3 summarizes the autonegotiation and half duplex support on EX4300-48MP switches.

**Table 3: 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 Gbps/5 Gbps/ 2.5 Gbps/1 G/100 Mbps)	No

**Table 3: Autonegotiation and Half-Duplex Support for EX4300-48MP Switches (Continued)**

Port Numbers	Interface Name	Autonegotiation Supported (YES/NO)?	Half Duplex Supported (YES/NO)?
PIC 2 PORT 0 – 3 (Uplink ports)	xe	No	No
PIC 0 PORTS 0 – 23	ge	1 G/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 G SFP and xe for 10 GbE SFP.	No
PIC 2 PORT 0, 1	Uplink 2x40G	No	No

## Autonegotiation Support for EX4400 Switches

**Table 4: Autonegotiation and Half-Duplex Support for EX4400 Switches**

Interface Name	Autonegotiation Supported (YES/NO)?	Half Duplex Supported (YES/NO)?
xe	No	No
ge	1 G/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.



**Table 4: Autonegotiation and Half-Duplex Support for EX4400 Switches (Continued)**

Interface Name	Autonegotiation Supported (YES/NO)?	Half Duplex Supported (YES/NO)?
mge (PIC 0)	Yes. Speed supported (10 Gbps/ 5 Gbps/2.5 Gbps/1 Gbps/ 100 Mbps)	No
2x40G, 2x100G (PIC 1)	No	No
Uplink 1x100G, 4x10G, and 4x25G (PIC 2)	No. Based on inserted transceiver, port is ge for 1 G SFP and xe for 10 GbE SFP.	No

## Autonegotiation Support for EX4100 Switches

**Table 5: Autonegotiation and Half-Duplex Support for EX4100 Switches**

Interface Name	Autonegotiation Supported (YES/NO)?	Half Duplex Supported (YES/NO)?
mge	Yes. Speed supported (10 Gbps/ 5 Gbps/2.5 Gbps/1 Gbps/ 100 Mbps)	No
xe	No	No
ge	1 Gbps/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 the JNP-SFPP-10GE-T transceivers come up based on the speed (100 Mbps, or 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 6 on page 7](#) discusses EX4600-40F, QFX5110-48S and QFX5100-48S switches with JNP-SFPP-10GE-T transceiver details.

**Table 6: 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	<p>Interface created - ge interface.</p> <p>On EX4600-40F, QFX5110-48S and QFX5100-48S, the ge interface supports 100 Mbps, 1 Gbps, and 10 Gbps speeds, which can be configured using the speed configuration statement.</p> <p>Use the set interfaces ge-0/0/0 speed (100M 1G 10G) command to configure the speed and autonegotiation.</p>
EX4600-40F, QFX5110-48S and QFX5100-48S switches with other transceivers	Interface created - ge or the xe.
Duplex Mode	Full duplex
Viewing Media Specific Information	<p>Execute the show interfaces media command to view the media-specific information. In the output of show interfaces <i>name</i> media, the output field speed displays the speed that is configured for the mge interface (with a default of 10 Gbps). The configured speed signifies the highest speed that the JNP-SFPP-10GE-T transceiver is capable of working at. You should enable autonegotiation for the transceiver unless it works in 100 Mbps. In 100 Mbps speed, the transceiver can use the parallel detect capability. This capability enables the transceiver to 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.</p>

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 7 on page 8](#):

**Table 7: Interface Speed Based on Configured Speed and Remote End Speed**

Configured Speed	The remote end speed of the interface (in the up state)
10G	10G, 1G, and 100M
1G	1G and 100M
100M	100M

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

For information about platforms support, see [Hardware Compatibility Tool](#).

[Table 8 on page 8](#) discusses QFX5120-48Y with JNP-SFPP-10GE-T transceiver details.

**Table 8: QFX5120-48Y with JNP-SFPP-10GE-T Transceiver Details and Description**

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

[Table 9 on page 9](#) configure 1-Gbps and 10-Gbps speeds on QFX5120-48Y with JNP-SFPP-10GE-T transceiver.

**Table 9: Configure and Delete 1-Gbps Speed**

Configure Speed.	Description
1 Gbps	<p>Use the <code>set chassis fpc 0 pic 0 port <i>port-number</i> speed 1G</code> command. Due to hardware limitations, you can configure the <i>port-number</i> 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.</p> <p>On QFX5120 switch, mge interfaces are not supported due to hardware limitations.</p>
To revert to 10-Gbps speed (after setting 1 Gbps speed.	Delete the 1G speed configuration.

## Autonegotiation on QFX5100-48T Switches

**Table 10: QFX5100-48T Details and Description**

Details	Description
Duplex Mode	Full duplex
Autonegotiation	<p>The autonegotiation option is to negotiate the speeds.</p> <p>10 Gbps, 1 Gbps, 100 Mbps-By default, autonegotiation is enabled.</p>

Following are guidelines for autonegotiation on QFX5100-48T switches:

- 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: Configure Speed and Autonegotiation

Configure Speed/Autonegotiation	Use Configuration
<p>To configure a particular speed, mention the speed.</p>	<p>For a port to only advertise a specific speed, start with a specific speed.</p> <p>It is mandatory:</p> <ul style="list-style-type: none"> <li>• to enable the autonegotiation option</li> <li>• to configure the interface with a specific speed.</li> </ul> <pre>set interfaces xe-0/0/0 ether-options auto-negotiation set interfaces xe-0/0/0 speed <i>speed</i></pre> <p>For example to configure 1-Gbps speed, execute the following command:</p> <pre>set interfaces xe-0/0/0 ether-options auto-negotiation set interfaces xe-0/0/0 speed 1g</pre>
<p>To enable autonegotiation and advertise all speeds.</p>	<p>With or without below, QFX5100-48T interface support autonegotiation to one of either 10 Gbps and 1 Gbps.</p> <pre>set interfaces xe-0/0/0 ether-options auto-negotiation set interfaces xe-0/0/0 speed auto</pre> <p>This configuration does not change any functionality. If the speed is set to 10 Gbps, the interface still operates as auto, and advertises 10 Gbps/1 Gbps/100 Mbps.</p> <p>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.</p>

# Configure Gigabit and 10-Gigabit Ethernet Interfaces

## IN THIS SECTION

- [Configure Gigabit Ethernet Interfaces for EX Series Switches with ELS Support | 16](#)

Learn how to configure link settings and IP options on Gigabit and 10-Gigabit Ethernet interfaces. Also, this topic includes how to configure Gigabit Ethernet interfaces with ELS support.

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 GbE and 10-Gigabit Ethernet interfaces.
- **autonegotiation**—Enable or disable autonegotiation of flow control, link mode, and speed for interfaces.
- **link-mode**—Specify **full-duplex**, **half-duplex**, or **automatic** for GbE interfaces.
- **loopback**—Enable or disable a lo0 for both GbE and 10 GbE interfaces.

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

[edit]

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

This topic describes:

## Configure the Link Settings

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 GbE and 10-Gigabit Ethernet interfaces.
- **autonegotiation**—Enable or disable autonegotiation of flow control, link mode, and speed for interfaces.
- **link-mode**—Specify **full-duplex**, **half-duplex**, or **automatic** for GbE interfaces.
- **loopback**—Enable or disable a lo0 for both GbE and 10 GbE interfaces.

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

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

### Link Settings Guidelines for 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.

Do not use the `set interfaces name ether-options no-auto-negotiate` 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.

## Link Setting Guidelines for QFX5100-48T Switches

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

- All the 10-GbE interfaces are set to **autonegotiation**.
- 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-GbE interfaces operate in full-duplex mode by default.
- Autonegotiation is enabled by default, and will autonegotiate the speed with the link partner. You cannot disable autonegotiation.

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. To connect to the link partner with 100-Mbps speed that supports autonegotiation at 100-Mbps, 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.



## Link Settings Guidelines for QFX5120-48T Switches

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

- The 10-Gigabit Ethernet interfaces are set to autonegotiation 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.

## Link Settings Guidelines for EX4100 and 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.

## Configure 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
```

## Configure Gigabit Ethernet Interfaces for EX Series Switches with ELS Support

### IN THIS SECTION

- [Configure VLAN Options and Interface Mode | 16](#)
- [Configure the Link Settings | 17](#)
- [Configure the IP Options | 20](#)

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 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](#).

This topic describes:

### Configure 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](#)

Configure the Link Settings

EX Series switches include a factory default configuration that enables interfaces with the link settings provided in [Table 12 on page 17](#).

Table 12: 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

**Table 12: Factory Default Configuration Link Settings for EX Series Switches (Continued)**

Ethernet Interface	Autonegotiation	Flow Control	Link Mode	Link Speed
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

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 EX4300 interface 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.

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.

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

```
[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 [Interface Naming Conventions](#).

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.
- `flow-control`—Enable or disable flow control.
- `link-mode`—Specify full duplex, half duplex, or autonegotiation.

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 EX4300 switches. 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 EX4300 link 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
```

- `loopback`—Enable or disable loopback mode.
- `speed`—Specify 10 Mbps, 100 Mbps, 1 Gbps, or autonegotiation.


## Configure 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.

### 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
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.

# 2

CHAPTER

## Aggregated Ethernet Interfaces for Switches

---

### IN THIS CHAPTER

- Aggregated Ethernet Interfaces Overview | 22
  - Configure Aggregated Ethernet Interfaces | 34
  - Aggregated Ethernet LACP for Switches | 38
  - Aggregated Ethernet Link Protection | 63
  - Local Link Bias | 66
-



# Aggregated Ethernet Interfaces Overview

## SUMMARY

Learn about aggregated Ethernet interfaces, LACP, and LAG.

## IN THIS SECTION

- [Overview | 22](#)
- [Example: Configure Link Aggregation Between a QFX Series Switches and an Aggregation Switch | 27](#)

## Overview

### IN THIS SECTION

- [Link Aggregation Group \(LAG\) | 23](#)
- [Link Aggregation Control Protocol \(LACP\) | 26](#)

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.

You can configure a mixed rate of link speeds for the aggregated Ethernet bundle. Link speeds of 10GbE, 40GbE, and 100GbE are supported. Load balancing does not work if you configure link speeds that are not supported.

Use [Aggregated Ethernet interface](#) to confirm platform and release support for specific features.

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.

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.

During a network Node group switchover, traffic might be dropped for a few seconds.

## Link Aggregation Group (LAG)

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 13 on page 23](#) 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.

For Junos OS Evolved, the software does not impose a limit on the maximum number of aex in a mixed-rate aggregated Ethernet bundle. All child logical interfaces belong to the same aggregated Ethernet physical interface and share the same selector. So, much less load-balance memory and mixed-rate aex configurations should go through even if number exceeds 64 logical interfaces.

**Table 13: Maximum Interfaces per LAG and Maximum LAGs per Switch (EX Series Switches)**

Switch	Maximum Interfaces per LAG	Maximum LAGs
EX4100-F Virtual Chassis	8	128
EX4200 and EX4200 Virtual Chassis	8	111
EX4300 and EX4300 Virtual Chassis	16	128

**Table 13: Maximum Interfaces per LAG and Maximum LAGs per Switch (EX Series Switches)**  
*(Continued)*

Switch	Maximum Interfaces per LAG	Maximum LAGs
EX4400	16	128
EX4600	32	128
EX4650 Virtual Chassis	64	72
EX6200	8	111
EX9200	64	150

**Table 14: 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

On QFX Series switches, if you try to commit a configuration containing more than 64 Ethernet interfaces in a LAG, you receive an error message. The error message says 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.

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.

- 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.

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 the physical link used for transmission. Thus, 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. This preserves 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.

## Example: Configure Link Aggregation Between a QFX Series Switches and an Aggregation Switch

### IN THIS SECTION

- [Requirements | 27](#)
- [Overview and Topology | 27](#)
- [Configuration | 28](#)
- [Verification | 32](#)
- [Troubleshooting | 33](#)

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.

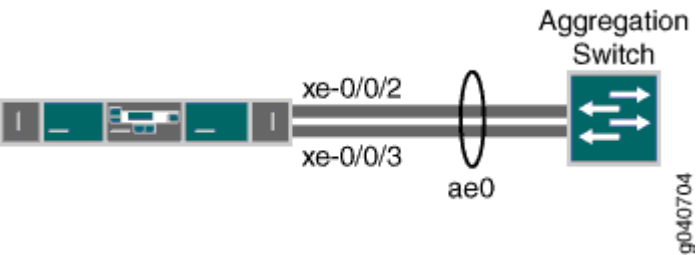


Table 15 on page 28 details the topology used in this configuration example.

Table 15: 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 | 29
- Results | 31

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

[Verify That LAG ae0.0 Has Been Created | 32](#)

[Verify That LAG ae0 Has Been Created | 33](#)

To verify that switching is operational and one LAG has been created, perform these tasks:

Verify 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.

Verify 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 | 33](#)

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

[Verify the Status of a LAG Interface](#)

*show lacp statistics interfaces (View)*

## Configure Aggregated Ethernet Interfaces

#### SUMMARY

Learn how to configure aggregated Ethernet interfaces. Includes a sample configuration as well.

To configure an aex:

1. Specify the number of aggregated Ethernet bundles you want on your device. If you specify the device-count value as 2, you can configure two aggregated bundles.

```
[edit chassis aggregated-devices ethernet]  
user@host# set device-count number
```

2. Specify that you want to configure the LAG interface.

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

3. Configure the aex.

```
[edit interfaces interface-name]
user@host# set ether-options 802.3ad aex
```

4. Specify the link speed for the aggregated Ethernet links. When you specify the speed, all the interfaces that make up the aggregated Ethernet bundle have the same speed. You can also configure the member links of an aggregated Ethernet bundle with mixed rates for efficient bandwidth utilization. See [link-speed \(Aggregated Ethernet\)](#).

```
[edit interfaces]
user@host# set aex aggregated-ether-options link-speed speed
```

5. Specify the minimum number of links for the aex—that is, the defined bundle— to be labeled *up*. By default, only one link must be up for the bundle to be labeled *up*.

```
[edit interfaces]
user@host# set aex aggregated-ether-options minimum-links number
```

You cannot use the minimum link in aggregated Ethernet with mixed speed. You cannot configure the minimum number of links and the minimum bandwidth at the same time.

6. (Optional) Specify the minimum bandwidth for the aggregated Ethernet links. You cannot configure link protection with minimum bandwidth.

```
[edit interfaces]
user@host# set aex aggregated-ether-options minimum-bandwidth
```

7. Configure tagged aggregated Ethernet. Specify the `vlan-tagging` statement at the `[edit interfaces aex]` hierarchy level.

```
[edit interfaces]
user@host# set aex vlan-tagging unit 0 vlan-id vlan-id
```

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]`

## 8. Configure untagged aggregated Ethernet.

```
[edit interfaces]
user@host# set aex unit 0 family inet address ip-address
```

- 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.

## 9. (Optional) Configure your device to collect multicast statistics for the aggregated Ethernet interface.

```
[edit interfaces]
user@host# set aex multicast-statistics
```

## 10. Verify and commit the configuration.

```
[edit interfaces]
user@host# run show configuration
user@host# commit
```

## 11. (Optional) Delete an aggregated Ethernet Interface.

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

Check the following guidelines while configuring aggregated Ethernet interfaces:

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, GbE IQ features and some newer GbE features are not supported in aggregated Ethernet bundles.

GbE 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.

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.

## Sample Aggregated Ethernet Interfaces Configuration

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;
    }
}
}

```

## Aggregated Ethernet LACP for Switches

### SUMMARY

Learn about aggregated Ethernet LACP, and how to configure LACP and LACP link protection.

### IN THIS SECTION

- [Force LAG Links or Interfaces with Limited LACP Capability to Be Up | 39](#)
- [Configure Aggregated Ethernet LACP \(CLI Procedure\) | 39](#)
- [Configure LACP Link Protection of Aggregated Ethernet Interfaces for Switches | 41](#)
- [Verify That LACP Is Configured Correctly and Bundle Members Are Exchanging LACP Protocol Packets | 47](#)
- [Example: Configuring Aggregated Ethernet High-Speed Uplinks with LACP Between an EX4200 Virtual Chassis Access Switch and an EX4200 Virtual Chassis Distribution Switch | 49](#)

- [Example: Configuring Link Aggregation with LACP Between a QFX Series Product and an Aggregation Switch | 57](#)

## Force LAG Links or Interfaces with Limited LACP Capability to Be Up

Ensure that the peer with limited LACP capability is active and accessible on the LAG network. To achieve this, configure one of the aggregated Ethernet links or interfaces on a PE device to be active. Use the appropriate hierarchy level on your device while configuring:

- `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):

- When an AE on a switch has multiple member links and one is in force-up state with its peer's LACP down, force-up is disabled if LACP partially comes up with a non-force-up member link. The member link is then ready for connection establishment through LACP. Force-up is eligible only if the server-side interface has LACP issues.

## Configure 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

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
```

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 \(CLI Procedure\)](#)

Configure LACP Link Protection of Aggregated Ethernet Interfaces for Switches

[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](#)

## Configure LACP Link Protection of Aggregated Ethernet Interfaces for Switches

### IN THIS SECTION

- [Configure LACP Link Protection for a Single Link at the Global Level | 43](#)
- [Configure LACP Link Protection for a Single Link at the Aggregated Interface Level | 44](#)

- [Configure Subgroup Bundles to Provide LACP Link Protection to Multiple Links in an Aggregated Ethernet Interface | 44](#)

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 \(CLI Procedure\)](#).

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.

## Configure 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
```

## Configure 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
```

## Configure 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.

Use [LACP Link Protection \(1:1\)](#) to confirm platform and release support for specific features.

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
```

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`

## Verify That LACP Is Configured Correctly and Bundle Members Are Exchanging LACP Protocol Packets

### IN THIS SECTION

- [Verify the LACP Setup | 47](#)
- [Verify That LACP Packets Are Being Exchanged | 48](#)

Verify that LACP has been set up correctly and that the bundle members are transmitting LACP protocol packets.

### Verify the LACP Setup

#### IN THIS SECTION

- [Purpose | 48](#)
- [Action | 48](#)
- [Meaning | 48](#)

**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.

**Verify That LACP Packets Are Being Exchanged**

IN THIS SECTION

[Purpose | 48](#)

[Action | 49](#)

[Meaning | 49](#)

**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 | 50](#)
- [Overview and Topology | 50](#)
- [Configuring LACP for the LAGs on the Virtual Chassis Access Switch | 51](#)
- [Configuring LACP for the LAGs on the Virtual Chassis Distribution Switch | 52](#)
- [Verification | 54](#)
- [Troubleshooting | 56](#)

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 | 51](#)

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 | 53](#)

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 | 54](#)

[Verifying That the LACP Packets Are Being Exchanged | 55](#)

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 aex 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 | 56](#)

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 | 57](#)
- [Overview and Topology | 58](#)
- [Configuring LACP for the LAG on the QFX Series | 58](#)
- [Verification | 60](#)
- [Troubleshooting | 62](#)

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 27:

### 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 | 59](#)

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 | 60](#)
- [Verifying That the LACP Packets Are Being Exchanged | 61](#)

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-0/0/2

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 aex 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 | 62](#)

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: Configure Link Aggregation Between a QFX Series Switches and an Aggregation Switch | 27](#)

---

*Example: Configuring an FCoE LAG on a Redundant Server Node Group*

---

*show lacp statistics interfaces (View)*

# Aggregated Ethernet Link Protection

## IN THIS SECTION

- [Configuring Link Protection for Aggregated Ethernet Interfaces | 64](#)
- [Configuring Primary and Backup Links for Link Aggregated Ethernet Interfaces | 64](#)
- [Reverting Traffic to a Primary Link When Traffic is Passing Through a Backup Link | 65](#)
- [Disabling Link Protection for Aggregated Ethernet Interfaces | 66](#)

Learn how to provide link protection for aggregated Ethernet Interfaces and configure the link protection for aggregated Ethernet interfaces. Learn also how to configure primary and backup links for aggregated Ethernet interfaces.

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	<ul style="list-style-type: none"> <li>ACX7000 Series routers that support link protection can operate in revertive mode. Enable auto-revertive operation using the following command:  set interfaces ae1 aggregated-ether-options link-protection revertive</li> </ul>

## 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
```

## 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 aex 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)*

## Local Link Bias

### SUMMARY

Learn about local link bias, local minimum links, and how to configure local link bias.

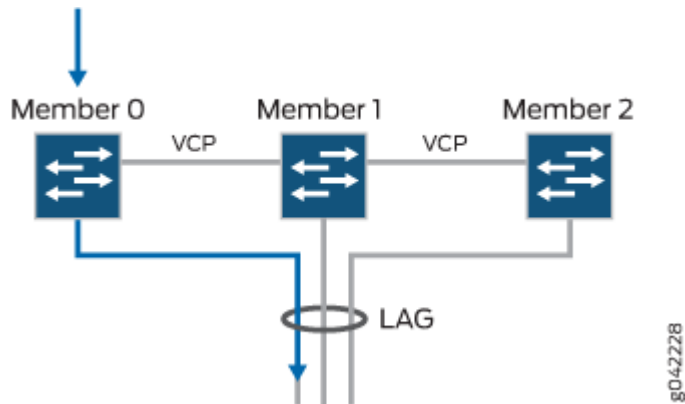
### IN THIS SECTION

- [Local Link Bias Overview | 66](#)
- [Configure Local Link Bias | 68](#)
- [Local Minimum Links Overview | 69](#)

## Local Link Bias Overview

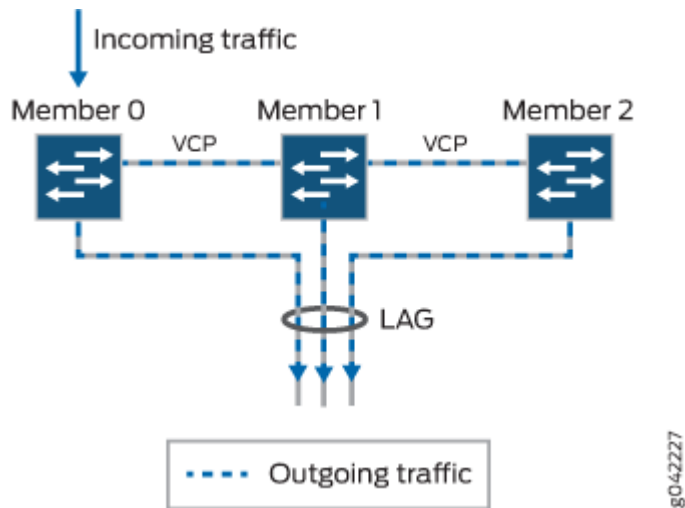
Local link bias conserves bandwidth on Virtual Chassis ports (VCPs) by using local links to forward unicast traffic exiting a Virtual Chassis or Virtual Chassis Fabric (VCF) that has a LAG bundle composed of member links on different member switches in the same Virtual Chassis or VCF. A local link is a member link in the LAG bundle that is on the member switch that received the traffic. Traffic is received and forwarded on the same member switch when local link bias is enabled. Hence, 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 1 on page 67](#).

**Figure 1: 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 2 on page 67](#).

**Figure 2: 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.

## Configure 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
```

## Local Minimum Links Overview

### IN THIS SECTION

- [Configuring Local Minimum Links | 71](#)
- [Local Minimum Links Effect on LAG Minimum Links | 71](#)
- [Local Minimum Links and Local Link Bias | 71](#)

A 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.

The local minimum links feature involves three components:

- Member links: Part of an aggregated Ethernet bundle (LAG)
- Member switches: Chassis in a Virtual Chassis or VCF
- Local member links (or local links): Member links of the same LAG local to a specific Virtual Chassis or VCF member switch.

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.

Use [LACP Minimum Link](#) to confirm platform and release support for specific features.

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 aex, 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 aex.
- 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.

Enable the local minimum links feature for a LAG. If one or more member links on a chassis fail, the feature compares the percentage of local member links still up to the threshold. If 'up' links fall below the threshold, the feature deactivates remaining active local links. No traffic for the aex will pass 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. Also, LAG traffic will continue to be distributed over available member links on that chassis.

For example, consider a member switch in a VCF 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. The feature also checks if 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 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 compare with the threshold.

A failed member link can force all local links down. When that link comes back up and increases the 'up' links percentage above the threshold, the system marks the forced-down links as up again.

Enable this feature only if your system manages ingress and egress traffic forwarding paths on LAGs for each chassis in a Virtual Chassis and VCFs. This is crucial when 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, 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
```

Enabling the feature on a LAG bundle applies to any chassis that has local member links in the LAG.

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. When local link bias is enabled, it favors forwarding traffic out of local links in the aggregated Ethernet bundle. If those links are down because the local minimum links threshold is not met, outgoing traffic redirects 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.

### 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
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.

# 3

CHAPTER

## Load Balancing for Aggregated Ethernet Interfaces

---

### IN THIS CHAPTER

- Load Balancing Overview | 74
  - Dynamic Load Balancing (DLB) | 91
  - Hashing Algorithms for LAG and ECMP | 119
  - Global Load Balancing (GLB) | 143
-

# Load Balancing Overview

## SUMMARY

Learn about load balancing on aggregated ethernet interfaces, and how to configure load balancing based on MAC addresses. reduces network congestion by dividing traffic among multiple interfaces. to

## IN THIS SECTION

- [Load Balancing and Ethernet Link Aggregation Overview | 74](#)
- [Configure Load Balancing Based on MAC Addresses | 75](#)
- [Configure Load Balancing on a LAG Link | 77](#)
- [Example: Configuring Load Balancing on a LAG Link | 78](#)
- [Understand Multicast Load Balancing on Aggregated 10-Gigabit Links for Routed Multicast Traffic on EX8200 Switches | 78](#)
- [Example: Configure Multicast Load Balancing for Use with Aggregated 10-Gigabit Ethernet Interfaces on EX8200 Switches | 83](#)

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 LAG for a group of Ethernet ports. L2 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 L2 payload by default. When the hash-mode is set to L2 payload, the hashing algorithm uses the IPv4 and IPv6 payload fields for hashing. You can also configure the load balancing hash key for L2 traffic to

use fields in the L3 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 77](#). In an L2 switch, one link is overutilized and other links are underutilized.

## Configure Load Balancing Based on MAC Addresses

### IN THIS SECTION

- [Platform-Specific MAC Address Based Load-Balancing Behavior | 76](#)

The hash key mechanism for load-balancing uses L2 MAC information such as frame source and destination address. To load-balance traffic based on L2 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 Explorer](#) to confirm platform and release support for specific features.

Review the ["Platform-Specific MAC Address Based Load-Balancing Behavior" on page 76](#) 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	<ul style="list-style-type: none"><li>• 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.</li></ul> <p>Note the following about hashing on ACX7000 Series Routers:</p> <ul style="list-style-type: none"><li>• 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.</li><li>• 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.</li><li>• Do not support weighted hashing.</li></ul>

**SEE ALSO**

| *multiservice*

## Configure Load Balancing on a LAG Link

You can configure the load balancing hash key for L2 traffic to use fields in the L3 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. You can also look at **layer-4** fields. You configure this statement at the `[edit forwarding-options hash-key family multiservice]` hierarchy level.

You can configure L3 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 L3 information carried in the packet.

For more information about LAG configuration, see the [Junos OS Network Interfaces Library for Routing Devices](#).

### Example:

This example configures the load-balancing hash key to use the source L3 IP address option and Layer 4 header fields. The example also includes the source and destination MAC addresses for load balancing on a LAG link.

```
[edit]
forwarding-options {
  hash-key {
    family multiservice {
      source-mac;
      destination-mac;
      payload {
        ip {
          layer-3 {
            source-ip-only;
          }
          layer-4;
        }
      }
    }
  }
}
```

Any change in the hash key configuration requires a reboot of the FPC for the changes to take effect.



## 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.

## Understand 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 | 79](#)
- [When Should I Use Multicast Load Balancing? | 81](#)

- [How Does Multicast Load Balancing Work? | 81](#)
- [How Do I Implement Multicast Load Balancing on an EX8200 Switch? | 82](#)

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. Streaming video technology was once limited to occasional company presentations. Multicasting now boosts the technology, enabling a constant stream of movies, real-time data, news clips, and amateur videos to flow 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 LAGs.

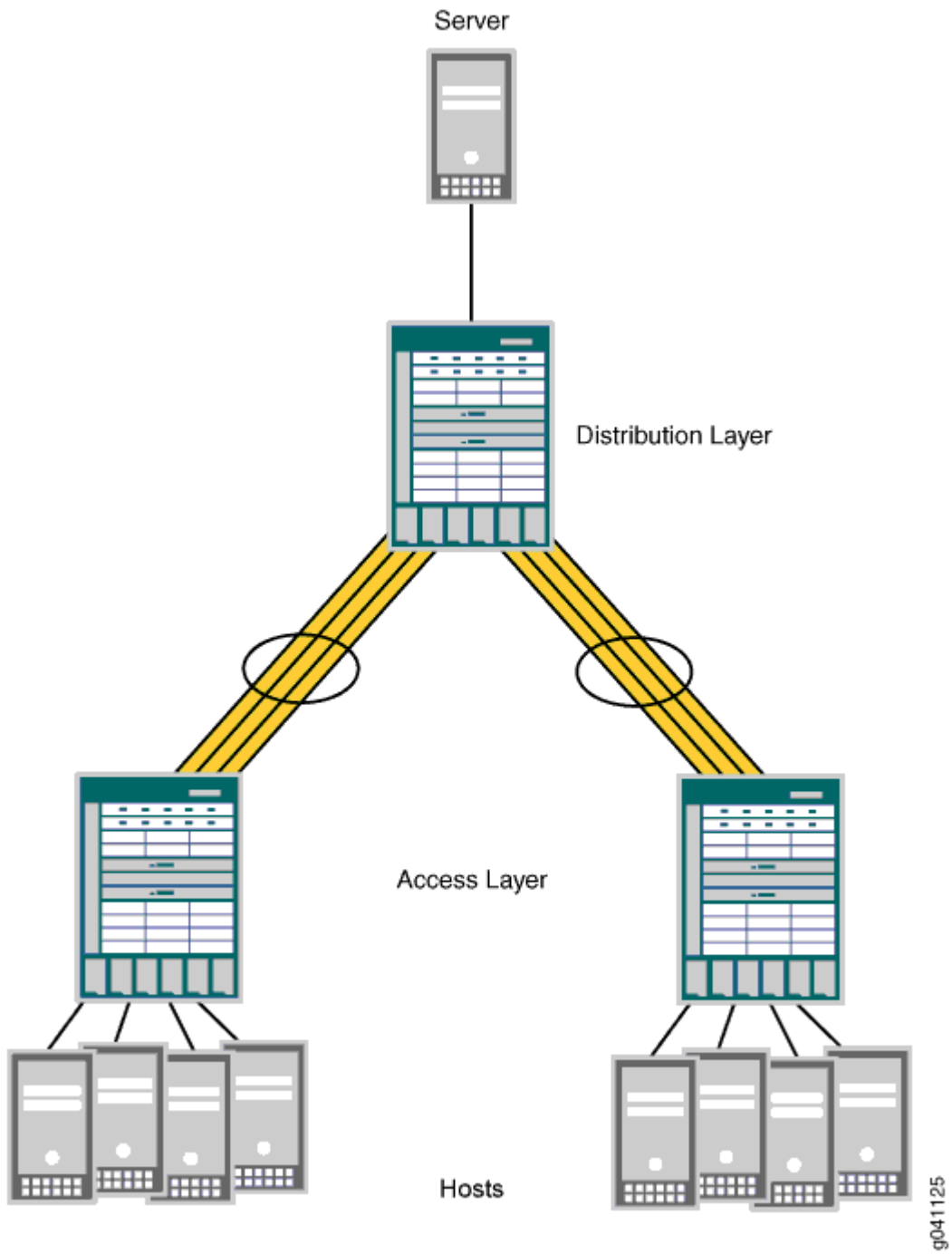
Multicast load balancing involves managing the individual links in each LAG to ensure that each link is used efficiently. Hashing algorithms continually evaluate the data stream, adjusting stream distribution over the links in the LAG, so that no link is underutilized or overutilized. Multicast load 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 3 on page 80](#), four 10-gigabit links have been aggregated to form each 40-gigabit link.

Figure 3: 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 16 on page 81](#) for more information.

**Table 16: Hashing Algorithms Used by Multicast Load Balancing**

Hashing Algorithms	Based On	Best Use
<code>crc-sgip</code>	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 16: 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: Configure Multicast Load Balancing for Use with Aggregated 10-Gigabit Ethernet Interfaces on EX8200 Switches

### IN THIS SECTION

- [Requirements | 84](#)
- [Overview and Topology | 84](#)
- [Configuration | 86](#)
- [Verification | 89](#)

EX8200 switches support multicast load balancing on LAGs. Multicast load balancing evenly distributes L3 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, provides 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 reenabling it.

An interface with an already configured IP address cannot form part of the LAG.

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

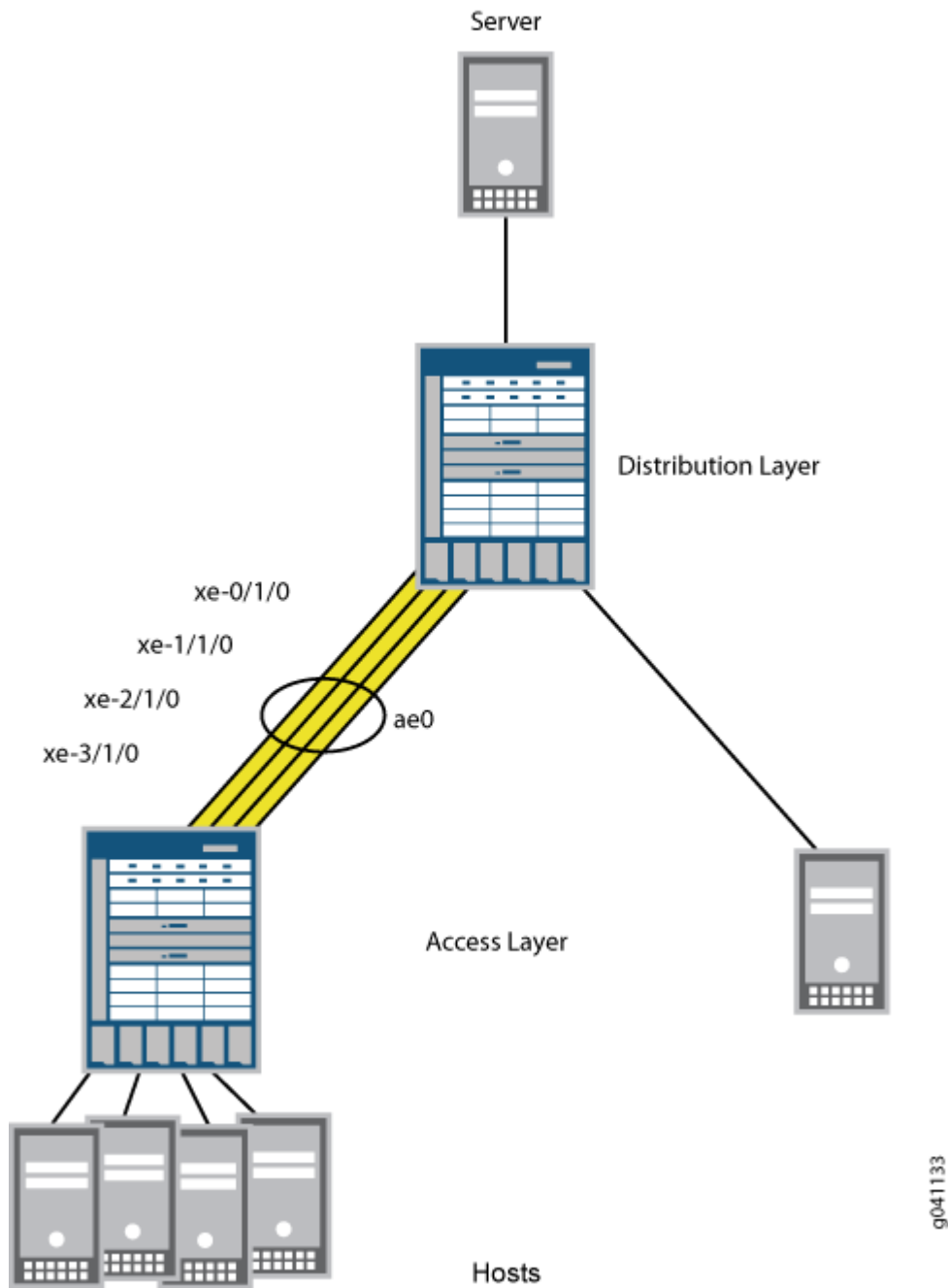
Multicast load balancing uses one of seven hashing algorithms to balance traffic between the individual 10-gigabit links in the LAG. For a description of the hashing algorithms, see *multicast-loadbalance*. The default hashing algorithm is *crc-sgip*. You can experiment with the different hashing algorithms until you determine the one that best balances your L3 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 4 on page 85](#), an EX8200 switch in the distribution layer is connected to an EX8200 switch in the access layer.

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 Gbps.

The default hashing algorithm, *crc-sgip*, involves a cyclic redundancy check (CRC) of both the multicast packet source and group IP addresses.

Figure 4: 40-Gigabit LAG Composed of Four 10-Gigabit Links



You will configure a LAG on each switch and reenenable multicast load balancing. When reenabled, multicast load balancing will automatically take effect on the LAG, and the speed is set to 10 Gbps for each link in the LAG. Link speed for the 40-gigabit LAG is automatically set to 40 Gbps.



## Configuration

### IN THIS SECTION

- [Procedure | 86](#)

## Procedure

### CLI Quick Configuration

```
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 (aex) to be created:

```
[edit chassis]
user@switch# set aggregated-devices ethernet device-count 1
```

2. Specify the minimum number of links for the aex, that is, the LAG, to be labeled up:

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
```

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 Gbps 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 L3 routed multicast traffic.

If you change the hashing algorithm when multicast load balancing is disabled, the new algorithm takes effect after you reenabling 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 | 89](#)
- [Verifying Multicast Load Balancing | 90](#)

To confirm that the configuration is working properly, perform these tasks:

### Verifying the Status of a LAG Interface

#### Purpose

Verify that a 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 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

Learn about Dynamic load balancing (DLB) and how to configure DLB. This topic also includes how to configure DLB for ECMP and LAG.

## IN THIS SECTION

- [Dynamic Load Balancing Overview | 91](#)
- [Configuring Dynamic Load Balancing | 94](#)
- [Example: Configure Dynamic Load Balancing | 95](#)
- [Selective Dynamic Load Balancing \(DLB\) | 104](#)
- [Customize Egress Port Link Quality Metrics for DLB | 110](#)
- [Configure Flowset Table Size in DLB Flowlet Mode | 112](#)
- [Reactive Path Rebalancing | 114](#)

## Dynamic Load Balancing Overview

### IN THIS SECTION

- [Benefits | 93](#)
- [DLB Modes | 93](#)

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.



## Configuring Dynamic Load Balancing

### IN THIS SECTION

- [Configure DLB for ECMP \(Flowlet mode\) | 94](#)
- [Configure DLB for LAG \(Flowlet mode\) | 95](#)

This topic describes how to configure dynamic load balancing (DLB) in flowlet mode.

### Configure 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 the *ether-type* statement:

```
[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.

## Configure DLB for LAG (Flowlet mode)

Before you begin, create an aggregated Ethernet bundle by configuring a set of router interfaces as aggregated Ethernet and with a specific aggregated Ethernet 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.

## Example: Configure Dynamic Load Balancing

### IN THIS SECTION

- [Requirements | 96](#)
- [Overview | 96](#)
- [Configuration | 97](#)
- [Verification | 102](#)

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 | 96](#)

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 EtherTypes such as IPv4, IPv6, and MPLS based on configuration. If you don't configure any EtherType, then DLB is applied to all EtherTypes. Note that you must explicitly configure the DLB mode because there is no default mode.



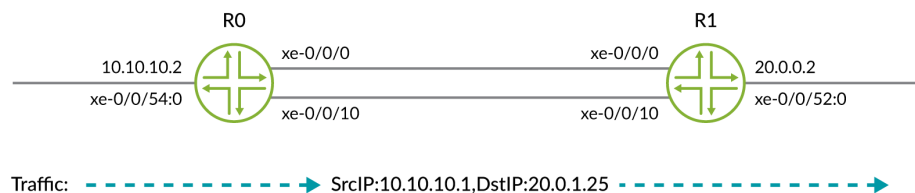
### NOTE:

- 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 5: Dynamic Load Balancing**



g300681



**NOTE:** This example shows static configuration. You can also add configuration with dynamic protocols.

## Configuration

### IN THIS SECTION

● [Verification](#) | 100

### 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.

**RO**

```
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 | 101](#)

- **Verify Traffic Load After Configuring Dynamic Load Balancing Feature on LAG | 101**

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 | 102](#)
- [Verify Load Balancing on R1 | 103](#)

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.

## Selective Dynamic Load Balancing (DLB)

### IN THIS SECTION

- [Selective DLB Overview | 104](#)
- [Selective DLB in AI-ML Data Centers | 105](#)
- [Configuration | 105](#)
- [Example: Selectively Enable DLB with a Firewall Filter Match Condition | 108](#)

### Selective DLB Overview

#### IN THIS SECTION

- [Benefits | 104](#)

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 | 105](#)
- [Topology | 106](#)
- [Disable DLB Globally and Selectively Enable DLB | 106](#)
- [Enable DLB Globally and Selectively Disable DLB | 107](#)

### 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 1 summarizes the load balancing configuration on the two flows shown and the results of the load balancing applied on the flows:

Figure 6: Per-Flow and Per-Packet Load Balancing

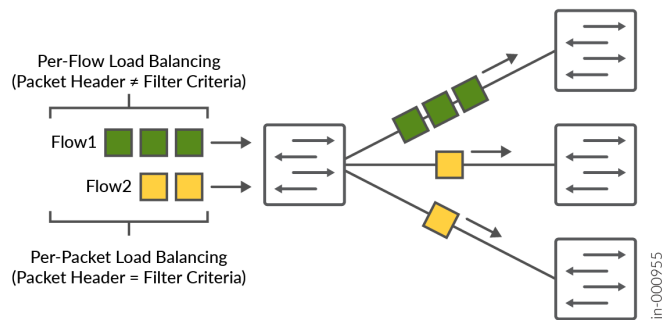


Table 17: Flow Behaviors

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.

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 | 110](#)
- [Configuration | 111](#)

## Overview

### IN THIS SECTION

- [Benefits | 111](#)

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 | 112
- Configuration | 113

## Overview

### IN THIS SECTION

- Benefits | 113

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:

$$32,768 / (\text{flowset size}) = \text{Number of ECMP groups}$$

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" on page 94](#).
  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 | 114](#)
- [Configuration | 115](#)

## Overview

### IN THIS SECTION

- [Benefits | 115](#)

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 | 115](#)
- [Topology | 116](#)
- [Configure Reactive Path Rebalancing | 117](#)

## 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.

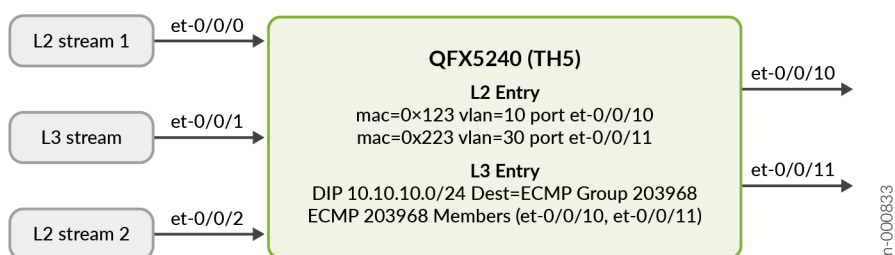
Be aware of the following when using this feature:

- 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 7: 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 stream 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" on page 94](#).
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

<i>dlb</i>
<i>dynamic-load-balance</i>
<i>egress-quantization</i>
<i>enhanced-hash-key</i>
<i>rdma-opcode</i>
<i>show forwarding-options enhanced-hash-key</i>

# Hashing Algorithms for LAG and ECMP

## SUMMARY

Learn about hashing algorithms used for LAG and ECMP , and how to configure the hashing algorithms.

## IN THIS SECTION

- [Understand the Algorithm Used to Hash LAG Bundle and Egress Next-Hop ECMP Traffic | 119](#)
- [Configure the Fields in the Algorithm Used To Hash LAG Bundle and ECMP Traffic | 128](#)
- [Example: Configure Link Aggregation Between a QFX Series Switches and an Aggregation Switch | 131](#)
- [Resilient Hashing on LAGs and ECMP groups | 139](#)

## Understand the Algorithm Used to Hash LAG Bundle and Egress Next-Hop ECMP Traffic

### IN THIS SECTION

- [Understand the Hashing Algorithm | 120](#)
- [IP \(IPv4 and IPv6\) | 121](#)
- [MPLS | 124](#)
- [MAC-in-MAC Packet Hashing | 125](#)
- [Layer 2 Header Hashing | 126](#)
- [Hashing Parameters | 127](#)

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.

This topic contains the following sections:

## Understand 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 18 on page 122](#) 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 18: IPv4 and IPv6 Hashing Fields

[illegible]



## 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. 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

Use [Feature Explorer](#) to confirm platform and release support for specific features.

[Table 19 on page 124](#) 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 19: 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	✓	✓	✓	✓	✓

Table 19: MPLS Hashing Fields (*Continued*)

Field	EX3400	EX4300		QFX5100	QFX5110 and QFX5120	QFX5200
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
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 20 on page 126](#).



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 20: 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 21: Layer 2 Header Hashing Fields**

Field	EX3400	EX4300	QFX5100	QFX5110 and QFX5120	QFX5200
Source MAC	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)	✓ (configurable)
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.

## Configure the Fields in the Algorithm Used To Hash LAG Bundle and ECMP Traffic

### IN THIS SECTION

- [Configure the Hashing Algorithm to Use Fields in the Layer 2 Header for Hashing | 128](#)
- [Configure the Hashing Algorithm to Use Fields in the IP Payload for Hashing | 129](#)
- [Configure the Hashing Algorithm to Use Fields in the IPv6 Payload for Hashing | 130](#)
- [Configure Other Hashing Parameters | 130](#)

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.

Use [Link aggregation group \(LAG\) bundle hashing configuration](#) to confirm platform and release support for specific features.

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.

### Configure 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
```

## Configure 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
```

## Configure 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
```

## Configure 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 \(QFX 10002 and QFX 10008 Switches\)](#)

## Example: Configure Link Aggregation Between a QFX Series Switches and an Aggregation Switch

### IN THIS SECTION

- [Requirements | 132](#)
- [Overview and Topology | 132](#)
- [Configuration | 133](#)
- [Verification | 136](#)
- [Troubleshooting | 138](#)

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.

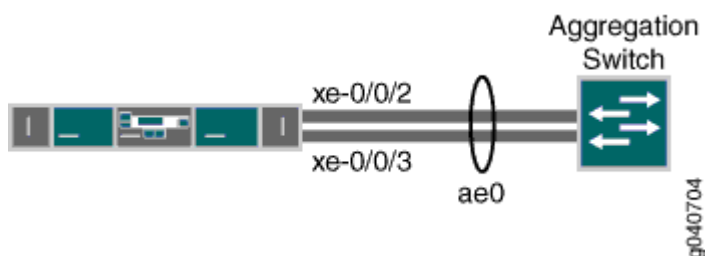


Table 22 on page 133 details the topology used in this configuration example.

Table 22: 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 | 133
- Results | 135

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

 [Verify That LAG ae0.0 Has Been Created](#) | 137

● [Verify That LAG ae0 Has Been Created | 137](#)

To verify that switching is operational and one LAG has been created, perform these tasks:

**Verify 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.

**Verify 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 | 138](#)

## 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

[Verify the Status of a LAG Interface](#)

*show lacp statistics interfaces (View)*

## Resilient Hashing on LAGs and ECMP groups

### IN THIS SECTION

- [Understand the Use of Resilient Hashing to Minimize Flow Remapping in LAGs/ECMP Groups | 139](#)
- [Configure Resilient Hashing for LAGs/ECMP Groups | 142](#)

Resilient hashing helps minimize the flow remapping across equal cost multipath (ECMP) groups and LAGs in a load-balanced system. The topics below discuss the working, usage and configuring of resilient hashing on link aggregation groups (LAGs) and ECMP groups.

### Understand the Use of Resilient Hashing to Minimize Flow Remapping in LAGs/ECMP Groups

#### IN THIS SECTION

- [Why You Might Want to Use Resilient Hashing and How It Works with Static Hashing | 139](#)
- [Limitations and Caveats for Resilient Hashing | 141](#)
- [Resilient Hashing on LAGs | 141](#)
- [Resilient Hashing on ECMP | 142](#)

You use resilient hashing to minimize flow remapping across members of a LAG/ECMP group in a load-balanced system. You can configure resilient hashing in LAG and in ECMP groups.

#### Why You Might Want to Use Resilient Hashing and How It Works with Static Hashing

Resilient hashing works with the default static hashing algorithm. When members are added to or deleted from a LAG/ECMP group, the static hashing algorithm might remap destination paths. With resilient hashing, the chances of a flow being remapped are minimal if its path is unaffected by the LAG/ECMP group's member change. When a flow is affected by a member change, the Packet Forwarding Engine rebalances the flow by reprogramming the FlowSet table.

Use [Resilient Hashing for Load Balancing](#) to confirm platform and release support for specific features.

Resilient hashing thus provides the following benefits:

- Minimizes traffic-distribution imbalances among members of a LAG/ECMP group when members are added to or deleted from the group.
- Minimizes the impact on flows bound to unaffected members when a new member is added or an existing member is deleted from the group.

In normal hash-based load balancing, with the static hashing algorithm used alone, flows are assigned to members through the mathematical mod (%) operation. Any increase or decrease in the number of group members results in a complete remapping of flows to member IDs, as shown in the following example:

- Member ID = Hash (key) mod (number of members in group)
- Example:
  - Hash (key) = 10
  - $10 \bmod 5 = 0$  (member with ID 0 is selected for flow)
  - $10 \bmod 4 = 2$  (member with ID 2 is selected for the same flow when the number of members is decreased by 1)

Resilient hashing minimizes the destination path remapping when a member in the LAG/ECMP group is added or deleted.

When the flow is affected by a member change in the group, resilient hashing rebalances the flow by reprogramming the FlowSet table.

**Table 23: Destination Path Results for Static Hashing and for Resilient Hashing When Members Are Added to or Deleted from LAGs**

LAG/ ECMP Group Size	Normal (Static) Hashing Result	Resilient Hashing Result	Notes
4	Hash(10) % 4 = 2 Flow is assigned to member ID 2.	Flow is assigned to one of four group members based on FlowSet table entries.	Original LAG/ECMP group size is 4.
3	Hash(10) % 3 = 1 Flow is assigned to member ID 1.	Flow is assigned to same member as in the previous case.	Delete one member from original LAG/ECMP group. LAG/ECMP group size is 3.

**Table 23: Destination Path Results for Static Hashing and for Resilient Hashing When Members Are Added to or Deleted from LAGs (*Continued*)**

LAG/ ECMP Group Size	Normal (Static) Hashing Result	Resilient Hashing Result	Notes
5	Hash(10) % 5 = 0 Flow is assigned to member ID 0.	There is minimal redistribution of flows from other members to this newly added member.	Add one member to original LAG group. LAG/ECMP group size is 5.

### Limitations and Caveats for Resilient Hashing

Notice the following limitation and caveats for the resilient hashing feature:

- Resilient hashing applies only to unicast traffic.
- Resilient hashing supports a maximum of 1024 LAGs, with each group having a maximum of 256 members.
- Resilient hashing does not guarantee that traffic distribution is even across all group members—it depends on the traffic pattern and on the organization of the resilient hashing FlowSet table in hardware. Resilient hashing *minimizes* remapping of flows to destination links when members are added to or deleted from the group.
- If resilient hashing is enabled on a LAG or ECMP group and if `set forwarding-options enhanced-hash-keyis` is used with one of the following options, some flows might change destination links. The reason is that the new hash parameters might generate new hash indexes for the flows.
  - `hash-mode`
  - `inet`
  - `inet6`
  - `layer2`
- Resilient hashing is not supported on VCP links.

### Resilient Hashing on LAGs

A LAG combines Ethernet interfaces (members) to form a logical point-to-point link that increases bandwidth, provides reliability, and allows load balancing. Resilient hashing minimizes destination remapping behavior when a new member is added or deleted from the LAG.



A resilient hashing configuration on LAGs is per-aggregated-Ethernet-interface-based.

## Resilient Hashing on ECMP

An ECMP group for a route contains multiple next-hop equal cost addresses for the same destination in the routing table. Routes of equal cost have the same preference and metric values.

Junos OS uses the static hashing algorithm to choose one of the next-hop addresses in the ECMP group to install in the forwarding table. Resilient hashing enhances ECMPs by minimizing destination remapping behavior when a new member is added or deleted from the ECMP group.

A resilient hashing configuration on ECMP is global—it applies to all ECMP groups.

## Configure Resilient Hashing for LAGs/ECMP Groups

### IN THIS SECTION

- [Configure Resilient Hashing on LAGs | 142](#)
- [Configure Resilient Hashing on ECMP Groups | 143](#)

You use resilient hashing to minimize flow remapping across members of a LAG/ECMP group in a load-balanced system. You can configure resilient hashing in LAGs and ECMP sets.

This topic includes:

### Configure Resilient Hashing on LAGs

To enable resilient hashing for a LAG:

- Configure resilient hashing on the aggregated Ethernet interface:

```
[edit interfaces]
user@switch# set aex aggregated-ether-options resilient-hash
```

- (Optional) Configure a specific value for the resilient-hash seed. This value will apply only to the HASH2 engine:

```
[edit]
user@switch# set forwarding-options enhanced-hash-key resilient-hash-seed seed-value
```

Configure Resilient Hashing on ECMP Groups

To enable resilient hashing for ECMP groups:

- Configure resilient hashing for ECMP:

```
[edit forwarding-options]
user@switch# set enhanced-hash-key ecmp-resilient-hash
```

When resilient hashing is added or removed, the traffic distribution across all members of an ECMP group for a given flow are reprogrammed and, as a result, some flows might be remapped to new ECMP group members.

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	on the QFX5000 line of switches, you can change hashing parameters for the existing algorithms implemented.

# Global Load Balancing (GLB)

SUMMARY

Learn about GLB and how to configure GLB.

IN THIS SECTION

- [GLB Overview | 144](#)
- [GLB in AI-ML Data Centers | 145](#)

## GLB Overview

### IN THIS SECTION

- [Benefits | 144](#)

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.

## Configure GLB

### IN THIS SECTION

- [Considerations | 145](#)
- [Configure GLB | 146](#)

### 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-options 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*

*global-load-balancing*

# 4

CHAPTER

## Energy Efficient Ethernet Interfaces for Switches

---

### IN THIS CHAPTER

- Energy Efficient Ethernet Interfaces | 149
  - Configure Energy Efficient Ethernet on Interfaces | 150
-

# Energy Efficient Ethernet Interfaces

## SUMMARY

Learn about Energy efficient Ethernet (EEE) interfaces, its benefits, and how it reduces power consumption on interfaces.

## IN THIS SECTION

- [Benefits of Energy Efficient Ethernet Interfaces | 149](#)
- [Reduce Power Consumption on Interfaces Using Energy Efficient Ethernet | 149](#)

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.

## Benefits of Energy Efficient Ethernet Interfaces

The benefits of interfaces on Juniper Networks switches include:

- **Reduced Power Consumption:** EEE helps in reducing power consumption on interfaces, which is crucial for maintaining operational efficiency and minimizing environmental impact.
- **Improved Network Performance:** By optimizing energy usage, EEE interfaces can help in maintaining consistent network performance, ensuring that the network runs at peak efficiency without compromising on speed or reliability.
- **Enhanced Sustainability:** Juniper Networks' focus on sustainability is reflected in the improved power efficiency of their EEE interfaces, contributing to a longer life cycle network architecture with power, space, and design optimization.

These benefits help in operational simplicity, making EEE interfaces a valuable feature for their switches.

## 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 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 physical layer devices to exchange LPI indications to signal the transition to low-power mode when traffic is nil. 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 physical layer devices 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 | 151](#)
- [Disable EEE on a Base-T Copper Ethernet Port | 151](#)
- [Verify EEE-Enabled Ports | 151](#)

Learn how to configure EEE on interfaces. This topic also includes how to enable EEE on an EEE-Capable Base-T Copper Ethernet Port, disable EEE on a Base-T Copper Ethernet Port, and verify EEE-enabled ports.

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 | 151
- Action | 152

### 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 using the `show chassis power-budget-statistics` command.

# 5

CHAPTER

## Switching Interface Features

---

### IN THIS CHAPTER

- Targeted Broadcast | 154
  - Uplink Failure Detection | 165
  - Generic Routing Encapsulation (GRE) | 178
-

# Targeted Broadcast

## SUMMARY

Learn about targeted broadcast and how to configure targeted broadcast.

## IN THIS SECTION

- [Overview | 154](#)
- [Understand Targeted Broadcast | 155](#)
- [Configure Targeted Broadcast | 157](#)
- [Configure Targeted Broadcast \(CLI Procedure\) | 161](#)
- [Example: Configure Targeted Broadcast on a Switch | 162](#)
- [Verify IP Directed Broadcast Status | 165](#)

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.

## Overview

Targeted broadcast is a process of flooding a target subnet with L3 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 VRF instances.

Regular L3 broadcast IP packets originating from a subnet are broadcast within the same subnet. When these IP packets reach a different subnet, the packets are forwarded to the Routing Engine (to be forwarded to other applications). Hence, 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.

L3 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 the packets 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. The forwarding is helpful when the router is flooded with packets to process, or to both an egress interface and the Routing Engine.

Any *firewall filter* that is configured on the Routing Engine lo0 cannot be applied to IP packets that are forwarded to the Routing Engine as a result of a targeted broadcast. The reason is broadcast packets are forwarded as flood next-hop traffic and not as local next-hop traffic. You can apply a firewall filter only to local next-hop routes for traffic directed toward the Routing Engine.

## Understand Targeted Broadcast

### IN THIS SECTION

- [Targeted Broadcast Overview | 156](#)
- [Targeted Broadcast Implementation | 156](#)
- [When to Enable Targeted Broadcast | 156](#)
- [When Not to Enable Targeted Broadcast | 157](#)

When packets reach the destination subnet and targeted broadcast is enabled on the receiving switch, the switch converts the targeted broadcast packet into a broadcast. The conversion floods the packet on the target subnet. All hosts on the target subnet receive the targeted broadcast packet.

This topic covers:

## Targeted Broadcast Overview

Targeted 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 a targeted broadcast is to flood the target subnet with the broadcast packets without broadcasting to the entire network. Targeted broadcast packets cannot originate from the target subnet.

When you send a targeted 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 targeted broadcast is enabled on the interface that is directly connected to the target subnet:

- If targeted 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 targeted broadcast is disabled on the interface that is directly connected to the target subnet, the switch drops the packet.

## Targeted Broadcast Implementation

You configure targeted broadcast on a per-subnet basis by enabling targeted broadcast on the L3 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, targeted broadcast is disabled.

## When to Enable Targeted Broadcast

Targeted broadcast is disabled by default. Enable targeted 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 targeted broadcast on a subnet affects only the hosts within that subnet. Only packets received on the subnet's L3 interface that have the subnet's broadcast IP address as the destination address is flooded on the subnet.

## When Not to Enable Targeted Broadcast

Typically, you do not enable targeted broadcast on subnets that have direct connections to the Internet. Disabling targeted broadcast on a subnet's L3 interface affects only that subnet. If you disable targeted 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 targeted broadcast on it increases the network's susceptibility to DoS attacks.

A malicious attacker can spoof a source IP address to deceive a network into identifying the attacker as legitimate. The attacker can then send targeted broadcasts with ICMP echo (ping) packets. When the hosts on the network with targeted broadcast enabled receive the ICMP echo packets, the hosts send replies to the victim that has the spoofed source IP address. The replies create a flood of ping replies in a DoS attack that can overwhelm the spoofed source address known as a *smurf* attack. Another common DoS attack on exposed networks with targeted broadcast enabled is a *fraggle* attack. The attack is similar to a smurf attack except that the malicious packet is a UDP echo packet instead of an ICMP echo packet.

## Configure Targeted Broadcast

### IN THIS SECTION

- [Configure Targeted Broadcast | 157](#)
- [Display Targeted Broadcast Configuration Options | 159](#)

## Configure Targeted Broadcast

You can configure targeted broadcast on an egress interface with different options.

Either of these configurations is acceptable:

- You can allow the IP broadcast packets destined for a Layer 3 address to be forwarded through the egress interface and to send a copy of the IP broadcast packets to the Routing Engine.
- You can allow the IP broadcast packets to be forwarded through 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 interface.

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

or

```
[edit]
user@host# set interfaces irb
```

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. Forward IP broadcast packets to a Layer 3 address:

- a. through the egress interface and send a copy of the same packets to the Routing Engine.

```
[edit interfaces interface-name unit interface-unit-number family inet targeted-broadcast]
user@host# forward-and-send-to-re;
```

or

- b. through the egress interface only.

```
[edit interfaces interface-name unit interface-unit-number family inet targeted-broadcast]
user@host# forward-only;
```

## Display Targeted Broadcast Configuration Options

### IN THIS SECTION

- [Forward IP Broadcast Packets on the Egress Interface and to the Routing Engine | 159](#)
- [Forward IP Broadcast Packets on the Egress Interface Only | 160](#)

The following example topics display targeted broadcast configuration options:

### Forward IP Broadcast Packets on the Egress Interface and to the Routing Engine

#### IN THIS SECTION

- [Purpose | 159](#)
- [Action | 160](#)

#### *Purpose*

Display the configuration when targeted broadcast is configured on the egress interface to forward the IP broadcast packets on the egress interface and to send a copy of the same 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-and-send-to-re;
}
```

To display the configuration for `irb`, run the `show` command at the `[edit interfaces irb unit interface-unit-number family inet]`.

```
[edit interfaces irb unit interface-unit-number family inet]
user@host#show
targeted-broadcast {
    forward-and-send-to-re;
}
```

### Forward IP Broadcast Packets on the Egress Interface Only

#### IN THIS SECTION

- Purpose | 160
- Action | 161

### Purpose

Display the configuration when targeted broadcast is configured on the egress interface to forward the IP broadcast 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;
}
```

To display the configuration, run the `show` command at the `[edit interfaces irb unit interface-unit-number family inet]`.

```
[edit interfaces irb unit interface-unit-number family inet]
user@host#show
targeted-broadcast {
    forward-only;
}
```

## Configure Targeted Broadcast (CLI Procedure)

Before you begin to configure targeted 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 \(CLI Procedure\)](#).

We recommend that you do not enable targeted broadcast on subnets that have a direct connection to the Internet because of increased exposure to DoS attacks.

This task uses Junos OS for EX Series switches that does not support the ELS configuration style. For ELS details, see *Using the Enhanced Layer 2 Software CLI*.

You can use targeted broadcast on an EX Series Switches switch to facilitate remote network management by sending broadcast packets to hosts on a specified subnet without broadcasting to the

entire network. Targeted broadcast packets are broadcast on only the target subnet. The rest of the network treats targeted broadcast packets as unicast packets and forwards the packets accordingly.

To enable targeted 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 L3 interface on the VLAN that is the target of the targeted broadcast packets:

```
[edit interfaces]
user@switch# set vlan.1 family inet address 10.1.2.1/24
```

3. Associate an L3 interface with the VLAN:

```
[edit vlans]
user@switch# set v1 l3-interface (VLAN) vlan.1
```

4. Enable the L3 interface for the VLAN to receive targeted broadcasts:

```
[edit interfaces]
user@switch# set vlan.1 family inet targeted-
broadcast
```

## Example: Configure Targeted Broadcast on a Switch

### IN THIS SECTION

- [Requirements | 163](#)
- [Overview and Topology | 163](#)

Targeted 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 targeted 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 targeted 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 [I3-interface](#).

## Overview and Topology

### IN THIS SECTION

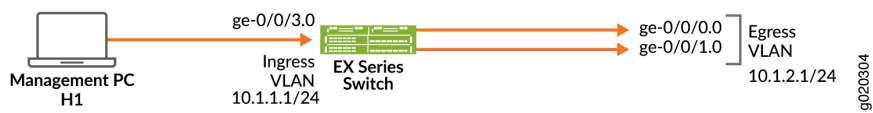
- [Topology | 164](#)

You might want to perform remote administration tasks such as backups and WOL application tasks to manage groups of clients on a subnet. One way to do the administration tasks is to send targeted broadcast packets targeted at the hosts in a particular target subnet.

The network forwards targeted broadcast packets as if the packets were unicast packets. When the targeted 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 164](#)), 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 Targeted Broadcast



Topology

Table 24 on page 164 shows the settings of the components in this example.

Table 24: Components of the Targeted Broadcast Topology

Property	Settings
Ingress VLAN name	v0
Ingress VLAN IP address	10.1.1.1/24
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

SEE ALSO

Configuring Targeted Broadcast for Switches

## Verify IP Directed Broadcast Status

### IN THIS SECTION

- Purpose | 165
- Action | 165

### 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 162](#).

## Uplink Failure Detection

### SUMMARY

Learn about the failure on uplink interfaces and conveying of this information to downward interfaces. This topic also includes how to configure interfaces for uplink failure.

### IN THIS SECTION

- Overview of Uplink Failure Detection | 166
- Configure Interfaces for Uplink Failure Detection | 168
- Example: Configure Interfaces for Uplink Failure Detection | 170
- Verify That Uplink Failure Detection Is Working Correctly | 176

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 | 166](#)
- [Failure Detection Pair | 167](#)
- [Debounce Interval | 168](#)

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 server NICs share the same IP address. The NICs are configured in a primary or secondary relationship. 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.

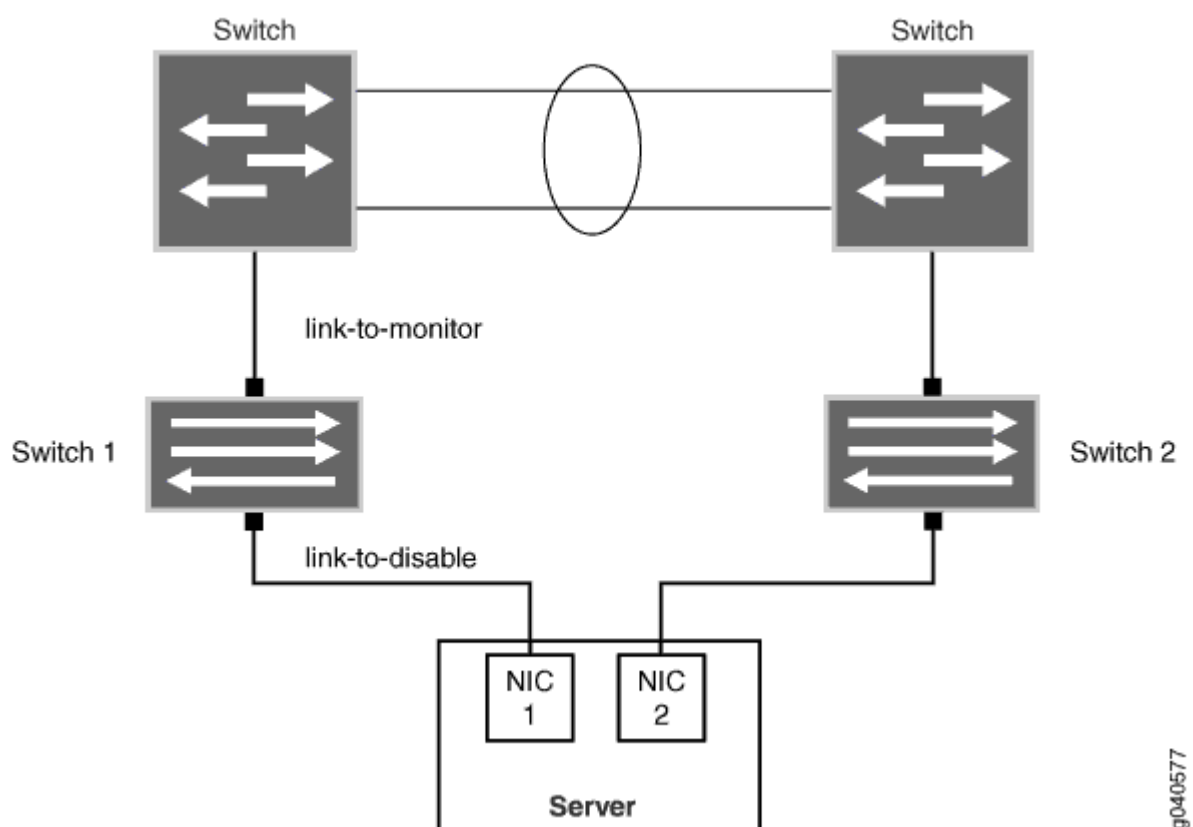
This topic describes:

### 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 9 on page 167](#) illustrates a typical setup for uplink failure detection.

Figure 9: Uplink Failure Detection Configuration on Switches



For uplink failure detection, you specify a group of uplink interfaces to be monitored. You also specify downlink interfaces to be shut 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 shuts 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.

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 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 powered up after corresponding state changes of the uplink interfaces. You can configure the debounce interval for the uplink failure detection group. Without debounce interval configuration, downlink interfaces are activated immediately after uplink state changes. This action might cause unnecessary downlink state changes and server failovers.

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 doesn't activate after the debounce timer expires, latency might occur between the timer's expiration and the interface's activation.

## Configure 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 48 groups 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.

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: Configure Interfaces for Uplink Failure Detection

### IN THIS SECTION

- [Requirements | 170](#)
- [Overview and Topology | 170](#)
- [Configure Uplink Failure Detection on both Switches | 172](#)
- [Verification | 175](#)

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 | 172](#)

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.

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 10 on page 171 illustrates a typical setup for uplink failure detection.

Figure 10: Uplink Failure Detection Configuration on Switches

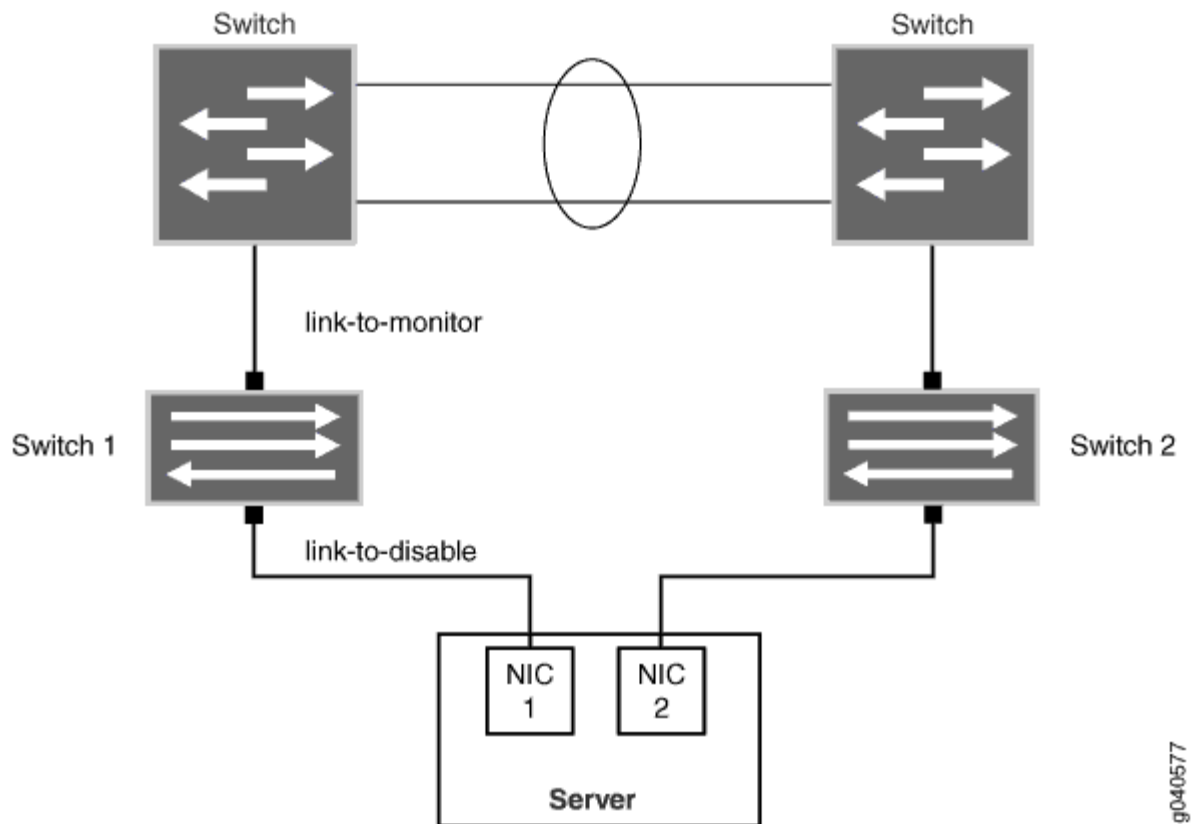


Table 25 on page 172 lists uplink failure settings for each QFX3500 switch.

Topology

Table 25: Settings for Uplink Failure Protection Example

Switch 1	Switch 2
<ul style="list-style-type: none"><li>Group name: Group1</li><li>Link-to-monitor interface: <b>xe-0/0/0</b></li><li>Link-to-disable interface: <b>xe-0/0/1</b></li><li>Debounce interval: <b>20</b></li></ul>	<ul style="list-style-type: none"><li>Group name: Group2</li><li>Link-to-monitor interface: <b>xe-0/0/0</b></li><li>Link-to-disable interface: <b>xe-0/0/1</b></li><li>Debounce interval: <b>20</b></li></ul>

Configure Uplink Failure Detection on both Switches

IN THIS SECTION

[Procedure | 172](#)

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 | 175](#)

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
```

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.

## Verify That Uplink Failure Detection Is Working Correctly

### IN THIS SECTION

- Purpose | 176
- Action | 177
- Meaning | 177

### 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
```

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.

# Generic Routing Encapsulation (GRE)

## SUMMARY

Learn about GRE, GRE tunneling, encapsulation and de-encapsulation, and configuration of GRE.

## IN THIS SECTION

- [Understand GRE | 178](#)
- [Configure Generic Routing Encapsulation \(GRE\) Tunneling | 182](#)
- [Verify That Generic Routing Encapsulation Tunneling Is Working Correctly | 184](#)

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.

## Understand GRE

### IN THIS SECTION

- [Overview of GRE | 178](#)
- [GRE Tunneling | 179](#)
- [Use a Firewall Filter to De-Encapsulate GRE Traffic | 181](#)
- [Configuration Limitations | 181](#)

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 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.

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. You can also rewrite rules on the egress port to control the transmission of a de-encapsulated GRE packet out from the egress port.

## Apply 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 26 on page 181](#) identifies these constraints.

**Table 26: 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

## Use a Firewall Filter to De-Encapsulate GRE Traffic

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 27 on page 181](#) lists features that are not supported with GRE.

**Table 27: 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.

## Configure Generic Routing Encapsulation (GRE) Tunneling

### IN THIS SECTION

- [Configure 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.

Use [GRE](#) to confirm platform and release support for specific features.

You can also use a firewall filter to de-encapsulate GRE traffic. 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 about 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
```

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:

## Configure 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
```

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.

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.

Indirect egress next-hops is currently not supported in the GRE implementation for QFX10000 switches.

## Verify That Generic Routing Encapsulation Tunneling Is Working Correctly

### IN THIS SECTION

- Purpose | 184
- Action | 185
- Meaning | 185

### 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.

# 6

CHAPTER

## Optical Transceivers for Switches

---

### IN THIS CHAPTER

- [Optical Transceivers | 187](#)
  - [400ZR Optics Support on QFX5220-32CD and QFX5130 Switches | 190](#)
-

# Optical Transceivers

## SUMMARY

Describes the optics support that the specific line cards and devices provide.

## IN THIS SECTION

- [Software Features | 187](#)
- [OTN Alarms and Defects | 188](#)
- [Supported PICs | 189](#)

## Software Features

The following interface features are supported:

- Compliant with ITU G.709 and G.798
- Performance monitoring features such as alarms, threshold-crossing alarms (TCA), 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% SD-FEC or 25% SD-FEC)
  - Differential and non-differential encoding modes
  - Transmit (TX) laser enable and disable
  - TX output power
  - Wavelength

- TCAs
- IEEE 802.1 ag 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 Packet Forwarding Engine.

## OTN Alarms and Defects

The following are OTN alarms and defects that are supported:

### 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 (TTIM)

#### 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 FEC 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

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)

## Supported PICs

[Table 1 on page 190](#) describes PICs that support optics.



Table 28: Supported PICs

PIC	Release
<a href="#">QFX10K-12C-DWDM</a>	Junos OS Release 17.2R1 and later

## 400ZR Optics Support on QFX5220-32CD and QFX5130 Switches

### SUMMARY

Learn about 400ZR optics and its configuration on QFX5220-32CD and QFX5130 switches.

### IN THIS SECTION

- [Configure 400ZR Optics | 190](#)

400ZR 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 400ZR optics fiber are the following:

- Data Center Interconnectivity (DCI) links
- Campus DWDM
- Metro DWDM

400ZR optics provides low latency and high speed.

## Configure 400ZR Optics

The following are the guidelines when you configure the 400ZR optics:

The number of ports supporting the 400ZR 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 400ZR optics. Each port supporting the 400ZR 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 commands to set the corresponding port (port 1) to unused, if the 400ZR 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`

The `set chassis` and `set interfaces` commands power on the port 0.

Use the following commands if the 400ZR 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:

**Table 29: Supported Ports and Corresponding Unused Ports**

Ports Supporting 400ZR Optics	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

Table 29: Supported Ports and Corresponding Unused Ports (*Continued*)

Ports Supporting 400ZR Optics	Corresponding Ports to Be Set Unused
20	21
23	22
24	25
27	26
28	29
31	30

- If the 400ZR 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 400ZR 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 400ZR 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 400ZR optics module, high-power mode and unused port settings are not required.

# 7

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 Series Switches | **200**
  - Port Speed on QFX Series Switches | **240**
-

# 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).

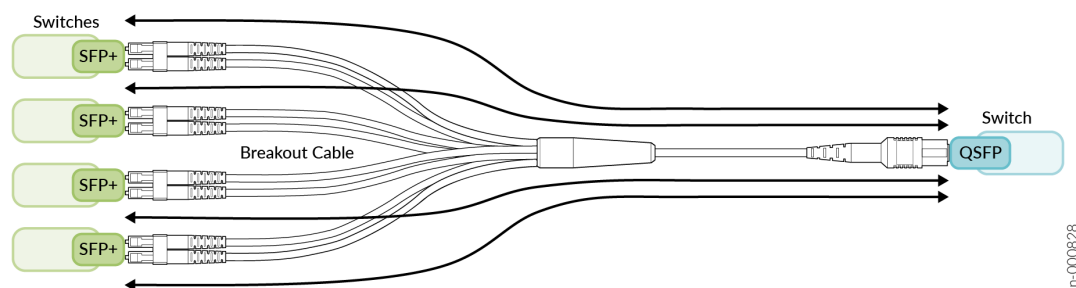
Port types:

- 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 11: 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 30: 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

Learn 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 31: Port Speed Configuration at Chassis Level

Configuration Steps	Details	Example
<b>Channelize Individual Port:</b> Configure an individual port to operate at a specific channel speed. Specify a port number and channel speed.	<pre>[edit chassis fpc fpc-slot pic pic-slot] user@host# set port port-number channel-speed speed</pre>	To configure an individual 40-Gigabit Ethernet (et) port to operate as 10-Gigabit Ethernet (xe) ports, specify a port number and channel speed.  <pre>[edit chassis fpc 0 pic 0] user@host# set port 3 channel-speed 10g</pre>
<b>Channelize Block of Ports:</b> Channelize a block of ports. Specify a port range and channel speed.	<pre>[edit chassis fpc fpc-slot pic pic-slot] user@host# set port-range port-range-low port-range-high channel-speed speed</pre>	To configure ports 0 through 3 on PIC 0 to operate as 50-Gigabit Ethernet ports:  <pre>[edit chassis fpc 0 pic 0] user@host# set port-range 0 3 channel-speed 50g</pre>
<b>Configure Speed per Quad:</b> 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.	<pre>[edit chassis fpc fpc-slot pic pic-slot] user@host# set port port-number speed speed</pre>	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.  <pre>[edit chassis fpc 0 pic 0] user@host# set port 4 speed 25g</pre>
<b>Configure Speed on an Individual Port</b>	<pre>[edit chassis fpc <i>fpc-slot</i> pic <i>pic-slot</i>] user@host# set port <i>port-number</i> speed <i>speed</i></pre>	<pre>[edit chassis fpc 0 pic 0] user@host# set port 3 speed 25g</pre>



## Configure Speed at Interfaces Level

Table 32: 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.	<pre>[edit interfaces interface-name] user@host# set speed (10G   25G   40G   50G   100G   400G)</pre> <p>For example:</p> <pre>[edit interfaces et-1/0/3] user@host# set speed 100g</pre>	<pre>[edit interfaces interface-name] user@host# set speed (10G   25G   40G   50G   100G   400G)</pre> <p>For example:</p> <pre>[edit interfaces et-1/0/3] user@host# set speed 100g</pre>
Step 2: To configure the speed for a group of ports.	<pre>[edit ] user@host# wildcard range set interfaces interface-name speed speed</pre> <p>For example:</p> <pre>[edit ] user@host# wildcard range set interfaces et-1/0/[0-5] speed 100g</pre>	<pre>[edit ] user@host# wildcard range set interfaces interface-name speed speed</pre> <p>For example:</p> <pre>[edit ] user@host# wildcard range set interfaces et-1/0/[7-12] speed 100g</pre>
Step 3: To specify the number of interfaces you want to configure per port.	Not applicable	<pre>[edit interfaces interface-name] user@host# set number-of-sub-ports number-of-sub-ports</pre> <p>For example:</p> <pre>[edit interfaces et-1/0/3] user@host# set number-of-sub-ports 4</pre> <p>In this example, in Step 1 and Step 2, you configure 4x100GE channelized interfaces.</p>

Table 32: 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.	<p><b>[edit] user@host# set interfaces interface-name unused</b></p> <p><b>For example:</b></p> <p>[edit] user@host# set interfaces et-2/0/3 unused</p> <p>In this example, no interfaces (channelized or non-channelized) are created on port 3 of the line card installed in the FPC slot 2.</p>	<p>[edit] user@host# set interfaces interface-name unused</p> <p><b>For example:</b></p> <p>[edit] user@host# set interfaces et-2/0/4 unused</p> <p>In this example, no interfaces (channelized or non-channelized) are created on port 4 of the line card installed in the FPC slot 2.</p>
Step 5: Verify the configuration.	<pre>et-x/y/z { speed 100g; unit 0 { ... } ... unit N { ... } } ... et-x/y/z { unused;</pre>	<pre>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;</pre>
Step 6: Commit the configuration.		

# Port Speed on EX Series Switches

## SUMMARY

Learn supported port speeds, channelization, and interface naming conventions of EX Series switches.

## IN THIS SECTION

- [Channelizing Interfaces 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 EX4100-H-24MP and EX4100-H-24F Switches | 231](#)
- [Port Speed on EX4000-12MP, EX4000-24MP, and EX4000-48MP Switches | 234](#)
- [Operating Speed of Interfaces on EX Switches | 236](#)

## Channelizing Interfaces 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 33: 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.

**Table 33: EX4650-48Y Details and Description (Continued)**

Details	Description
Auto speed detection mode (Enabled by default)	<p>If you have disabled auto-channelization, manually channelize the port speed using the <code>set chassis fpc slot-number port port-number channel-speed speed</code> command. You can set the speed 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. You cannot commit check this, however.</p> <p>On EX4650 switches, the extension module ports support auto-channelization.</p>

[Table 34 on page 201](#) summarizes the supported port speeds on EX4650-48Y.

**Table 34: 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 35 on page 202 lists the interface naming conventions of SFP+ ports (labeled 0 through 47) for the EX4650-48Y switch.

**Table 35: 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
	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 36 on page 202 lists the interface naming conventions of extension module ports (labeled 48 through 55) for the EX4650-48Y switch.

**Table 36: 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

**Table 36: Interface Naming Convention for the EX4650-48Y Switch (Extension Module Ports)**  
(Continued)

PIC	10-Gigabit Ethernet Interface	25-Gigabit Ethernet Interface	40-Gigabit Ethernet Interface	100-Gigabit Ethernet Interface
	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 37: 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

Table 37: Port Speed for EX4400-24T and EX4400-24P (Continued)

PIC	Ports	Supported Port Speeds	Default Speed
		4x25 Gbps	
		4x10 Gbps	
PIC 2	1x100 GbE QSFP28 extension module (model number: EX4400-EM-1C)	100 Gbps	100 Gbps
	<p>The extension module port can operate as a VCP using HGoE.</p> <p><b>NOTE:</b> 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 upgrade that is required on EX4400-24X to support the 1x100 GbE QSFP28 extension module".</p> <p>See <a href="#">Installing and Upgrading Firmware</a> and <a href="#">request system firmware upgrade</a> for steps to upgrade the firmware.</p>	40 Gbps	40 Gbps
		4x25 Gbps	
		4x10 Gbps	
	<p>4x25 GbE SFP28 extension module (model number: EX4400-EM-4Y)</p> <p>When the extension module ports operate at 25 Gbps speed, you can configure them to operate as VCPs using HGoE.</p>	25 Gbps	25 Gbps
		10 Gbps	<p>4x25G extension module can also support 10 GbE or 1 GbE upon setting the first port to 10 GbE/1 GbE using the command: <code>set chassis fpc &lt;fpc-slot&gt; pic-slot 2 port 0 speed &lt;10g/1g&gt;</code></p> <p>You can revert to the default 25-gigabit mode by using <code>set chassis fpc &lt;fpc-slot&gt; pic-slot 2 port 0 speed 25g</code> or <code>delete chassis fpc &lt;fpc-slot&gt; pic-slot 2 port 0 speed &lt;1g/10g&gt;</code> command.</p> <p>All the ports in the extension module operate in the same speed.</p>
		1 Gbps	

Table 37: Port Speed for EX4400-24T and EX4400-24P (Continued)

PIC	Ports	Supported Port Speeds	Default Speed
	4x10 GbE SFP+ extension module (model number: EX4400-EM-4S)	10 Gbps	10 Gbps or 1 Gbps.
		1 Gbps	<p>The default speed depends on the plugged-in transceiver and is the default behavior.</p> <p>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.</p> <p>To set speed explicitly to 1G or 10G, use the following command:</p> <pre>set chassis fpc <i>fpc-slot</i> pic <i>pic-number</i> port <i>port-number</i> speed <i>port speed</i>.</pre> <p>For example: set chassis fpc 0 pic 2 port 0 speed 1G.</p> <p>The set chassis command overrides the default behavior and all the ports in the PIC operates in the configured speed.</p>



Table 38: 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	
PIC 2	1x100 GbE QSFP28 extension module (model number: EX4400-EM-1C)  The extension module port can operate as a VCP using HGoE.	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.	25 Gbps	25 Gbps
		10 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>
		1 Gbps	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.

Table 38: Port Speed for EX4400-48T and EX4400-48P (Continued)

PIC	Ports	Supported Port Speeds	Default Speed
	4x10 GbE SFP+ extension module (model number: EX4400-EM-4S)	10 Gbps	10 Gbps or 1 Gbps.
		1 Gbps	<p>The default speed depends on the plugged-in transceiver and is the default behavior.</p> <p>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.</p> <p>To set speed explicitly to 1G or 10G, use the following command:</p> <pre>set chassis fpc <i>fpc-slot</i> pic <i>pic-number</i> port <i>port-number</i> speed <i>port speed</i>.</pre> <p>For example: set chassis fpc 0 pic 2 port 0 speed 1G.</p> <p>The set chassis command overrides the default behavior and all the ports in the PIC operates in the configured speed.</p>

Table 39: Port Speed for EX4400-48F

PIC	Ports	Supported Port Speeds	Default Speed
PIC 0	36 built-in (SFP) ports (Numbered 0 through 35)	100 Mbps/ 1 Gbps	1 Gbps
	12 built-in (SFP+) ports (Numbered 36 through 47)	10 Gbps	10 Gbps

Table 39: Port Speed for EX4400-48F (Continued)

PIC	Ports	Supported Port Speeds	Default Speed
	Auto-negotiation is only supported with copper SFPs for 1 Gbps speed.	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.	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.	25 Gbps	25 Gbps
		10 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>
		1 Gbps	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.

Table 39: Port Speed for EX4400-48F (Continued)

PIC	Ports	Supported Port Speeds	Default Speed
		1 Gbps	<p>The default speed depends on the plugged-in transceiver and is the default behavior.</p> <p>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.</p> <p>To set speed explicitly to 1G or 10G, use the following command:</p> <pre>set chassis fpc fpc-slot pic pic-number port port-number speed port speed.</pre> <p>For example: set chassis fpc 0 pic 2 port 0 speed 1G.</p> <p>The set chassis command overrides the default behavior and all the ports in the PIC operates in the configured speed.</p>

Table 40: Port Speed for EX4400-24X

PIC	Port Number/Module	Port Speed Supported	Default Speed
PIC 0	24 (0-23) fixed ports	1 GbE 1G copper ports support autonegotiation. 1G SFP ports do not support autonegotiation.	1 GbE
		10 GbE (SFP+ ports) SFP+ ports do not support autonegotiation.	10 GbE

Table 40: Port Speed for EX4400-24X *(Continued)*

PIC	Port Number/Module	Port Speed Supported	Default Speed
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 40: 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	<p>10 Gbps or 1 Gbps.</p> <p>The default speed depends on the plugged-in transceiver and is the default behavior.</p> <p>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.</p> <p>To set speed explicitly to 1G or 10G, use the following command:</p> <pre>set chassis fpc <i>fpc-slot</i> pic <i>pic-number</i> port <i>port-number</i> speed <i>port speed</i>.</pre> <p>For example: set chassis fpc 0 pic 2 port 0 speed 1G.</p> <p>The set chassis command overrides the default behavior and all the ports in the PIC operates in the configured speed.</p>

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 41: 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.	100 Gbps	100 Gbps
		40 Gbps	40 Gbps
		4x25 Gbps	

Table 41: Port Speed for EX4400-24MP (Continued)

PIC	Ports	Port Speeds Supported	Default Speed
		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.	25 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: <code>set chassis fpc &lt;fpc-slot&gt; pic-slot 2 port 0 speed &lt;10g/1g&gt;</code>  You can revert to the default 25-gigabit mode by using <code>set chassis fpc &lt;fpc-slot&gt; pic-slot 2 port 0 speed 25g</code> or <code>delete chassis fpc &lt;fpc-slot&gt; pic-slot 2 port 0 speed &lt;1g/10g&gt;</code> 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.
		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:  <code>set chassis fpc fpc-slot pic pic-number port port-number speed port speed.</code>  For example: <code>set chassis fpc 0 pic 2 port 0 speed 1G.</code>  The <code>set chassis</code> command overrides the default behavior and all the ports in the PIC operates in the configured speed.



Table 42: 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)  The extension module port can operate as a VCP using HGoE.	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	25 Gbps	25 Gbps
		10 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>

Table 42: Port Speed for EX4400-48MP (Continued)

PIC	Ports	Supported Port Speeds	Default Speed
	speed, you can configure them to operate as VCPs using HGoE.	1 Gbps	<p>You can revert to the default 25-gigabit mode by using <code>set chassis fpc &lt;fpc-slot&gt; pic-slot 2 port 0 speed 25g</code> or <code>delete chassis fpc &lt;fpc-slot&gt; pic-slot 2 port 0 speed &lt;1g/10g&gt;</code> command.</p> <p>All the ports in the extension module operate in the same speed.</p>
	4x10 GbE SFP+ extension module (model number: EX4400-EM-4S)	10 Gbps	10 Gbps or 1 Gbps.
		1 Gbps	<p>The default speed depends on the plugged-in transceiver and is the default behavior.</p> <p>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.</p> <p>To set speed explicitly to 1G or 10G, use the following command:</p> <pre>set chassis fpc fpc-slot pic pic-number port port-number speed port speed.</pre> <p>For example: <code>set chassis fpc 0 pic 2 port 0 speed 1G.</code></p> <p>The <code>set chassis</code> command overrides the default behavior and all the ports in the PIC operates in the configured speed.</p>

**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 43: 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 44 on page 217, Table 45 on page 219, Table 49 on page 225, Table 48 on page 223, Table 46 on page 221, Table 47 on page 222, and Table 50 on page 227 summarizes the supported port speeds on EX4100 switches.

**Table 44: 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	<p>Downlink ports support 1-Gbps speed with full-duplex. Both full-duplex and half-duplex are supported on 100-Mbps and 10-Mbps speeds.</p> <p>By default, the ports come up with 1-Gbps speed. You can configure the other port speeds at the [edit chassis] hierarchy level.</p> <p>Autonegotiation is supported and enabled by default.</p>

Table 44: 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	<p>Switches support 25-Gbps and 10-Gbps speed in both virtual chassis and network modes. See <a href="#">"Virtual Chassis Ports and Network Ports"</a> on page 228. 1-Gbps speed is supported only in network mode.</p> <p>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.</p> <p>Autonegotiation is not supported. The <code>show interfaces interface-name</code> command displays incorrect autonegotiation status when you disable autonegotiation.</p> <p>You can configure the network ports in mixed speed. For example, 2x10 GbE on ports 0 and 1, 1x1 GbE on port 2,</p>

Table 44: 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 45: 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	<p>Downlink ports support 1 Gbps speed with full-duplex. Both full-duplex and half-duplex are supported on 100 Mbps and 10 Mbps speeds.</p> <p>By default, the ports come up with 1 Gbps speed. You can configure the other port speeds at the [edit chassis] hierarchy level.</p> <p>Autonegotiation is supported and enabled by default.</p>

Table 45: 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	<p>Switches support 25 Gbps and 10 Gbps speed in both VC and network modes. See <a href="#">"Virtual Chassis Ports and Network Ports"</a> on page 228. 1 Gbps speed is supported only in network mode.</p> <p>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.</p> <p>Autonegotiation is not supported. The <code>show interfaces interface-name</code> command displays incorrect autonegotiation status when you disable autonegotiation.</p> <p>You can configure the network ports in mixed speed. For example, 2x10G on ports 0 and 1, 1x1G</p>

Table 45: 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 46: Port Speed for EX4100-24MP

PIC	Port Number	Port Speed Supported	Default Speed
PIC 0	Downlink ports (0 - 23)	<ul style="list-style-type: none"> <li>Ports (0 - 7) are multi-rate ports that support 100 Mbps, 1 Gbps, 2.5 Gbps, 5 Gbps, and 10 Gbps.</li> <li>Ports (8-23) are GigE ports that support 10 Mbps, 100 Mbps, and 1 Gbps speed.</li> </ul>	<ul style="list-style-type: none"> <li>For multi-rate ports - 10 Gbps</li> <li>For GigE ports - 1 Gbps</li> </ul>
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 <a href="#">"Virtual Chassis Ports and Network Ports"</a> on page 228.	No default value



**Table 46: 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 47: Port Speed for EX4100-48MP**

PIC	Port Number	Port Speed Supported	Default Speed
PIC 0	Downlink ports (0 - 47)	<ul style="list-style-type: none"> <li>Ports (0 - 15) are multi-rate ports that support 100 Mbps, 1 Gbps, and 2.5 Gbps.</li> <li>Ports (16-47) are GigE ports that support 10 Mbps, 100 Mbps, and 1 Gbps speed.</li> </ul>	<ul style="list-style-type: none"> <li>For multi-rate ports - 2.5 Gbps</li> <li>For GigE ports - 1 Gbps</li> </ul>
PIC 1	VCPs ports (0 - 3)	<p>4x25 Gbps</p> <p>4x10 Gbps</p> <p>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.)</p> <p>If you convert the VCPs to network ports, ports 0 through 3 on PIC1 support 1 Gbps, , 10 Gbps, or 25 Gbps speed. See <a href="#">"Virtual Chassis Ports and Network Ports" on page 228</a>.</p>	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 48: 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 48: 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	<p>Fixed switches support 10-Gbps speed in both VC and network modes. See <a href="#">"Virtual Chassis Ports and Network Ports"</a> on page 228. 1-Gbps speed is supported only in network mode.</p> <p>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.</p> <p>Autonegotiation is not supported. The <code>show interfaces interface-name</code> command displays incorrect autonegotiation status when you disable autonegotiation.</p> <p>You can configure the network ports in mixed speed. For example, you can configure 2x10G on ports 0 and 1 and</p>

Table 48: 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 49: 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	<p>Downlink ports support 1-Gbps speed with full-duplex. Both full-duplex and half-duplex are supported on 100-Mbps and 10-Mbps speeds.</p> <p>By default, the ports come up with 1-Gbps speed. You can configure the other port speeds at the [edit chassis] hierarchy level.</p> <p>Autonegotiation is supported and enabled by default.</p>

Table 49: 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	<p>Fixed switches support 10-Gbps speed in both VC and network modes. See <a href="#">"Virtual Chassis Ports and Network Ports"</a> on page 228. 1-Gbps speed is supported only in network mode.</p> <p>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.</p> <p>Autonegotiation is not supported. The <code>show interfaces interface-name</code> command displays incorrect autonegotiation status when you disable autonegotiation.</p> <p>You can configure the network ports in mixed speed. For example, you can configure 2x10G on ports 0 and 1 and</p>

Table 49: 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 50: 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 " <a href="#">Virtual Chassis Ports and Network Ports</a> " 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:

- 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.
- The EX4100 switches support HiGig over Ethernet (HGoE) mode. HGoE enables mixed mode operation where you can configure some ports as VCPs and other ports as standard network ports. HGoE is also the default VC mode for EX4100-H switches. The HGoE mode provides flexibility in network configuration.
- 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 51: 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 51: 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 52: 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 53: 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 LAG. You can form aggregated Ethernet LAG only if all ports in the LAG are of the same speed. For example: ge ports, 2.5 Gbps mge ports, or xe ports.
- 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 EX4100-H-24MP and EX4100-H-24F Switches

### SUMMARY

Learn about port speed, and auto negotiation information of EX4100-H-24MP and EX4100-H-24F switches.

To view the supported transceivers, optical interfaces, and DAC cables on EX4100-H-24MP and EX4100-H-24F, see [Hardware Compatibility Tool \(HCT\)](#).

Network Interfaces Support. EX4100-H-24MP and EX4100-H-24F includes three PICs with speeds as given below:

- Downlink ports
  - EX4100-H-24MP PIC 0 (ports 0 - 23). The first 8 are multi-rate ports that support 100-Mbps, 1-Gbps, and 2.5-Gbps. The remaining 16 ports are GigE ports that support 10-Mbps, 100-Mbps, and 1Gbps.
  - EX4100-H-24F PIC 0 (ports 0 - 23). The 24 ports support 100-Mbps, and 1-Gbps (10-Mbps, 100-Mbps and 1-Gbps on tri-rate SFP).
- Stacking/Network ports on EX4100-H-24MP and EX4100-H-24F PIC 1 (ports 0-3) support 1-Gbps and 10-Gbps speeds on network mode. If you convert the network mode to virtual chassis (VC)

mode, ports 0 through 3 on PIC 1 support 10-Gbps speeds. On PIC 1 ports, we support HiGig over Ethernet (HGoE) as default and HiGig for virtual chassis formation.

- Uplink ports on EX4100-H-24MP and EX4100-H-24F PIC 2 (ports 0-3) support 1-Gbps and 10-Gbps speeds on network mode. If you convert the network mode to virtual chassis mode, ports 0 through 3 on PIC 2 support 10-Gbps speeds. On PIC 2 ports, we support only HiGig over Ethernet (HGoE) for virtual chassis formation.

**Table 54: Port speed support on EX4100-H-24MP switches.**

PIC	Port Number and Type of Ports	Supported Port Speed	Default Speed
PIC 0	8 RJ45 ports (Downlink ports)	100 Mbps, 1 Gbps, and 2.5 Gbps	2.5 Gbps
	16 RJ45 ports (Downlink ports)	10 Mbps, 100 Mbps, and 1 Gbps	1 Gbps
PIC 1	4 SFP+ ports	1 Gbps and 10 Gbps	Depends on the plugged-in transceiver.
PIC 2	4 SFP+ ports (Uplink ports)	1 Gbps and 10 Gbps	Depends on the plugged-in transceiver.

**Table 55: Interface Naming Conventions for EX4100-H-24MP switches.**

PIC	Interface Type	Interfaces
PIC 0	RJ45	mge-0/0/0 – mge-0/0/7
		ge-0/0/8 – ge-0/0/23
PIC 1	SFP	ge-0/1/0 – ge-0/1/3
	SFP+	xe-0/1/0 - xe-0/1/3
PIC 2	SFP	ge-0/2/0 - ge-0/2/3
	SFP+	xe-0/2/0 - xe-0/2/3

**Table 56: Port speed support on EX4100-H-24F switches.**

PIC	Port Number and Type of Ports	Supported Port Speed	Default Speed
PIC 0	24 SFP ports (Downlink ports)	10 Mbps, 100 Mbps, and 1 Gbps	1 Gbps
PIC 1	4 SFP+ ports	1 Gbps and 10 Gbps	Depends on the plugged-in transceiver.
PIC 2	4 SFP+ ports (Uplink ports)	1 Gbps and 10 Gbps	Depends on the plugged-in transceiver.

**Table 57: Interface Naming Conventions for EX4100-H-24F switches.**

PIC	Interface Type	Interfaces
PIC 0	SFP	ge-0/0/0 – ge-0/0/23
PIC 1	SFP	ge-0/1/0 – ge-0/1/3
	SFP+	xe-0/1/0 - xe-0/1/3
PIC 2	SFP	ge-0/2/0 - ge-0/2/3
	SFP+	xe-0/2/0 - xe-0/2/3

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.
- 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-24MP and EX4100-H-24F 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-24MP and EX4100-H-24F switches do not support interface hold timer in sub seconds.
- The mge ports in PIC 0 do not support auto-MDIX disable.

SEE ALSO

- [Understanding HiGig and HGoE Modes in a Virtual Chassis](#)
- [Network Interfaces for EX Series](#)
- [Port Speed on EX4100 Switches](#)

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 58: 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 59: 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		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.

## 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

## Port Speed on QFX Series Switches

### SUMMARY

Learn supported port speeds, channelization, and interface naming conventions of QFX Series switches.

### IN THIS SECTION

- [Port Speed on QFX5100-24Q Switches | 241](#)
- [Port Speed on QFX5110-48S Switches | 242](#)
- [Port Speed on QFX5120-32C Switches | 243](#)
- [Port Speed on QFX5120-48T Switches | 244](#)
- [Port Speed on QFX5120-48Y Switches | 245](#)
- [Port Speed on QFX5120-48YM Switches | 246](#)
- [Port Speed on QFX5130-32CD Switches | 247](#)
- [Port Speed on QFX5130-48C/QFX5130-48CM Switches | 249](#)
- [Port Speed on QFX5200-32C Switches | 256](#)
- [Port Speed on QFX5210-64C Switches | 256](#)
- [Port Speed on QFX5230-64CD Switches | 257](#)
- [Port Speed on QFX5240 Switches | 261](#)
- [Port Speed on QFX5700 Switches | 265](#)

## 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 60: Port Speed on QFX5100-24Q Switches**

PIC 0	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 | 243](#)

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 61: 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 61: Port Speed on QFX5110-48S Switches (Continued)**

PIC	Port Number	Port Speeds Supported
	48-51 (QSFP28 ports)	<p>40 Gbps</p> <p>100 Gbps</p> <p>You can configure each port as an independent 100-GbE port or as an independent 40-GbE port.</p> <p>QSFP28 ports support channelization of 40 Gbps into 4x10 Gbps interfaces using breakout cables.</p>

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) connection to an 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 62: 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 63: 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 63: 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 Switch Hardware Guide](#).

**Table 64: 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 64: 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 Switch Hardware Guide](#).

**Table 65: 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 Switch Hardware Guide](#).

**Table 66: 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 up to 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 | 251](#)
- [Channelization | 251](#)
- [Supported FEC Modes | 255](#)

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 67 on page 249](#) for details.

The QSFP-DD ports support the following channelizations:

- 4x100GbE
- 2x200GbE
- 8x50GbE
- 4x25GbE
- 4x10GbE

The SFP-DD ports support 2x50G channelization.

**Table 67: 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 67: 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 68: 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-0/0/24 as *unused*.
- To channelize QSFP-DD port into 2x200G, do not mark any port as *unused*.

Table 69: Unused SFP Ports for 8x50G QSFP Port Channelization

Port to be channelized	Ports Unused
48	0
	1

Table 69: 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 69: 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 69: 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 70: 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 70: Unused SFP Ports for 4x100G, 4x25G, and 4x10G QSFP Port Channelization (Continued)**

Port to be channelized	Ports Unused
54	36
55	41

See [Table 67 on page 249](#) for details.

## Supported FEC Modes

See [Table 71 on page 255](#) for supported FEC modes on different transceivers.

**Table 71: 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 Switch Hardware Guide](#).

**Table 72: 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 73: Port Speed on QFX5210-64C Switches**

PIC	Port Number	Port Speeds Supported
PIC 0	0-63 (QSFP28 ports)	<p>40 Gbps</p> <p>100 Gbps</p> <p>Supports channelization of 100 Gbps into 2x50 Gbps interfaces or 4x25 Gbps using breakout cables.</p> <p>Supports channelization of 40 Gbps into 4x10 Gbps interfaces using breakout cables.</p>

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 | 259](#)
- [Channelization | 260](#)

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 74 on page 258](#) for details.

**Table 74: 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 74: 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 75 on page 259 lists the interface naming conventions for QFX5230-64CD switches.

Table 75: 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 74 on page 258](#) for details.

# Supported FEC Modes

[Table 76 on page 260](#) lists the supported FEC modes

**Table 76: FEC Modes Supported on Transceivers**

Optical Transceiver	FEC Mode
QSFP56-DD-400GbE	FEC119
QSFP56-DD-200GbE	FEC91-RS544
QSFP28-DD-100GbE	FEC91

Table 76: FEC Modes Supported on Transceivers *(Continued)*

Optical Transceiver	FEC Mode
QSP+-40GbE	None
SFP+-10GbE	None

## Port Speed on QFX5240 Switches

SUMMARY	<p>IN THIS SECTION</p> <ul style="list-style-type: none"> <li>Channelization   262</li> </ul>
---------	---

To view the supported transceivers, optical interfaces, and DAC cables on QFX5240-64OD, see [Hardware Compatibility Tool](#).

QFX5240-64OD supports the following speeds:

- 800 Gbps on octal small form factor pluggable (OSFP) ports
- 10 Gbps on SFP ports

Table 77: 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 77: 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 78: 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 79: 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 79: 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 80: 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](#).

# 8

CHAPTER

## Monitor Interfaces

---

### IN THIS CHAPTER

- [Monitor Interface Status and Traffic | 267](#)
  - [Monitor System Process Information | 268](#)
  - [Monitor System Properties | 269](#)
  - [Monitor Statistics for a Fast Ethernet or Gigabit Ethernet Interface | 272](#)
  - [Trace Operations of the Interface Process | 274](#)
-

# Monitor Interface Status and Traffic

## SUMMARY

Learn how to monitor interface status and traffic on Junos devices.

## IN THIS SECTION

- [Purpose | 267](#)
- [Action | 267](#)
- [Meaning | 267](#)

## 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*.

# Monitor System Process Information

## SUMMARY

Learn how to monitor system process information on a Junos device.

## IN THIS SECTION

- [Purpose | 268](#)
- [Action | 268](#)
- [Meaning | 268](#)

## 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 81 on page 268](#) summarizes the output fields in the system process information display.

The display includes the total CPU load and total memory utilization.

**Table 81: Summary of System Process Information Output Fields**

Field	Values
PID	Identifier of the process.
Name	Owner of the process.

Table 81: Summary of System Process Information Output Fields *(Continued)*

Field	Values
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.

RELATED DOCUMENTATION

| *show system uptime*

# Monitor System Properties

SUMMARY

Learn about how to monitor system properties such as uptime, users, and storage.

IN THIS SECTION

- [Purpose | 269](#)
- [Action | 270](#)
- [Meaning | 270](#)

## 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 82 on page 270](#) summarizes key output fields in the system properties display.

**Table 82: Summary of Key System Properties Output Fields**

Field	Values	Additional Information
<b>General Information</b>		
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.	

**Table 82: Summary of Key System Properties Output Fields (Continued)**

Field	Values	Additional Information
Time		
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.	

Table 82: Summary of Key System Properties Output Fields *(Continued)*

Field	Values	Additional Information
From	System from which the user has logged in. A hyphen indicates that the user is logged in through the console.	
Login Time	Time when the user logged in.	This is the <b>user@switch</b> field in <code>show system users</code> command output.
Idle Time	How long the user has been idle.	

RELATED DOCUMENTATION

| [show system processes](#)

# Monitor Statistics for a Fast Ethernet or Gigabit Ethernet Interface

SUMMARY

Learn about how to monitor real-time statistics on Fast Ethernet and Gigabit Ethernet interfaces.

IN THIS SECTION

- [Purpose | 272](#)
- [Action | 273](#)
- [Meaning | 274](#)

## 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:
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

---

## SUMMARY

Learn about how to trace interface process operations using traceoptions.

---

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](#).

## RELATED DOCUMENTATION

---

[Tracing Interface Operations Overview](#)

---

[Tracing Operations of an Individual Router Interface](#)

---

*traceoptions*

# 9

CHAPTER

## Troubleshoot Interfaces

---

### IN THIS CHAPTER

- [Troubleshoot Network Interfaces | 278](#)
  - [Diagnose a Faulty Twisted-Pair Cable \(CLI Procedure\) | 280](#)
  - [Troubleshoot Uplink Ports on EX2300 Switches | 284](#)
  - [Troubleshoot an Aggregated Ethernet Interface | 286](#)
-



# Troubleshoot Network Interfaces

## SUMMARY

Learn about how to resolve inaccurate statistics for logical interfaces on Layer 2 interfaces and fix down interfaces on ports with SFP or SFP+ transceivers

## IN THIS SECTION

- [Statistics for logical interfaces on Layer 2 interfaces are not accurate | 278](#)
- [The interface on the port in which an SFP or SFP+ transceiver is installed in an SFP or SFP+ module is down | 279](#)

## Statistics for logical interfaces on Layer 2 interfaces are not accurate

### IN THIS SECTION

- [Problem | 278](#)
- [Cause | 278](#)
- [Solution | 279](#)

### Problem

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.

### Cause

By default, QFX5000 switches do not collect statistics for logical interfaces on Layer 2 interfaces. This limitation also affects the child interfaces of Layer 2 AE interfaces, which leads to incomplete or missing I/O statistics.

## 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 | 279](#)
- [Cause | 279](#)
- [Solution | 279](#)

## Problem

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.

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.

# Diagnose a Faulty Twisted-Pair Cable (CLI Procedure)

## SUMMARY

Learn about how to diagnose cable faults using the TDR test.

## IN THIS SECTION

- [Platform-Specific Time Domain Reflectometry \(TDR\) Behavior | 283](#)

## Problem

A 10/100/1000BASE-T Ethernet interface has connectivity problems that you suspect might be caused by a faulty cable.

## Cause

Ethernet cables can become faulty or degrade due to physical damage, poor termination, electromagnetic interference, or wear over time. Issues such as open circuits, short circuits, impedance mismatches, pair swaps, polarity reversals, or excessive pair skew can disrupt proper signal transmission which leads to intermittent failed connectivity on the interface.

## 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.

Use [Feature Explorer](#) to confirm platform and release support for specific features.

Review the Platform-Specific Time Domain Reflectometry (TDR) Behavior section for notes related to your platform.

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.

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

## Platform-Specific Time Domain Reflectometry (TDR) Behavior

Use [Feature Explorer](#) to confirm platform and release support for specific features.

Use the following table to review platform-specific time domain reflectometry behaviors for your platform:

Platform	Difference
EX Series	<ul style="list-style-type: none"> <li>EX2300, EX3200, EX3400, and EX4300 switches support the Time Domain Reflectometry (TDR) test on RJ-45 network interfaces. However, the TDR test is not supported on management interfaces or SFP interfaces for these platforms.</li> <li>EX6200 and EX8200 switches support the TDR test on RJ-45 network interfaces located on the line cards. We recommend to perform the TDR test only when there is no traffic on the interface to ensure accurate results.</li> </ul>

RELATED DOCUMENTATION

[Troubleshooting Interface Configuration and Cable Faults](#)

*request diagnostics tdr*

*show diagnostics tdr*

# Troubleshoot 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 | 284](#)

Learn about how to troubleshoot 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

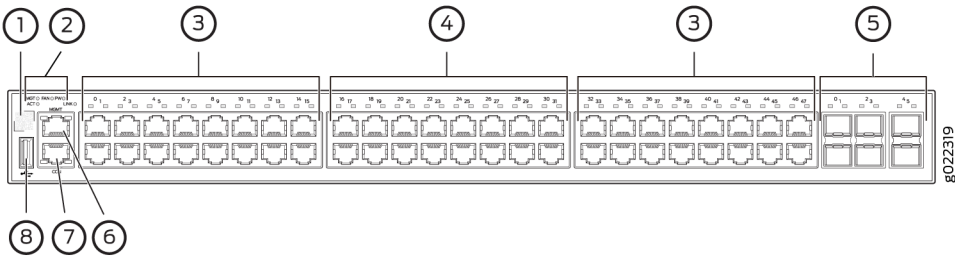
- [Problem | 284](#)
- [Cause | 285](#)
- [Solution | 285](#)

### Problem

### Description

The uplink ports 4 and 5 (see [Figure 12 on page 285](#)) do not support the speeds 10-Mbps and 100-Mbps.

Figure 12: 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 UI, 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.



# Troubleshoot an Aggregated Ethernet Interface

## IN THIS SECTION

- [Show Interfaces Command Shows the LAG Is Down | 286](#)
- [Logical Interface Statistics Do Not Reflect all Traffic | 287](#)
- [IPv6 Interface Traffic Statistics Are Not Supported | 287](#)
- [SNMP Counters ifHCInBroadcastPkts and ifInBroadcastPkts Are Always 0 | 288](#)

Troubleshooting issues for aggregated Ethernet interfaces:

## Show Interfaces Command Shows the LAG Is Down

### IN THIS SECTION

- [Problem | 286](#)
- [Solution | 286](#)

### 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 | 287](#)
- [Solution | 287](#)

### 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 | 288](#)
- [Solution | 288](#)

## 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 | [288](#)

● Solution | [288](#)

## 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](#)

# 10

CHAPTER

## Configuration Statements and Operational Commands

---

### IN THIS CHAPTER

- Common Output Fields Description | 290
  - Junos CLI Reference Overview | 300
-

# Common Output Fields Description

## IN THIS SECTION

- [Damping Field | 290](#)
- [Destination Class Field | 291](#)
- [Enabled Field | 291](#)
- [Filters Field | 292](#)
- [Flags Fields | 292](#)
- [Label-Switched Interface Traffic Statistics Field | 297](#)
- [Policer Field | 298](#)
- [Protocol Field | 298](#)
- [RPF Failures Field | 299](#)
- [Source Class Field | 299](#)

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:

Destination class	Packets (packet-per-second)	Bytes (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-fpcl/picport-in-policer, Output: type-fpcl/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)    (
```

## Junos CLI Reference Overview

We've consolidated all Junos CLI commands and configuration statements in one place. Read this guide to learn about the syntax and options that make up the statements and commands. Also 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](#)