



# Configuring Queuing and Flow Control

---

This chapter contains the following sections:

- [Information About Queues, page 1](#)
- [Information About Flow Control, page 3](#)
- [Configuring Queuing, page 4](#)
- [Configuring Flow Control, page 7](#)
- [Verifying the Queue and Flow Control Configurations, page 9](#)

## Information About Queues

### Ingress Queuing Policies

You can associate an ingress policy map with an Ethernet interface to guarantee bandwidth for the specified traffic class or to specify a priority queue.

The ingress policy is applied in the adapter to all outgoing traffic that matches the specified CoS value.

When you configure an ingress policy for an interface, the switch sends the configuration data to the adapter. If the adapter does not support the DCBX protocol or the ingress policy type-length-value (TLV), the ingress policy configuration is ignored.

### Egress Queuing Policies

You can associate an egress policy map with an Ethernet interface to guarantee the bandwidth for the specified traffic class or to configure the egress queues.

The bandwidth allocation limit applies to all traffic on the interface including any FCoE traffic.

Each Ethernet interface supports up to eight queues, one for each system class. The queues have the following default configuration:

- In addition to these queues, control traffic that is destined for the CPU uses strict priority queues. These queues are not accessible for user configuration.

- FCoE traffic (traffic that maps to the FCoE system class) is assigned a queue. This queue uses weighted round-robin (WRR) scheduling with 50 percent of the bandwidth.
- Standard Ethernet traffic in the default drop system class is assigned a queue. This queue uses WRR scheduling with 100 percent of the bandwidth.

If you add a system class, a queue is assigned to the class. You must reconfigure the bandwidth allocation on all affected interfaces. Bandwidth is not dedicated automatically to user-defined system classes.

You can configure one strict priority queue. This queue is serviced before all other queues except the control traffic queue (which carries control rather than data traffic).

To increase the burst absorption capability run the **debug hard int multicast-port-prune-threshold** command.

## Buffering and Queue Limits on the Cisco Nexus 5500 Platform

On the Cisco Nexus device, the packet buffer per port is 640KB. 16MB in ingress across all ports per ASIC. 9MB in egress across all ports per ASIC.

The following default buffer allocations per port exist for the Cisco Nexus device:

**Table 1: Cisco Nexus Device Platform Default Buffer Allocations Per Port**

Traffic Class	Ingress Buffer (KB)
Class-fcoe at 300M for a 10G Port	161.25
Class-fcoe at 300M for a 40G Port	182.5
Class-fcoe at 3000M for a 10G Port	412.5
Class-fcoe at 3000M for a 40G Port	1300
User-defined no-drop at 300M for a 10G Port	160
User-defined no-drop at 300M for a 40G Port	181.25
User-defined no-drop at 3000M for a 10G Port	411.875
User-defined no-drop at 3000M for a 40G Port	1298.125

The default buffer allocation varies depending on the type of class. For example, if you create a regular tail drop traffic class the default allocation is 22.7KB, unless you specify a larger size using the **queue-limit** command.

To increase the ingress buffer space available to a user-created qos-group, from a network-qos policy-map, use the **queue-limit** command.

In addition to ingress buffer allocated for each user-created qos-group there is an additional 29.76KB buffer required at egress for each qos-group.

With the default QoS configuration, all of the available buffer (470KB) is allocated to the class-default. When you create a new qos-group, the buffer required for the new qos-group will be taken away from class-default.

The amount of buffer that is left for class-default equals 470 minus the ingress buffer used by other qos-groups minus 29.76KB and times the number of qos-groups.

**Note**

Each new class requires an additional 29.76KB, so the exact amount of buffer that is left in the class default equals 470 minus the buffer used by other qos-groups minus 29.76KB times the number of qos-groups.

The default QoS policy for the Cisco Nexus device does not create class-fcoe and does not reserve buffer and qos-group for FCoE traffic.

The **show queuing interface** command can display the amount of ingress buffer allocated for each qos-group

## Information About Flow Control

### Link-Level Flow Control

IEEE 802.3x link-level flow control allows a congested receiver to communicate a transmitter at the other end of the link to pause its data transmission for a short period of time. The link-level flow control feature applies to all the traffic on the link.

The transmit and receive directions are separately configurable. By default, link-level flow control is disabled for both directions.

On the Cisco Nexus device, Ethernet interfaces do not automatically detect the link-level flow control capability. You must configure the capability explicitly on the Ethernet interfaces.

On each Ethernet interface, the switch can enable either priority flow control or link-level flow control (but not both).

### Priority Flow Control

Priority flow control (PFC) allows you to apply pause functionality to specific classes of traffic on a link instead of all the traffic on the link. PFC applies pause functionality based on the IEEE 802.1p CoS value. When the switch enables PFC, it communicates to the adapter which CoS values to apply the pause.

Ethernet interfaces use PFC to provide lossless service to no-drop system classes. PFC implements pause frames on a per-class basis and uses the IEEE 802.1p CoS value to identify the classes that require lossless service.

In the switch, each system class has an associated IEEE 802.1p CoS value that is assigned by default or configured on the system class. If you enable PFC, the switch sends the no-drop CoS values to the adapter, which then applies PFC to these CoS values.

The default CoS value for the FCoE system class is 3. This value is configurable.

By default, the switch negotiates to enable the PFC capability. If the negotiation succeeds, PFC is enabled and link-level flow control remains disabled regardless of its configuration settings. If the PFC negotiation fails, you can either force PFC to be enabled on the interface or you can enable IEEE 802.x link-level flow control.

If you do not enable PFC on an interface, you can enable IEEE 802.3X link-level pause.



Note

Ensure that pause no-drop is configured on a class map for link-level pause.

By default, link-level pause is disabled.

# Configuring Queuing

## Configuring the Queue Limit for a Specified Fabric Extender

At the Fabric Extender configuration level, you can control the queue limit for a specified Fabric Extender for egress direction (from the network to the host). You can use a lower queue limit value on the Fabric Extender to prevent one blocked receiver from affecting traffic that is sent to other noncongested receivers ("head-of-line blocking"). A higher queue limit provides better burst absorption and less head-of-line blocking protection. You can use the **no** form of this command to allow the Fabric Extender to use all available hardware space.



Note

At the system level, you can set the queue limit for Fabric Extenders by using the **fex queue-limit** command. However, configuring the queue limit for a specific Fabric Extender will override the queue limit configuration set at the system level for that Fabric Extender.

You can specify the queue limit for the following Fabric Extenders:

- Cisco Nexus 2148T Fabric Extender (48x1G 4x10G SFP+ Module)
- Cisco Nexus 2224TP Fabric Extender (24x1G 2x10G SFP+ Module)
- Cisco Nexus 2232P Fabric Extender (32x10G SFP+ 8x10G SFP+ Module)
- Cisco Nexus 2248T Fabric Extender (48x1G 4x10G SFP+ Module)
- Cisco Nexus N2248TP-E Fabric Extender (48x1G 4x10G Module)
- Cisco Nexus N2348UPQ Fabric Extender (48x10G SFP+ 6x40G QSFP Module)

Procedure

	Command or Action	Purpose
Step 1	switch# <b>configure terminal</b>	Enters global configuration mode.
Step 2	switch(config)# <b>fex <i>fex-id</i></b>	Specifies the Fabric Extender and enters the Fabric Extender mode.
Step 3	switch(config-fex)# <b>hardware <i>fex_card_type</i> queue-limit <i>queue-limit</i></b>	Configures the queue limit for the specified Fabric Extender. The queue limit is specified in bytes. The range is from 81920 to 652800 for a Cisco Nexus 2148T Fabric Extender and from 2560 to 652800 for all other supported Fabric Extenders.

This example shows how to restore the default queue limit on a Cisco Nexus 2248T Fabric Extender:

```
switch# configure terminal
switch(config-if)# fex 101
switch(config-fex)# hardware N2248T queue-limit 327680
```

This example shows how to remove the queue limit that is set by default on a Cisco Nexus 2248T Fabric Extender:

```
switch# configure terminal
switch(config)# fex 101
switch(config-fex)# no hardware N2248T queue-limit 327680
```

## Configuring No-Drop Buffer Thresholds

You can configure the no-drop buffer threshold settings for 3000m lossless Ethernet.



### Note

To achieve lossless Ethernet for both directions, the devices connected to the device must have the similar capability. The default buffer and threshold value for the no-drop can ensure lossless Ethernet for up to 300 meters.

### Procedure

	Command or Action	Purpose
<b>Step 1</b>	switch# <b>configure terminal</b>	Enters global configuration mode.
<b>Step 2</b>	switch(config)# <b>policy-map type network-qos</b> <i>policy-map name</i>	Enters policy-map network-qos class mode and identifies the policy map assigned to the type network-qos policy map.
<b>Step 3</b>	switch(config-pmap-nq)# <b>class type network-qos</b> <i>class-map name</i>	References an existing network QoS class map in a policy map and enters class mode.
<b>Step 4</b>	switch(config-pmap-nq-c)# <b>pause no-drop buffer-size</b> <i>buffer-size</i> <b>pause-threshold</b> <i>xoff-size</i> <b>resume-threshold</b> <i>xon-size</i>	Specifies the buffer threshold settings for pause and resume for 3000m lossless Ethernet: <ul style="list-style-type: none"> <li>• <b>buffer-size</b>—Buffer size for ingress traffic, in bytes. Valid values are from 10240 to 490880.</li> <li>• <b>pause-threshold</b>—Specifies the buffer limit at which the port pauses the peer.</li> <li>• <b>xoff-size</b>—Buffer limit for pausing, in bytes. Valid values are 0 to 490880.</li> <li>• <b>resume-threshold</b>—Specifies the buffer limit at which the port resumes the peer.</li> <li>• <b>xon-size</b>—Buffer limit at which to resume, in bytes. Valid values are 0 to 490880.</li> </ul>

	Command or Action	Purpose
<b>Step 5</b>	switch(config-pmap-nq-c)# <b>no pause</b> <b>no-drop buffer-size</b> <i>buffer-size</i> <b>pause-threshold</b> <i>xoff-size</i> <b>resume-threshold</b> <i>xon-size</i>	(Optional) Removes the buffer threshold settings for pause and resume for 3000m lossless Ethernet.
<b>Step 6</b>	switch(config-pmap-nq-c)# <b>exit</b>	Exits class mode.
<b>Step 7</b>	switch(config-pmap-nq)# <b>exit</b>	Exits policy-map network-qos mode.

This example shows how to configure the no-drop buffer threshold for 3000 meters.

```
switch(config-pmap-nq)# policy-map type network-qos nqos_policy
switch(config-pmap-nq)# class type network-qos nqos_class
switch(config-pmap-nq-c)# pause no-drop buffer-size 152000 pause-threshold 103360
resume-threshold 83520
switch(config-pmap-nq-c)# exit
switch(config-pmap-nq)# exit
switch(config)# exit
switch#
```

## Configuring the Buffer Threshold for the Cisco Nexus 2148T Fabric Extender

In the Fabric Extender configuration mode, you can configure the buffer threshold for the Cisco Nexus 2148T Fabric Extender. The buffer threshold sets the consumption level of input buffers before an indication is sent to the egress queue to start observing the tail drop threshold. If the buffer usage is lower than the configured buffer threshold, the tail drop threshold is ignored.

### Procedure

	Command or Action	Purpose
<b>Step 1</b>	switch# <b>configure terminal</b>	Enters global configuration mode.
<b>Step 2</b>	switch(config)# <b>fex</b> <i>fex-id</i>	Specifies the Fabric Extender and enters the Fabric Extender mode.
<b>Step 3</b>	switch(config-fex)# <b>hardware</b> <b>N2148T buffer-threshold</b> <i>buffer limit</i>	Configures the buffer threshold for the Cisco Nexus 2148T Fabric Extender. The buffer threshold is specified in bytes. The range is from 81920 to 316160 for the Cisco Nexus 2148T Fabric Extender.

This example shows how to restore the default buffer threshold on the Cisco Nexus 2148T Fabric Extender:

```
switch# configure terminal
switch(config)# fex 101
switch(config-fex)# hardware N2148T buffer-threshold 163840
```

This example shows how to remove the default buffer threshold on the Cisco Nexus 2148T Fabric Extender:

```
switch# configure terminal
switch(config)# fex 101
switch(config-fex)# no hardware N2148T buffer-threshold
```

## Enabling Virtual Output Queuing Limits for Unicast Traffic on the Cisco Nexus Device

You can enable the Virtual Output Queuing (VOQ) limit for unicast traffic. To alleviate congestion and blocking, use VOQ to prevent one blocked receiver from affecting traffic that is sent to other noncongested blocking receivers.

### Procedure

	Command or Action	Purpose
<b>Step 1</b>	switch# <b>configure terminal</b>	Enters configuration mode.
<b>Step 2</b>	switch(config)# <b>hardware unicast voq-limit</b>	Enables the VOQ limit for unicast traffic. The default is disabled.
<b>Step 3</b>	switch(config)# <b>no hardware unicast voq-limit</b>	Disables the VOQ limit for unicast traffic.

This example shows how to enable the VOQ limits for unicast packets on a switch:

```
switch(config)# hardware unicast voq-limit  
switch(config)#
```

## Configuring Flow Control

### Link-Level Flow Control

IEEE 802.3x link-level flow control allows a congested receiver to communicate a transmitter at the other end of the link to pause its data transmission for a short period of time. The link-level flow control feature applies to all the traffic on the link.

The transmit and receive directions are separately configurable. By default, link-level flow control is disabled for both directions.

On the Cisco Nexus device, Ethernet interfaces do not automatically detect the link-level flow control capability. You must configure the capability explicitly on the Ethernet interfaces.

On each Ethernet interface, the switch can enable either priority flow control or link-level flow control (but not both).

### Configuring Priority Flow Control

By default, Ethernet interfaces negotiate PFC with the network adapter using the DCBX protocol. When PFC is enabled, PFC is applied to traffic that matches the CoS value configured for the no-drop class.

You can override the negotiation result by forcing the interface to enable PFC.

Beginning in Cisco NX-OS Release 5.0(3)N1(1), you can configure priority flow control on a Layer 2 or Layer 3 interface.

**Note**

Use the **no switchport** command to configure an interface as a Layer 3 interface

**Procedure**

	Command or Action	Purpose
<b>Step 1</b>	switch# <b>configure terminal</b>	Enters global configuration mode.
<b>Step 2</b>	switch(config)# <b>interface</b> <i>type slot/port</i>	Specifies the interface to be changed.  <b>Note</b> If this is a 10G breakout port, the <i>slot/port</i> syntax is <i>QSFP-module/port</i> .
<b>Step 3</b>	switch(config-if)# <b>no switchport</b>	(Optional) Selects a Layer 3 interface.
<b>Step 4</b>	switch(config-if)# <b>priority-flow-control mode {auto   on}</b>	Sets PFC mode for the selected interface.  Specifies auto to negotiate PFC capability. This is the default.  Specifies on to force-enable PFC.
<b>Step 5</b>	switch(config-if)# <b>no priority-flow-control mode on</b>	(Optional) Disables the PFC setting for the selected interface.

This example shows how to force-enable PFC on an interface:

```
switch# configure terminal
switch(config)# interface ethernet 1/2
switch(config-if)# priority-flow-control mode on
```

This example shows how to force-enable PFC on a Layer 3 interface:

```
switch# configure terminal
switch(config)# interface ethernet 1/5
switch(config-if)# no switchport
switch(config-if)# priority-flow-control mode on
```

## Configuring Link-Level Flow Control

By default, LLC on Ethernet interfaces is disabled. You can enable LLC for the transmit and receive directions. You can configure flow control on a Layer 2 or Layer 3 interface.

**Note**

Use the **no switchport** command to configure an interface as a Layer 3 interface.



**Procedure**

	Command or Action	Purpose
<b>Step 1</b>	switch# <b>configure terminal</b>	Enters global configuration mode.
<b>Step 2</b>	switch(config)# <b>interface</b> <i>type slot/port</i>	Specifies the interface to be changed. <b>Note</b> If this is a 10G breakout port, the <i>slot/port</i> syntax is <i>QSFP-module/port</i> .
<b>Step 3</b>	switch(config-if)# <b>no switchport</b>	(Optional) Selects a Layer 3 interface.
<b>Step 4</b>	switch(config-if)# <b>flowcontrol</b> [ <b>receive</b> { <b>on</b>   <b>off</b> }] [ <b>send</b> { <b>on</b>   <b>off</b> }]	Enables LLC for the selected interface. Set <b>receive</b> and/or <b>send on</b> or <b>off</b> .
<b>Step 5</b>	switch(config-if)# <b>no flowcontrol</b> [ <b>receive</b> { <b>on</b>   <b>off</b> }] [ <b>send</b> { <b>on</b>   <b>off</b> }]	(Optional) Disables LLC for the selected interface.

This example shows how to enable LLC on an interface:

```
switch# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
switch(config)# interface e1/48
switch(config-if)# flowcontrol receive on
switch(config-if)# flowcontrol send on
```

This example shows how to enable LLC on a Layer 3 interface:

```
switch# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
switch(config)# interface e1/48
switch(config-if)# no switchport
switch(config-if)# flowcontrol receive on
switch(config-if)# flowcontrol send on
```

## Verifying the Queue and Flow Control Configurations

Use one of the following commands to verify the configuration:

Command	Purpose
<b>show queuing interface</b> [ <i>interface slot/port</i> ]	Displays the queue configuration and statistics. <b>Note</b> If this is a 10G breakout port, the <i>slot/port</i> syntax is <i>QSFP-module/port</i> .
<b>show interface flowcontrol</b> [ <i>module number</i> ]	Displays the detailed listing of the flow control settings on all interfaces.
<b>show interface</b> [ <i>interface slot/port</i> ] <b>priority-flow-control</b> [ <i>module number</i> ]	Displays the priority flow control details for a specified interface. <b>Note</b> If this is a 10G breakout port, the <i>slot/port</i> syntax is <i>QSFP-module/port</i> .

Command	Purpose
<b>show wrr-queue cos-map</b> [ <i>var</i> ]	
<b>running-config ipqos</b>	Displays information about the running configuration for QoS.
<b>startup-config ipqos</b>	Displays information about the startup configuration for QoS.