Configuring Priority Flow Control

- About Priority Flow Control, on page 1
- Licensing Requirements for Priority Flow Control, on page 3
- Prerequisites for Priority Flow Control, on page 4
- Guidelines and Limitations for Priority Flow Control, on page 4
- Default Settings for Priority Flow Control, on page 7
- Configuring Priority Flow Control, on page 7
- Enabling Priority Flow Control on a Traffic Class, on page 8
- Configuring a Link Level Flow Control Watchdog and Priority Flow Control Watchdog, on page 12
- Configuring Pause Buffer Thresholds and Queue Limit Using Ingress Queuing Policy, on page 16
- Verifying the Priority Flow Control Configuration, on page 18
- Configuration Examples for Priority Flow Control, on page 19

About Priority Flow Control

Priority Flow Control (PFC) is used in lossless Ethernet to control the flow of data from a link partner for specific traffic priorities or classes specified as 'no-drop'. PFC Pause frames are transmitted to the link partner when certain queue thresholds are reached for a specific class or priority. PFC Pause frames are only local to the specific link but when traffic is suspended the congestion can cause PFC Pause frames to be generated on other links spreading the congestion. This can cause traffic for the priority or class to be suspended throughout the entire network for a time.

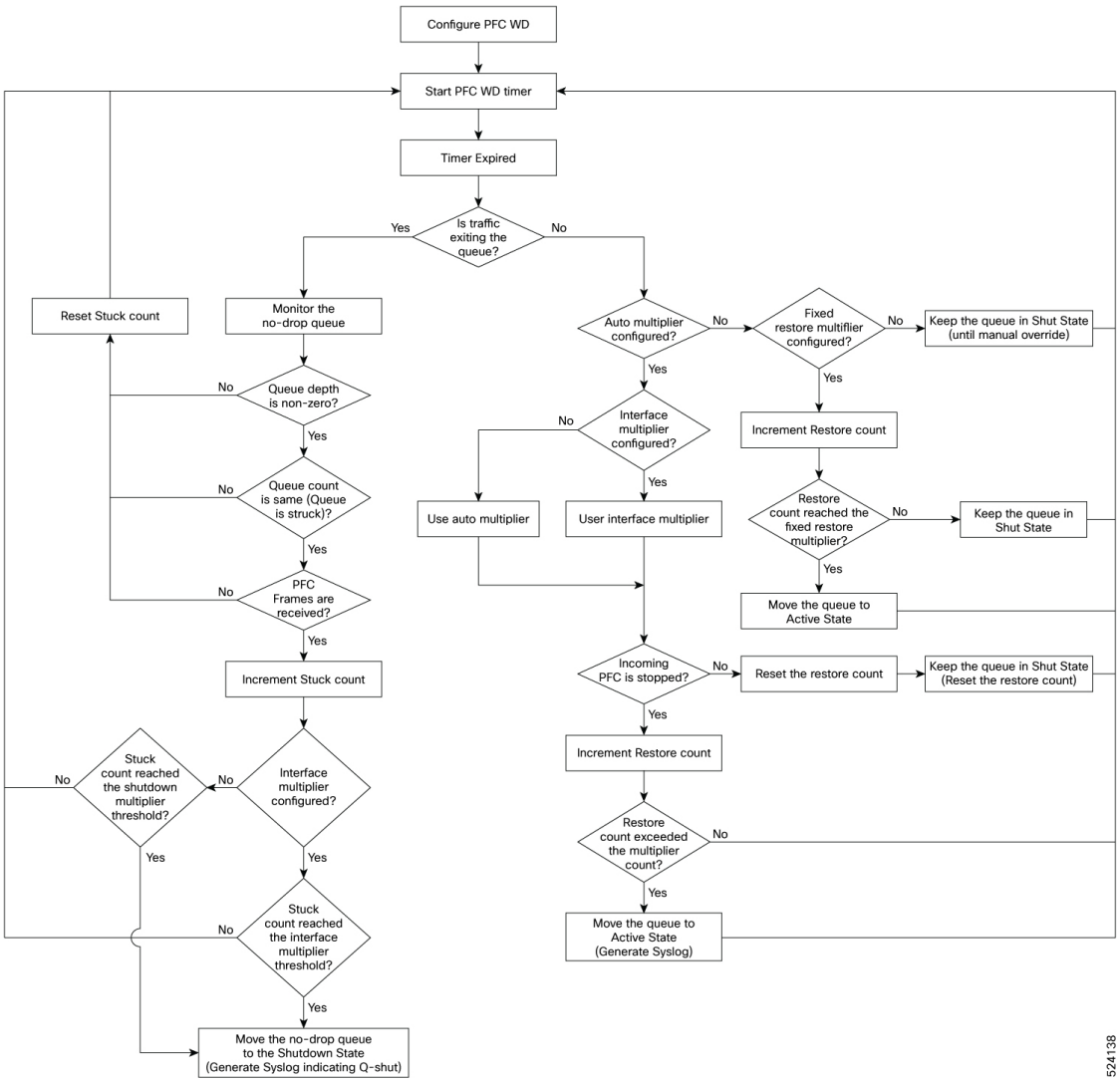
About Priority Flow Control Watchdog

Priority Flow Control Watchdog (PFCWD) is a mechanism designed to detect and resolve any PFC storms (queue-stuck condition) in the network. You can configure a PFC watchdog interval to detect whether packets in a no-drop queue are drained within a specified time period. When this time period is exceeded, all outgoing packets are dropped on interfaces that match the PFC queue that is not being drained.

Note The PFC watchdog feature is supported only for no-drop queues.

Workflow of Priority Flow Control Watchdog

- Monitors the PFC-enabled queues (no-drop queues) to identify the reception of an excessive number of PFC pause frames in a given interval (Watchdog interval).
- Monitors when an excessive number of PFC frames are received and traffic on the corresponding queues is halted for a specified time interval (auto + fixed multiplier).
- Initiates the shutdown timer and changes the queue's state to **wait-to-shutdown**.
- Drops all data packets when the queue transitions to a **shut** state if interface multiplier time exceeds (if the interface multiplier is configured) or shutdown multiplier timer expire exceeds (if the interface multiplier is not configured).
- Checks the queue for PFC frames and whether the traffic in the queue is still stuck at regular intervals (poll timer of 100ms) during the shutdown interval.
  - If traffic is stuck in the queue as PFC packets continue to arrive, the queue stays in the drop or shutdown state.
  - If the traffic is not stuck because the queue didn't receive any PFC frames, the queue reverts to the monitored state.
- Checks if the queue is stuck because of PFC frames when the traffic is no longer stuck at regular intervals, the auto-restore timer starts.
  - If the queue receives PFC frames during the last auto-restore interval (poll timer * auto-restore multiplier), the auto-restore timer (secs) is reset at its expiration.
  - If the queue receives no PFC frames during the last auto-restore interval, the watchdog module restores the queue, and traffic resumes.



Licensing Requirements for Priority Flow Control

The following table shows the licensing requirements for this feature:

Product	License Requirement
Cisco NX-OS	The PFC feature does not a require license. Any feature not included in a license package is bundled with the NX-OS image and is provided at no extra charge to you. For a complete explanation of the Cisco NX-OS licensing scheme, see the Cisco NX-OS Licensing Guide.

Prerequisites for Priority Flow Control

PFC has the following prerequisites:

- You must be familiar with using modular QoS CLI.
- You are logged on to the device.

Guidelines and Limitations for Priority Flow Control

PFC has the following configuration guidelines and limitations:

- PFC is not supported on the Cisco Nexus 9508 switch.
- The **show** commands with the **internal** keyword are not supported.
- Adding the "pause buffer size threshold" configuration is optional for cable lengths that are less than 100 meters and it does need not to be configured.
- Input queuing policy maps cannot have pause buffer and priority/bandwidth together.
- For cable lengths greater than 100m, the "pause buffer size threshold" configuration is mandatory and it is required as part of the QoS policy configuration.
- If PFC is enabled on a port or a port channel, it does not cause a port flap.
- PFC configuration enables PFC in both the send (Tx) and receive (Rx) direction.
- Configuration time quanta of the pause frames is not supported.
- You can configure a PFC watchdog interval to detect whether packets in a no-drop queue are being drained within a specified time period. When the time period is exceeded, all outgoing packets are dropped on interfaces that match the PFC queue that is not being drained. This feature is supported only for Cisco Nexus 9200 Series switches, Cisco Nexus 93108TC-EX, and 93180YC-EX switches, and Cisco Nexus 9508 switches with the X9732C-EX line cards.

  This feature is supported on Cisco Nexus 9508 switches with N9K-X9636PQ line cards and Cisco Nexus 3164Q switches.
- The configuration does not support pausing selected streams that are mapped to a particular traffic-class queue. All flows that are mapped to the class are treated as no-drop. It blocks out scheduling for the entire queue, which pauses traffic for all the streams in the queue. To achieve lossless service for a no-drop class, we recommend that you have only the no-drop class traffic on the queue.
- When a no-drop class is classified based on 802.1p CoS x and assigned an internal priority value (qos-group) of y, we recommend that you use the internal priority value x to classify traffic on 802.1p CoS only, and not on any other field. The packet priority that is assigned is x if the classification is not based on CoS, which results in packets of internal priority x and y to map to the same priority x.
- The PFC feature supports up to three no-drop classes of any maximum transmission unit (MTU) size. However, there is a limit on the number of PFC-enabled interfaces, based on the following factors:
  - MTU size of the no-drop class
  - Number of 10G and 40G ports

- You can define the upper limit of any MTU in the system using the **systemjumbomtu** command. The MTU range is from 1500 to 9216 bytes, and the default is 9216 bytes.
- The interface QoS policy takes precedence over the system policy. PFC priority derivation also happens in the same order.
- Ensure that you apply the same interface-level QoS policy on all PFC-enabled interfaces for both ingress and egress.

Caution Irrespective of the PFC configuration, we recommend that you stop traffic before applying or removing a queuing policy that has strict-priority levels at the interface level or the system level.

- To achieve end-to-end lossless service over the network, we recommend that you enable PFC on each interface through which the no-drop class traffic flows (Tx/Rx).
- We recommend that you change the PFC configuration when there is no traffic. Otherwise, packets already in the Memory Management Unit (MMU) of the system may not get the expected treatment.
- We recommend that you use default buffer sizes for no-drop classes or configure different input queuing policies suitable to 10G and 40G interfaces and the no-drop class MTU size. If the buffer size is specified through the CLI, it allocates the same buffer size for all ports irrespective of the link speed and MTU size. Applying the same pause buffer-size on 10G and 40G interfaces is not supported.
- Do not enable WRED on a no-drop class because it results in drops in the egress queue.
- Dynamic load balancing cannot be enabled for internal links with PFC. Disable DLB and enable RTAG7 load-balancing for internal links with the port-channel load-balance internal rtag7 command.
- The dynamic load balancing (DLB) based hashing scheme is enabled by default on all internal links of a linecard. When DLB is enabled, no-drop traffic may experience an out-of-order packet delivery when congestion on internal links occurs and PFC is applied. If applications on the system are sensitive to out-of-order delivery, you can adjust for this event by disabling DLB at the qos-group level. Disable DLB by using the **set dlb-disable** action in the QoS policy-maps and the **set qos-group** action for no-drop classes.

  In the following example, assume that qos-group 1 is a no-drop class. DLB is disabled for this no-drop class by adding the **set dlb-disable** action and the **set qos-group** action.

```
switch(config)# policy-map p1
switch(config-pmap-qos)# class c1
switch(config-pmap-c-qos)# set qos-group 1
switch(config-pmap-c-qos)# set dlb-disable
switch(config-pmap-c-qos)# end
switch# show policy-map p1


  Type qos policy-maps
  ====================

  policy-map type qos p1
    class  c1
      set qos-group 1
      set dlb-disable
```

Note The following Cisco Nexus platform switches do not support the **set-dlb-disable** command:

- Cisco Nexus 9200-series platform switches
- Cisco Nexus 9300-EX/FX/FX2 platform switches
- Cisco Nexus 9500-series platform switches with -EX and -FX line cards

- For VLAN-tagged packets, priority is assigned based on the 802.1p field in the VLAN tag and takes precedence over the assigned internal priority (qos-group). DSCP or IP access-list classification cannot be performed on VLAN-tagged frames.
- For non VLAN-tagged frames, priority is assigned based on the **set qos-group** action provided by the ingress QoS policy. Classification is based on a QoS policy-allowed match condition such as precedence, DSCP, or access-list. Ensure that the **pfc-cos** value that is provided in the network-qos policy for this class is the same as the **qos-group** value in this case.
- PFC is not supported for the N9K-X9408PC-CFP2 line card on Cisco Nexus 9500 Series switches.
- Link level flow control and PFC are supported on Cisco Nexus 9300 Series switches and line cards that contain the ALE (Application Leaf Engine).
- PFC on mode is used to support the hosts that support PFC but do not support the Data Center Bridging Capability Exchange Protocol (DCBXP).
- Only an exact match of the no-drop CoS is considered as a successful negotiation of PFC by the DCBXP.
- Beginning with Cisco NX-OS Release 7.0(3)I3(1), DCBXP is supported on the following Cisco Nexus switches:
  - Nexus 9332PQ switch
  - Nexus 9372PX switch
  - Nexus 9372PX-E switch
  - Nexus C9396PX switch
  - Nexus 9500 Series switches with the following line cards:
    - X9432PQ
    - X9464PX
    - X9464TX
    - X9536PQ
    - X9564PX
    - X9564TX
    - X9636PQ
- DCBXP is supported on Cisco Nexus 9200 and 9300-EX Series switches.

- The **no lldp tlv-select dcbxp** command is enhanced so that PFC is disabled for interfaces on both sides of back-to-back switches.
- BUM traffic is not supported in no-drop PFC queues. Avoid marking multicast traffic as no-drop and sending it to these queues.

Default Settings for Priority Flow Control

Table 1: Default PFC Setting

Parameter	Default
PFC	Auto

Configuring Priority Flow Control

You can configure PFC on a per-port basis to enable the no-drop behavior for the CoS as defined by the active network QoS policy. PFC can be configured in one of these modes:

- auto—Enables the no-drop CoS values to be advertised by the DCBXP and negotiated with the peer. A successful negotiation enables PFC on the no-drop CoS. Any failures because of a mismatch in the capability of peers causes the PFC not to be enabled. (Cisco NX-OS Release 7.0(3)I3(1) and later)
- on—Enables PFC on the local port regardless of the capability of the peers.
- off—Disables PFC on the local port.

Note You can use the **priority-flow-control override-interface mode off** command to globally disable PFC on all interfaces regardless of the current interface configuration. This command, which is meant to be used during troubleshooting, allows you to quickly disable PFC without having to disable PFC on each interface. It is supported beginning with Cisco NX-OS Release 7.0(3)I4(2) and only for Cisco Nexus 9200 platform switches, Cisco Nexus 93108TC-EX and 93180YC-EX switches, and Cisco Nexus 9508 switches with the Cisco Nexus 9732C-EX line card.

Beginning with Cisco NX-OS Release 7.0(3)I4(5), this feature is supported on Cisco Nexus 9508 switches with Cisco Nexus 9636PQ line cards and Cisco Nexus 3164Q switches.

SUMMARY STEPS

1. **configure terminal**
2. **interface** *type slot/port*
3. **priority-flow-control mode** [**auto** | **off** |**on**]
4. (Optional) **show interface priority-flow-control**

DETAILED STEPS

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**interface** *type slot/port*
	Example:
	`switch(config)# interface ethernet 2/5`
	`switch(config-if)#`
		Enters interface mode on the interface specified.

Step 3	**priority-flow-control mode** [**auto** | **off** |**on**] **Example:** `switch(config-if)# priority-flow-control mode on`	Sets PFC to the on mode.
Step 4	(Optional) **show interface priority-flow-control** **Example:** `switch# show interface priority-flow-control`	Displays the status of PFC on all interfaces.

Enabling Priority Flow Control on a Traffic Class

You can enable PFC on a particular traffic class.

SUMMARY STEPS

1. **configure terminal**
2. **class-map type qos match** { **all** | **any** } *class-name*
3. **match cos** *cos-value*
4. **match dscp** *dscp-value*
5. **exit**
6. **policy-map type qos** *policy-name*
7. **class** *class-name*
8. **set qos-group** *qos-group-value*
9. **exit**
10. **exit**
11. **policy-map type network-qos** *policy-name*
12. **class type network-qos** *class-name*
13. **pause pfc-cos** *value* [**receive**]
14. **exit**
15. **exit**
16. **system qos**

17. **service-policy type network-qos** *policy-name*
18. **exit**
19. **interface ethernet** *slot* / *number*
20. **priority-flow-control mode** { **auto** | **on** | **off** }
21. **service-policy type qos input** *policy-name*
22. **exit**

DETAILED STEPS

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**class-map type qos match** { **all** | **any** } *class-name*
	Example:
	`switch(config)# class-map type qos c1`
	`switch(config-cmap-qos)#`
		Creates a named object that represents a class of traffic. Class-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.
		match { **all** | **any** }: Default is **match all** (if multiple matching statements are present all of them must be matched).

Step 3	**match cos** *cos-value*
	Example:
	`switch(config-cmap-qos)# match cos 2`
	`switch(config-cmap-qos)#`
		Specifies the CoS value to match for classifying packets into this class. You can configure a CoS value in the range of 0 to 7.

Step 4	**match dscp** *dscp-value*
	Example:
	`switch(config-cmap-qos)# match dscp 3`
	`switch(config-cmap-qos)#`
		Specifies the DSCP value to match for classifying packets into this class. You can configure a DSCP value in the range of 0 to 63 or the listed values.

Step 5	**exit**
	Example:
	`switch(config-cmap-qos)# exit`
	`switch(config)#`
		Exits class-map mode and enters global configuration mode.

Step 6	**policy-map type qos** *policy-name*
	Example:
	`switch(config)# policy-map type qos p1`
	`switch(config-pmap-qos)#`
		Creates a named object that represents a set of policies that are to be applied to a set of traffic classes. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.

Step 7	**class** *class-name* **Example:**	Associates a class map with the policy map and enters the configuration mode for the specified system class.

	Command or Action	Purpose
	`switch(config-pmap-qos)# class c1` `switch(config-pmap-c-qos)#`	**Note** The associated class map must be the same type as the policy map type.

Step 8	**set qos-group** *qos-group-value*
	Example:
	`switch(config-pmap-c-qos)# set qos-group 3`
	`switch(config-pmap-c-qos)#`
		Configures one or more qos-group values to match on for classification of traffic into this class map. There is no default value.

Step 9	**exit**
	Example:
	`switch(config-pmap-c-qos)# exit`
	`switch(config-pmap-qos)#`
		Exits the system class configuration mode and enters policy-map mode.

Step 10	**exit**
	Example:
	`switch(config-pmap-qos)# exit`
	`switch(config)#`
		Exits policy-map mode and enters global configuration mode.

Step 11	**policy-map type network-qos** *policy-name*
	Example:
	`switch(config)# policy-map type network-qos pfc-qos`
	`switch(config-pmap-nqos)#`
		Creates a named object that represents a set of policies that are to be applied to a set of traffic classes. Policy-map names can contain alphabetic, hyphen, or underscore characters, are case sensitive, and can be up to 40 characters.

Step 12	**class type network-qos** *class-name*
	Example:
	`switch(config-pmap-nqos)# class type network-qos nw-qos3`
	`switch(config-pmap-nqos-c)#`
		Associates a class map with the policy map, and enters the configuration mode for the specified system class.
		Note
		The associated class map must be the same type as the policy map type.

Step 13	**pause pfc-cos** *value* [**receive**]
	Example:
	`switch(config-pmap-nqos-c)# pause pfc-cos 3 receive`
	`switch(config-pmap-nqos-c)#`
		PFC sends a pause frame that indicates which CoS value needs to be paused. Only PFC receive is enabled for the list of PCF CoS values.
		receive: When this optional keyword is used, PFC only receives and honors pause frames. PFC will never send pause frames. This is known as "Asymmetric PFC".
		Note
		Although not required, the **pause pfc-cos** *value* should match the *qos-group-value* in the **set qos-group** command. See the **set qos-group** command in steps 8 above.

Step 14	**exit** **Example:**	Exits configuration mode and enters policy-map mode.

	Command or Action	Purpose
	`switch(config-pmap-nqos-c)# exit` `switch(config-pmap-nqos)#`	

Step 15	**exit**
	Example:
	`switch(config-pmap-nqos)# exit`
	`switch(config)#`
		Exits policy-map mode and enters global configuration mode.

Step 16	**system qos**
	Example:
	`switch(config)# system qos`
	`switch(config-sys-qos)#`
		Enters system class configuration mode.

Step 17	**service-policy type network-qos** *policy-name* **Example:** `switch(config-sys-qos)# service-policy type network-qos pfc-qos`	Applies the policy map of type network-qos at the system level or to the specific interface.

Step 18	**exit**
	Example:
	`switch(config-sys-qos)# exit`
	`switch(config)#`
		Exits policy-map mode and enters global configuration mode.

Step 19	**interface ethernet** *slot* / *number*
	Example:
	`switch(config)# interface ethernet 1/1`
	`switch(config-if)#`
		Enters the ethernet interface configuration mode for the selected slot and chassis number.

Step 20	**priority-flow-control mode** { **auto** | **on** | **off** }
	Example:
	`switch(config-if)# priority-flow-control mode on`
	`switch(config-if)#`
		Enables the priority flow control policy for the interface.

Step 21	**service-policy type qos input** *policy-name* **Example:** `switch(config-if)# service-policy type qos input p1`	Adds classification to the interface ensuring that packets matching the previously configured CoS or DSCP values are classified in the correct QoS group.

Step 22	**exit**
	Example:
	`switch(config-if)# exit`
	`switch(config)#`
		Exits the ethernet interface mode and enters the global configuration mode.

Configuring a Link Level Flow Control Watchdog and Priority Flow Control Watchdog

Link Level Flow Control Watchdog (LLFCWD) is enabled globally by default. LLFCWD on an interface is enabled automatically when PFC and PFCWD are configured on the interface. If an LLFC packet is seen on a PFC/PFCWD configured interface that doesn't have LLFC configured, the LLFC watchdog is triggered.

PFCWD interval and PFCWD multiplier CLI commands are used to configure the LLFCWD interval and multipliers. Use this procedure to configure the LLFC watchdog interval and the multiplier used to restore the no-drop queue.

Note PFC watchdog is not supported on Cisco Nexus 9500 platform switches with Cisco Nexus 9400, 9500 and 9600 line cards, with the exception of Cisco Nexus 9636PQ line cards (that support the PFC watchdog feature).

Note Ingress drops provide statistics of PFC watchdog dropped packets on the front panel ports.

Note For Cisco Nexus 9200 platform switches, Cisco Nexus 9300-EX/FX/FX2 platform switches, and Cisco Nexus 9500 platform switches with -EXX97160YC-EX or -FX line cards, one of the following calculations is performed to determine when the queue is moved to the shutdown state:

If the interface multiplier is configured, the following calculation is used:

priority-flow-control watch-dog interval *value* * **priority-flow-control watch-dog internal-interface-multiplier** *multiplier*

If the interface multiplier is not configured, the watchdog shutdown multiplier is used instead:

priority-flow-control watch-dog interval *value* * **priority-flow-control watch-dog shutdown-multiplier** *multiplier*

Before you begin

Consider the following before configuring the Link Level Flow Control Watchdog Interval:

- Link Level Flow Control Watchdog is supported on the following Cisco Nexus 9000 Series platform switches and line cards:
  - N9K-C9232C
  - N9K-C9236C
  - N9K-C92304QC
  - N9K-X9736C-EX
  - N9K-X9732C-EX

- N9K-X9732C-EXM
- N9K-X97160YC-EX

- PFC must be enabled at the interface. PFCWD must be enabled at the interface and globally. LLFC shouldn't be configured on the same interface.

Note PFC watchdog uses a command to send a syslog message indicating that the queue is "stuck" (**priority-flow-control watch-dog-interval on disable-action**). If this command is invoked on a PFC interface, the queue isn't shut down but instead, the syslog message is generated. When the LLFC watchdog feature is enabled, and if a link level flow control packet is received on an interface, the queue is shut even with the **disable-action** command for PFC watchdog is enabled.

- Auto-restore and fixed restore should never be configured to 0.

SUMMARY STEPS

1. **configure terminal**
2. **priority-flow-control auto-restore multiplier** *value*
3. **priority-flow-control fixed-restore multiplier** *value*
4. **priority-flow-control watch-dog-interval** {**on** | **off**}
5. **priority-flow-control watch-dog interval** *value*
6. **priority-flow-control watch-dog shutdown-multiplier** *multiplier*
7. (Optional) **priority-flow-control watch-dog internal-interface-multiplier** *multiplier*
8. (Optional) **show queuing pfc-queue [interface** *interface-list***] [module** *module***] [detail]**
9. (Optional) **show queuing llfc-queue [interface** *interface-list***] [module** *module***] [detail]**
10. (Optional) **clear queuing pfc-queue [interface] [ethernet|ii] [intf-name]**
11. (Optional) **clear queuing llfc-queue [interface** *interface-list***] [module** *module***]**
12. (Optional) **priority-flow-control recover interface [ethernet|ii] [intf-name] [qos-group <0-7>]**

DETAILED STEPS

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**priority-flow-control auto-restore multiplier** *value*	Configures a value for the auto-restore multiplier, which is calculated by multiplying the set PFC WD interval time. The range is from 0 to 100. **Note**

	Command or Action	Purpose
		The auto-restore multiplier should never be configured to 0. When the LLFC watchdog no-drop queue is restored, a system logging message entry is created to record the conditions of the queue. The following is an example of the message: `Error Message TAHUSD-SLOT#-2-TAHUSD_SYSLOG_LLFCWD_QUEUE_RESTORED : [chars] Description : NO DROP Queue Restored due to LLFC WatchDog timer expiring message` This command is applicable for both LLFCWD and PFCWD.
Step 3	**priority-flow-control fixed-restore multiplier** *value*	Configures a value for the PFC fixed-restore multiplier.

Step 4	**priority-flow-control watch-dog-interval** {**on** | **off**}
	Example:
	`switch(config)# priority-flow-control watch-dog-interval on`
		Globally enables or disables the PFC watchdog interval for all interfaces. This command should be configured at global and also at an interface.
		See the following example of the command configured at global:
		`switch(config)# priority-flow-control watch-dog-interval on`
		See the following example of the command configured at an interface:
		`switch(config)# interface ethernet 7/5`
		`switch(config-if)# priority-flow-control watch-dog-interval on`
		Note
		You can use this same command in interface configuration mode to enable or disable the PFC watchdog interval for a specific interface.
		This command is applicable for both LLFCWD and PFCWD.

Step 5	**priority-flow-control watch-dog interval** *value* **Example:** `switch(config)# priority-flow-control watch-dog interval 200`	Specifies the watchdog interval value of all queues and ports for which this configuration is enabled. The range is from 100 to 1000 milliseconds. **Note** This command is applicable for both LLFCWD and PFCWD.
Step 6	**priority-flow-control watch-dog shutdown-multiplier** *multiplier* **Example:**	Specifies when to declare the PFC queue as stuck shutdown multiplier * poll interval. The range is from 1 to 10, and the default value is 1.

	Command or Action	Purpose
	`switch(config)# priority-flow-control watch-dog shutdown-multiplier 5`	
Step 7	(Optional) **priority-flow-control watch-dog internal-interface-multiplier** *multiplier* **Example:** `switch(config)# priority-flow-control watch-dog internal-interface-multiplier 5`	Configures a PFC watchdog poll-interval multiplier for HiGig™ interfaces. The range is from 0 to 10, and the default value is 2. A value of 0 disables this feature on HiGig™ interfaces. **Note** This command is only applicable for EoR switches.
Step 8	(Optional) **show queuing pfc-queue [interface** *interface-list***] [module** *module***] [detail]** **Example:** `switch(config)# sh queuing pfc-queue interface ethernet 1/1 detail`	Displays the PFCWD statistics.
Step 9	(Optional) **show queuing llfc-queue [interface** *interface-list***] [module** *module***] [detail]** **Example:** `switch(config)# show queuing llfc-queue interface ethernet 1/1 detail`	Displays the LLFCWD statistics. See the output example at the end of this procedure.
Step 10	(Optional) **clear queuing pfc-queue [interface] [ethernet|ii] [intf-name]** **Example:** `switch(config)# clear queuing pfc-queue interface ethernet 1/1`	Clears the environment variable PFCWD statistics.
Step 11	(Optional) **clear queuing llfc-queue [interface** *interface-list***] [module** *module***]** **Example:** `switch(config)# clear queuing llfc-queue interface ethernet 1/1`	Clears the LLFCWD queue statistics.
Step 12	(Optional) **priority-flow-control recover interface [ethernet|ii] [intf-name] [qos-group <0-7>]** **Example:** `switch# priority-flow-control recover interface ethernet 1/1 qos-group 3`	Recovers the interface manually.

Example

Beginning with Cisco NX-OS Release 7.0(3)I6(1), on Cisco Nexus 9200, 9300, 9300-EX, and 9500 platform switches, using the detail option, you can account for Ingress drops.

```
| QOS GROUP 1 [Active] PFC [YES] PFC-COS [1]
+-------------------------------------------------+
|                                |  Stats         |
```

```
+-----------------------------------------------------+
|                        Shutdown|                   0|
|                        Restored|                   0|
|             Total pkts drained|                   0|
|             Total pkts dropped|                   0|
|   Total pkts drained + dropped|                   0|
|          Aggregate pkts dropped|                   0|
|     Total Ingress pkts dropped|                   0| ===>>>>>Ingress
| Aggregate Ingress pkts dropped|                   0|===>>>>Ingress
+-------------------------------+
```

The following example shows detail output of the **show queuing llfc-queue** command for an Ethernet 1/1 interface:

```
switch# show queuing llfc-queue interface 1/1 detail

slot  1
=======
+-----------------------------------------------------+
Global watch-dog interval [Enabled]
+-----------------------------------------------------+
+-----------------------------------------------------+
Global LLFC watchdog configuration details

LLFC watchdog poll interval                : 100 ms
LLFC watchdog auto-restore multiplier      : 10
LLFC watchdog fixed-restore multiplier     : 0
+-----------------------------------------------------+

+-----------------------------------------------------+
Ethernet1/1 Interface LLFC watchdog: [Enabled]
+-----------------------------------------------------+
+-----------------------------------------------------+
| QOS GROUP 6 [Active] LLFC [YES] LLFC-COS [6]
+-----------------------------------------------------+
|                                |  Stats              |
+-----------------------------------------------------+
|                        Shutdown|                   1|
|                        Restored|                   1|
|             Total pkts drained|                 554|
|             Total pkts dropped|            56093783|
|   Total pkts drained + dropped|            56094337|
|          Aggregate pkts dropped|            56094337|
|     Total Ingress pkts dropped|                   0|
| Aggregate Ingress pkts dropped|                   0|
+-----------------------------------------------------+
```

Configuring Pause Buffer Thresholds and Queue Limit Using Ingress Queuing Policy

The pause buffer thresholds specified in the network-qos policy are shared by all the ports in the system. However, there are situations where a few ports may need different thresholds (such as long distance connections). An ingress queuing policy can be used for this purpose.

An ingress queuing policy also allows the configuration of the queue-limit to restrict the amount of shared buffer that can be used in addition to the reserved pause buffer by the no-drop class.

Each no-drop class is mapped internally to one of the port's priority-group in the ingress direction. The configured pause buffer thresholds and queue-limit are applied to the priority-group associated with the class.

Note Adding pause buffer size threshold configuration is optional for cable lengths that are less than 100 meters and it need not be configured.

For cable lengths that are greater than 100m, the pause buffer size threshold configuration is mandatory and it is required as part of the QoS policy configuration.

Note About queue limits for 100G enabled devices (such as the Cisco Nexus 9300 platform switch with the N9K-M4PC-CFP2 GEM):

- The maximum dynamic queue-limit alpha value supported by the device might be greater that 8. However 8 is the maximum alpha value supported. Configuring the alpha value to a value greater than 8 is overridden by the maximum alpha value of 8.

  No message is issued when the alpha value is overridden.

- The static queue-limit has a maximum of 20,000 cells. Any value specified greater than the maximum 20,000 cell limit is overridden by the 20,000 cell limit.

  No message is issued when the cell limit is overridden.

SUMMARY STEPS

1. **configure terminal**
2. **policy-map type queuing** *policy-map-name*
3. **class type queuing** *c-in-q1*
4. **pause buffer-size** *buffer-size* **pause threshold** *xoff-size* **resume threshold** *xon-size*
5. **no pause buffer-size** *buffer-size* **pause threshold** *xoff-size* **resume threshold** *xon-size*
6. **queue-limit** *queue size* [**dynamic** *dynamic threshold*]

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**configure terminal**	Enters global configuration mode.
Step 2	**policy-map type queuing** *policy-map-name*	Enters policy-map queuing class mode and identifies the policy map assigned to the type queuing policy map.
Step 3	**class type queuing** *c-in-q1*	Attaches the class map of type queuing and then enters policy-map class queuing mode. Class queuing names are listed in the System-Defined Type queuing Class Maps table. **Note**

	Command or Action	Purpose
		The qos-group associated with the class must be defined as a no-drop class in the network-qos policy applied in the system qos. **Note** Up to eight ingress queues are supported for the Cisco Nexus 9636C-R and 9636Q-R line cards and the Cisco Nexus 9508-FM-R fabric module (in a Cisco Nexus 9508 switch). The range is from c-in-8q-q-default to c-in-8q-q1 through 7.
Step 4	**pause buffer-size** *buffer-size* **pause threshold** *xoff-size* **resume threshold** *xon-size*	Specifies the buffer threshold settings for pause and resume.
Step 5	**no pause buffer-size** *buffer-size* **pause threshold** *xoff-size* **resume threshold** *xon-size*	Removes the buffer threshold settings for pause and resume.
Step 6	**queue-limit** *queue size* [**dynamic** *dynamic threshold*]	(Optional) Specifies either the static or dynamic shared limit available to the ingress priority-group. The static queue limit defines the fixed size to which the priority-group can grow. The dynamic queue limit allows the priority-group's threshold size to be decided depending on the number of free cells available, in terms of the alpha value. **Note** Cisco Nexus 9200 platform switches only support a class level dynamic threshold configuration with respect to the alpha value. This means that all ports in a class share the same alpha value. **Note** The queue limit for the Cisco Nexus 9636C-R and 9636Q-R line cards and the Cisco Nexus 9508-FM-R fabric module (in a Cisco Nexus 9508 switch) can be entered as a percent or in bytes/kbytes/mbytes/gbytes. For example, **queue-limit percent 1** or **queue-limit bytes 100**.

Verifying the Priority Flow Control Configuration

To display the PFC configuration, perform the following task:

Command	Purpose
show interface priority-flow-control [**module** *number*]	Displays the status of PFC on all interfaces or on specific modules.

Configuration Examples for Priority Flow Control

The following example shows how to configure PFC:

```
configure terminal
interface ethernet 5/5
priority-flow-control mode on
```

The following example shows how to enable PFC on a traffic class:

```
switch(config)# class-map type qos c1
switch(config-cmap-qos)# match cos 3
switch(config-cmap-qos)# exit
switch(config)# policy-map type qos p1
switch(config-pmap-qos)# class type qos c1
switch(config-pmap-c-qos)# set qos-group 3
switch(config-pmap-c-qos)# exit
switch(config-pmap-qos)# exit
switch(config)# class-map type network-qos match-any c1
switch(config-cmap-nqos)# match qos-group 3
switch(config-cmap-nqos)# exit
switch(config)# policy-map type network-qos p1
switch(config-pmap-nqos)# class type network-qos c-nq1
switch(config-pmap-nqos-c)# pause pfc-cos 3
switch(config-pmap-nqos-c)# exit
switch(config-pmap-nqos)# exit
switch(config)# system qos
switch(config-sys-qos)# service-policy type network-qos p1
```

The following example shows how to configuring the PFC mode and its policies which is a prerequisite for PFC watchdog:

```
Watchdog is enabled by default, with system default values of:

Watchdog interval = 100 ms

Shutdown multiplier = 1

Auto-restore multiplier = 10
```

The following example shows how to check PFC watchdog statistics:

```
switch# show queuing pfc-queue interface ethernet 1/23

slot  1
=======
+-----------------------------------------------------+
Global watch-dog interval [Enabled]
Forced Global watch-dog   [Enabled]
+-----------------------------------------------------+

+-----------------------------------------------------+
Global PFC watchdog configuration details

PFC watchdog poll interval                : 100 ms
PFC watchdog shutdown multiplier          : 1
PFC watchdog auto-restore multiplier      : 10
PFC watchdog fixed-restore multiplier     : 0
PFC watchdog internal-interface multiplier : 2
+-----------------------------------------------------+

+-------------------------------------------------------------+
```

```
| Port              PFC Watchdog  (VL bmap)  State (Shutdown) |
+-------------------------------------------------------------+
  Ethernet1/23      Enabled       ( 0x8 )   - - - - Y - - -  >>>>>>>>>>>>>>>> The Queue is
 marked as SHUT


switch# show queuing pfc-queue interface ethernet 1/23 detail

slot  1
=======
+---------------------------------------------------+
Global watch-dog interval [Enabled]
Forced Global watch-dog   [Enabled]
+---------------------------------------------------+

+---------------------------------------------------+
Global PFC watchdog configuration details

PFC watchdog poll interval                : 100 ms
PFC watchdog shutdown multiplier          : 1
PFC watchdog auto-restore multiplier      : 10
PFC watchdog fixed-restore multiplier     : 0
PFC watchdog internal-interface multiplier : 2
+---------------------------------------------------+

+---------------------------------------------------+
Ethernet1/23 Interface PFC watchdog: [Enabled]
Disable-action                      : No
PFC watch-dog interface-multiplier : 0
+---------------------------------------------------+
+---------------------------------------------------+
| QOS GROUP 3 [Shutdown] PFC [YES] PFC-COS [3]
+---------------------------------------------------+
|                              |  Stats             |
+---------------------------------------------------+
|                       Shutdown|                  1|
|                       Restored|                  0|
|             Total pkts drained|                  0|
|             Total pkts dropped|                  0|
|   Total pkts drained + dropped|                  0|
|         Aggregate pkts dropped|                  0|
|     Total Ingress pkts dropped|               1924|    >>>>>>>>> Account for Ingress
drops here
| Aggregate Ingress pkts dropped|               1924|
+---------------------------------------------------+
```

Configuring a No-Drop Policy

The following example shows how to configure a no-drop policy and attach the policy to a session policy:

```
Device# configure terminal
Device(config)# class-map type network-qos class1
Device(config-cmap-nq)# match qos-group 1
Device(config-cmap-nq)# policy-map type network-qos my_network_policy
Device(config-pmap-nq)# class type network-qos class1
Device(config-pmap-nq-c)# pause pfc-cos 2
Device(config-pmap-nq-c)# system qos
Device(config-sys-qos)# service-policy type network-qos my_network_policy
Device# show running ipqos
```

Classifying Traffic to a No-Drop Class

The following example shows how to create a QoS policy to map all the traffic to the no-drop class:

```
Device# configure terminal
Device(config)# class-map type qos class1
Device(config-cmap-qos)# match cos 2
Device(config-cmap-qos)# policy-map type qos my_qos_policy
Device(config-pmap-qos)# class type qos class1
Device(config-pmap-c-qos)# set qos-group 1
Device(config-pmap-c-qos)# interface e1/5
Device(config-sys-qos)# service-policy type qos input my_qos_policy
Device(config-sys-qos)#
```

Add the queuing policy that guarantees the bandwidth for qos-group 1 and apply that under system-qos as outlined in the following example:

```
policy-map type queuing my_queuing_policy
class type queuing c-out-q-default
bandwidth percent 1
class type queuing c-out-q3
bandwidth percent 0
class type queuing c-out-q2
bandwidth percent 0
class type queuing c-out-q1
bandwidth percent 99

system qos
  service-policy type queuing output my_queuing_policy
```

In the above example, c-out-q1 by default matches the traffic on qos-group 1. Therefore, the non-default class-map for queuing which matches qos-group 1 is not needed. For further information on configuring queuing, see Configuring Queuing.

For LLFC to be enabled, you need to configure the no-drop policy on network-qos. The buffering module needs to inform the MAC module to generate pause (either LLFC or PFC based on the interface level configuration). PFC negotiation to the adapter is by using DCBX. LLFC or PFC is controlled by the configuration on the interfaces. For example, the **flow-control send and receive on** enables LLFC on the interfaces and the **priority-flow-control mode on** enables PFC on the interfaces.

If DCBX is supported, auto mode negotiates the PFC with the adapter. This is the interface level configuration to enable LLFC or PFC but regardless of it, you have to configure network-qos level pause configuration for LLFC to work. Even if the traffic is classified to qos-group 1 but when it generates pause, it generates LLFC based on the interface level configuration.