Configuring Link Level Flow Control

Link Level Flow Control

Link-level flow control is a congestion management technique that pauses data transmission until the congestion in the system is resolved. When a receiving device becomes congested, it communicates with the transmitter by sending a PAUSE frame. When the transmitting device receives a Pause frame it stops the transmission of any further data frames 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.

Guidelines and Restrictions for Link Level Flow Control

- **show** commands with the **internal** keyword are not supported.
- Beginning with Cisco NX-OS Release 7.0(3)I7(8), syslog messages are generated at 5 minute intervals for accounting all the incoming global and link level pause frames.
- Changing or configuring LLFC on FEX HIF or FEX HIF PO interfaces is not supported.
- Link-level flow control (LLFC) is supported on Cisco Nexus 9500 platform switches with Network Forwarding Engine (NFE) (and Cisco Nexus 3164Q switch with NFE).
- Link-level flow control (LLFC) is supported on Cisco Nexus 9200 platform switches (7.0(3)I3(1) and later releases.
- Beginning with Cisco NX-OS Release 7.0(3)I5(1) and later releases, Link-level flow control (LLFC) is supported on the Cisco Nexus 9300 and 9300-EX platform switches.
- The 100G Cisco Nexus 9408PC-CFP2 line card does not support link-level flow control (LLFC).
- Ethernet interfaces do not auto-detect the link-level flow control capability. You must configure the capability explicitly.

- Enabling link level flow control requires a part of the buffer to be reserved. This reduces the available shared buffer space.
- Data Center Bridging Exchange Protocol (DCBX) is not supported.
- Configuration time quanta of the pause frames is not supported.
- On each Ethernet interface, the switch can enable either PFC or LLFC, but not both.

Note When both PFC and LLFC are enabled, LLFC is selected.

- Only pure CoS-based classification of traffic classes is supported.
- Setting of pause threshold values is restricted.
- Configuring Link Level Flow Control on the interfaces will flap the interfaces which results in a momentary traffic loss.
- When a no-drop QoS group is configured, you must ensure that packets received on ports that do not have flow control send-on configured are not classified to a no-drop QoS group.
- Only a no-drop QoS group is capable of generating link level pause frames.
- Weighted Random Early Detection (WRED) should not be enabled on a no-drop class because it can cause egress queue drops.
- It is recommended to use default buffer sizes for no-drop classes because if the buffer size is specified through CLI, it will allocate the same buffer size for all ports irrespective of the link speed, and MTU size.
- It is recommended to change the LLFC configuration when there is no traffic, otherwise packets already in the MMU of the system may not get the expected treatment.
- Beginning with NX-OS 7.0(3)I1(2), link level flow control and PFC are supported on Cisco Nexus 9300 platform switches and line cards that contain the ALE (Application Leaf Engine).
- 3232C does not support a combination of cut-through and LLFC enabled ports. Cut-through and LLFC are mutually exclusive and will work without the presence of the other feature. Post 9.3(8) release, on a cut-through enabled switch, if LLFC is enabled on a port, that port will operate in store and forward mode.
- First generation Cisco Nexus 9000 switches (NFE-based) can properly operate in cut-through switching mode only when there is no back pressure, i.e. LLFC pause frames should not be coming in. If RX LLFC is enabled alongside the default switching mode of cut-through, in some circumstances this can lead to output errors on the port configured with RX LLFC. To mitigate that, the user needs to either configure store and forward switching mode via the CLI "**switching-mode store-forward**" or disable RX LLFC on the given port. This limitation applies to the following PIDs:
  - N9K-C93120TX, N9K-C93128TX, N9K-C9332PQ, N9K-C9372PX, N9K-C9372PX-E, N9K-C9372TX, N9K-C9372TX-E, N9K-C9396PX, N9K-C9396TX, N9K-X9408PC-CFP2, N9K-X9432PQ, N9K-X9464PX, N9K-X9464TX, N9K-X9464TX2, N9K-X9536PQ, N9K-X9564PX, N9K-X9564TX, and N9K-X9636PQ.

Information About Link Level Flow Control

Link Level Flow Control on Interfaces

When link level flow control is configured the system changes the interface state to Down if the specified interface is in UP state and then applies the flow control configuration. After the configuration is successfully applied to the interface, the system restores the interface to the UP state.

Link Level Flow Control on Ports

During a port shutdown event, the flow-control settings on an interface are retained, however no traffic is received or transmitted on the link. During a port startup event the flow-control settings are reinstated on to the hardware.

Mismatched Link Level Flow Control Configurations

The transmit and receive directions can be configured separately, and each device on the network can have a different Link Level Flow Control (LLFC) configuration. The following table describes how devices with mis-matched configurations interact.

Switch A	Switch B	Description
LLFC configured to receive and transmit PAUSE frames.	LLFC configured to receive PAUSE frames.	Switch A can transmit 802.3x PAUSE frames and honor 802.3x PAUSE frames. Switch B can only receive 802.3x PAUSE frames.
LLFC configured to receive and transmit PAUSE frames.	LLFC configured to transmit PAUSE frames.	Switch A can transmit 802.3x PAUSE frames and honor 802.3x PAUSE frames. Switch B can transmit 802.3x PAUSE frames but will drop all received PAUSE frames.

How to Configure Link Level Flow Control

Configuring Link Level Flow Control Receive

SUMMARY STEPS

1. **configure terminal**
2. **interface ethernet 1/1**
3. **flowcontrol receive on**
4. **exit**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 2	**interface ethernet 1/1** **Example:** `Device(config)# interface ethernet 1/1`	Configures an interface type and enters interface configuration mode.
Step 3	**flowcontrol receive on** **Example:** `Device(config-if)# flowcontrol receive on`	Enables the interface to receive and process pause frames.
Step 4	**exit** **Example:** `Device(config-if)# exit`	Exits interface configuration mode.

Configuring Link Level Flow Control Transmit

To configure link-level flow control transmit on an interface, you enable flow control on the interface, configure a network-qos type QoS policy to enable a no-drop QoS group, and apply a qos type QoS policy to classify the traffic that requires no-drop behavior to the no-drop class.

You must ensure that bandwidth is allocated for the No-Drop QoS class using a queuing policy when you define a no-drop class. For more information, see the "Configuring Type Queuing Policies" section.

Note When a no-drop QoS Group is configured you must ensure that packets received on ports that do not have flow-control send-on configured, are not classified to a no-drop QoS group. This is required as any ingress port that does not have flow-control send-on configured, can not generate a link level pause frame and there is no way to request the transmitting device to stop the transmission. Therefore, if flow-control send-on is not configured on all the interfaces you should not use a system policy to classify the packets to the no-drop QoS group. Instead, you should apply an interface QoS policy to the interfaces that having flow-control send-on enabled.

SUMMARY STEPS

1. **configure terminal**
2. **interface ethernet 1/1**
3. **flowcontrol send on**
4. **exit**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 2	**interface ethernet 1/1** **Example:** `Device(config)# interface ethernet 1/1`	Configures an interface type and enters interface configuration mode.
Step 3	**flowcontrol send on** **Example:** `Device(config-if)# flowcontrol transmit on`	Enables the interface to send pause frames to remote devices.
Step 4	**exit** **Example:** `Device(config-if)# exit`	Exits interface configuration mode and returns to global configuration mode.

Configuring a Link Level Flow Control Watchdog Interval

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 To disable the LLFC watchdog, enter the following in the global configuration mode:

```
switch(config)# link-level-flow-control watch-dog interval off
```

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 watch-dog interval** *value*
3. **priority-flow-control auto-restore multiplier** *value*
4. (Optional) **show queuing llfc-queue [interface** *interface-list***] [module** *module***] [detail]**
5. (Optional) **clear queuing llfc-queue [interface** *interface-list***] [module** *module***]**

DETAILED STEPS

	Command or Action	Purpose

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

Step 2	**priority-flow-control watch-dog interval** *value* **Example:** `switch(config)# priority-flow-control watch-dog interval 200`	Specifies the watchdog interval value. The range is from 100 to 1000 milliseconds.
Step 3	**priority-flow-control auto-restore multiplier** *value* **Example:** `switch(config)# priority-flow-control auto-restore multiplier 50`	Configures a value for the auto-restore multiplier. The range is from 0 to 100. **Note** The auto-restore multiplier should never be configured to 0.

	Command or Action	Purpose

		Note 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`

Step 4	(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 5	(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.

Example

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

```
switch# show queuing llfc-queue interf e 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|
+-----------------------------------------------------+
```

Configuration Examples for Link Level Flow Control

Example: Configuring a No-Drop Policy

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.

Example: Configuring Link Level Flow Control Receive and Send

Configuring Link Level Flow Control Receive and Send

The following examples show how to configure Link Level Flow Control receive and send on the device.

- When only LLFC receive is enabled, no-drop class does not need to be configured on the system network-qos.

Note You must configure the no-drop class under system network-qos on the Cisco Nexus 9200 and 9300-EX/FX platforms for releases earlier than NX-OS 7.0(3)I7(3).

```
Device# configure terminal
Device(config)# interface ethernet 1/1
Device(config-if)# flowcontrol receive on
Device(config-if)# exit
```

- When both LLFC receive and send are enabled, no-drop class needs to be configured on the system network-qos. (Refer to the Configuring a No-Drop Policy example for information about configuring the no-drop class.)

```
Device# configure terminal
Device(config)# interface ethernet 1/1
```

```
Device(config-if)# flowcontrol receive on
Device(config-if)# flowcontrol send on
Device(config-if)# exit
```

- When only LLFC send is enabled, no-drop class needs to be configured on the system network-qos. (Refer to the Configuring a No-Drop Policy example for information about configuring the no-drop class.)

```
Device# configure terminal
Device(config)# interface ethernet 1/1
Device(config-if)# flowcontrol send on
Device(config-if)# exit
```