Congestion Management

This chapter provides information about devices that cause congestion in a Fibre Channel or Fibre Channel over Ethernet (FCoE) network and provides information about how to identify and avoid or isolate such devices. These devices can be both slow devices and devices that are attempting to over utilize the bandwidth of their links or interfaces.

Finding Feature Information

Your software release might not support all the features documented in this module. For the latest caveats and feature information, see the Bug Search Tool at https://tools.cisco.com/bugsearch/ and the release notes for your software release. To find information about the features documented in this module, and to see a list of the releases in which each feature is supported, see the New and Changed chapter or the Feature History table below.

Feature History for Congestion Management

Feature Name	Release	Feature Information
Congestion Isolation	8.5(1)	This feature is now handled by Fabric Performance Monitor (FPM). The following commands were introduced: • **feature fpm** • **fpm congested-device** {**exclude** | **static**} **list** • **member pwwn** *pwwn* **vsan** *id* [**credit-stall**] • **fpm congested-device recover pwwn** *pwwn* **vsan** *id* The following commands were deprecated: • **congestion-isolation** {**include** | **exclude**} **pwwn** *pwwn* **vsan** *vsan-id* • **feature congestion-isolation** • **show congestion-isolation** {**exclude-list** | **global-list** | **ifindex-list** | **include-list** | **pmon-list** | **remote–list** | **status**} • **congestion-isolation remove interface** *slot*/*port*

Feature Name	Release	Feature Information
Congestion Isolation Recovery	8.5(1)	The Congestion Isolation Recovery feature automatically recovers the flow which was moved to low-priority VL after it was detected as slow back to normal VL; thereby, recovering the flow. The following commands were introduced: • **feature fpm** • **fpm congested-device** {**exclude** | **static**} **list** • **member pwwn** *pwwn* **vsan** *id* [**credit-stall**] • **fpm congested-device recover pwwn** *pwwn* **vsan** *id* • **port-monitor cong-isolation-recover** {**recovery-interval** *seconds* | **isolate-duration** *hours* **num-occurrence** *number*} The **counter** port monitor command was modified to add the **cong-isolate-recover** port-guard action.

Feature Name	Release	Feature Information
Fabric Notifications	8.5(1)	Fabric Notifications are used to notify end devices of performance impacting conditions and behaviors that affect the normal flow of IO such as link integrity degradation and congestion. The following commands were introduced: • **feature fpm** • **counter txwait warning-signal-threshold** *count1* **alarm-signal-threshold** *count2* **portguard congestion-signals** • **fpm congested-device** {**exclude** | **static**} **list** • **member pwwn** *pwwn* **vsan** *id* [**credit-stall**] • **fpm congested-device recover pwwn** *pwwn* **vsan** *id* • **fpm fpin period** *seconds* • **fpm congestion-signal period** *seconds* • **show fpm** {**fpin** | **registration** {**congestion-signal** | **summary**} | **congested-device database** [**exclude** | **local** | **remote** | **static**]} **vsan** *id* • **port-monitor fpin** {**recovery-interval** *seconds* | **isolate-duration** *hours* **num-occurrence** *number*} The **counter** port monitor command was modified to add the **FPIN** port-guard action.

Feature Name	Release	Feature Information
Dynamic Ingress Rate Limiting (DIRL)	8.5(1)	DIRL is used to automatically limit the amount of traffic that is flowing through a switch port that is congested. The following commands were introduced: • **feature fpm** • **fpm dirl** {**exclude list** | **reduction** *percentage* **recovery** *percentage* • **member** {**fc4-feature target** | **interface fc** *slot/port*} • **fpm dirl recover interface fc** *slot/port* • **show fpm** {**dirl exclude** | **fpin vsan** *id* | **ingress-rate-limit** {**events** | **status**} **interface fc***slot/port*} • **port-monitor dirl recovery-interval** *seconds* The **counter** port monitor command was modified to add the **DIRL** port-guard action.

Feature Name	Release	Feature Information
Fibre Channel and Fibre Channel over Ethernet (FCoE)	8.4(1)	The following commands were modified: • The **show hardware internal rxwait-history** [**module** *number* | **port** *number*] command was changed to **show interface** [*interface-range*] **rxwait-history**. • The **show hardware internal txwait-history** [**module** *number* | **port** *number*] command was changed to **show interface** [*interface-range*] **txwait-history**. • The **show process creditmon txwait-history** [**module** *number* [**port** *number*]] command was changed to **show interface** [*interface-range*] **txwait-history**. The following command outputs were modified: • **show interface** *interface-range* **aggregate-counters** • **show interface** *interface-range* **counters** • **show interface** *interface-range* **counters detailed** • **show interface priority-flow-control** • **show interface vfc** *interface-range* **counters detailed**

Feature Name	Release	Feature Information
Fibre Channel over Ethernet (FCoE)	8.2(1)	New FCoE commands were introduced and some FCoE commands were modified to align with the commands used in Fibre Channel. The following commands were modified: • The congestion drop timeout command has changed from **system default interface congestion timeout** *milliseconds* **mode** {**core** | **edge**} to **system timeout fcoe congestion-drop** {*milliseconds* | **default**} **mode** {**core** | **edge**} • The pause drop timeout command has changed from **system default interface pause timeout** *milliseconds* **mode** {**core** | **edge**} to **system timeout fcoe pause-drop** {*milliseconds* | **default**} **mode** {**core** | **edge**} • The output for the **show interface vfc** *interface-range* **counters detailed** and **show interface priority-flow-control** commands were modified to add the receive and transmit pause frame information in the output. • The **show logging onboard** command was modified to add the **txwait**, **rxwait**, and **error-stats** keywords. The following commands were introduced: • **show hardware internal txwait-history** [**module** *number* | **port** *number*] • **show hardware internal rxwait-history** [**module** *number* | **port** *number*]

Feature Name	Release	Feature Information
Extended Receiver Ready	8.1(1)	This feature allows each Inter-Switch Link (ISL) between supporting switches to be split into four separate virtual links, with each virtual link assigned its own buffer-to-buffer credits. The following commands were introduced: • **show flow-control** {**er_rdy** | **r_rdy**} [**module** *number*] • **switchport vl-credit** {**default** | **vl0** *value* **vl1** *value* **vl2** *value* **vl3** *value*} • **system fc flow-control** {**default** | **er_rdy** | **r_rdy**}
Congestion Isolation	8.1(1)	This feature allows devices to be categorized as slow by either configuration command or by the port monitor. The following commands were introduced: • **congestion-isolation** {**include** | **exclude**} **pwwn** *pwwn* **vsan** *vsan-id* • **feature congestion-isolation** • **show congestion-isolation** {**exclude-list** | **global-list** | **ifindex-list** | **include-list** | **pmon-list** | **remote–list** | **status**} The *cong-isolate* portguard action was added to the following commands: • **counter credit-loss-reco** • **counter tx-credit-not-available** • **counter tx-slowport-oper-delay** • **counter tx-wait**

Feature Name	Release	Feature Information
Congestion Drop Timeout, No-Credit Frame Timeout, and Slow-Port Monitor Timeout Values for Fibre Channel	8.1(1)	The following features were modified: • The link connecting a core switch to a Cisco NPV switch should be treated as an ISL (core port) for the purpose of congestion-drop, no-credit-drop, slow-port monitor, and port monitor. To accomplish this, **logical-type** {**all** | **core** | **edge**} feature was introduced. • The Fibre Channel congestion drop timeout value's range was changed from 100-500 ms to 200-500 ms. The following commands were modified: **Note** From Cisco MDS NX-OS Release 8.1(1), E ports are treated as core and F ports are treated as edge in the **system timeout congestion-drop**, **system timeout no-credit-drop**, and **system timeout slowport-monitor** commands. • **system timeout congestion-drop** *milliseconds* **logical-type** {**core** | **edge**} • **system timeout no-credit-drop** *milliseconds* **logical-type edge** • **system timeout slowport-monitor** *milliseconds* **logical-type** {**core** | **edge**} • **switchport logical-type** {**auto** | **core** | **edge**}

Information About SAN Congestion

Information About SAN Congestion Caused by Slow-Drain Devices

Most SAN edge devices use Class 2 or Class 3 Fibre Channel services that have link-level flow control. This flow control feature allows a receiving port to back-pressure the upstream-sending port whenever the receiving port reaches its capacity to accept frames. When an edge device does not accept frames from the fabric for an extended time, it creates a congestion condition in the fabric that is known as slow drain. If the upstream source of a slow edge device is an ISL, it results in credit starvation or slow drain in that ISL. This credit starvation then affects any unrelated flows that use the same shared ISL. This type of congestion can occur in both Fibre Channel and FCoE although the flow control mechanisms are different in each of them. Regardless of the protocol of the device causing the congestion, the congestion can propagate back to the source of the frames via both Fibre Channel and FCoE links.

Fibre Channel uses buffer-to-buffer credits (BB_credits). This is a flow-control mechanism to ensure that each side of the Fibre Channel link is able to control the rate of incoming frames. BB_credits are set on a per-hop basis. Each side of a Fibre Channel connection informs the other side of the number of buffers that are available for it to receive frames. The sender can only send frames if the receiver has buffers. For each frame received, the receiver transmits an R_RDY (also known as BB_credit) to the sender of that frame. If there is some processing delay in the receiver, it can withhold the BB_credits from the sender, thereby limiting the rate at which it is receiving frames. If the receiver withholds the BB_credits for a significant amount, it causes congestion on that link. It may also cause congestion in the SAN as well. This BB_credit mechanism works independently in each direction of the traffic flow.

Frames and BB_credits are not sent reliably. If a frame is received that is so corrupt that it cannot be recognized, the receiver of that frame does not return a BB_credit. Or, if a frame is received intact and the BB_credit is returned but it is corrupted in transmission on the link, the receiver of that BB_credit does not recognize it as a BB_credit. In both cases, a transmit credit is lost. Credit Loss Recovery (LR or LRR) results when all the transmit credits are lost over time. The BB_SCN feature is used to recover such lost credits before completely running out of credits and causing congestion. Counts of frames and credits that are returned are periodically exchanged and if there is any discrepancy in the count then credits can be recovered. BB_SCN is available on all ISLs and is extended to F ports from Cisco MDS NX-OS Release 8.2(1). For F ports, the attached device must indicate support for BB_SCN in the FLOGI sent.

In FCoE, the flow control mechanism is called Priority Flow Control (PFC). PFC consists of a receiver sending class-based pause frames to a sender when it wants the sender to cease sending any frames of that class. PFC pause frames contain a value that is called a quanta. The quanta determines how long a class of traffic is paused. There are two types of PFC pause frames—nonzero quanta and zero quanta. A PFC pause frame with a nonzero quanta signals the receiver to stop sending frames immediately for a specified amount of time. A PFC pause frame with a zero quanta signals the receiver that it can resume sending frames immediately. As the receiver experiences some processing delay or its buffers reach a defined threshold, it can transmit a PFC pause frame with a nonzero quanta. After the buffers are sufficiently available, the receiver can transmit another PFC pause frame containing a zero quanta which in turn signals the sender to resume traffic. This PFC pause mechanism works in each direction of the traffic flow independently of the other.

Devices that do not accept frames at the rate that is generated by the sender can be both Fibre Channel and FCoE. The underlying flow control mechanism is different between the Fibre Channel and FCoE. But, Fibre Channel and FCoE can equally cause congestion in the SAN. These devices are referred to as slow-drain devices.

Slow-drain devices can be detected, and actions can be taken to mitigate the resulting congestion.

These actions include:

- Drop all or old frames that are queued to the slow drain interface that exceed the configured thresholds.
- Isolate the slow device to a separate logical virtual link on an ISL.
- Reset credits on the affected ports.
- Flap the affected ports.
- Error disable the affected ports.

These Congestion Detection, Congestion Avoidance, and Congestion Isolation features are used to detect slow-drain devices and take appropriate actions on them.

The slow drain condition can be classified in the following four levels:

- Level 3—Indicates severe congestion. Ports are without credits for a continuous amount of time and Credit Loss Recovery is initiated. For an F port, the duration when ports are without credits for a continuous amount of time is 1 second and for an E port it is 1.5 seconds. Credit Loss Recovery involves sending a Fibre Channel Link Credit Reset (LR) primitive to restore the BB_credits on the link in both directions. If the receiver responds with a Link Credit Reset Response (LRR), the credits are restored and the link resumes normal operation.

  If the congestion is severe, LRR may not be returned and the link fails with the *LR failed due to timeout* error. Credit Loss Recovery can be initiated from either side of the link. If MDS is the receiver of the LR (because the adjacent device initiated the Credit Loss Recovery), the only way MDS can return an LRR is when the input buffers of an interface are empty. If the interface still has frames that it had received but was unable to forward to the destination interface, the link fails with the *LR failed nonempty receive queue* error. If LR or LRR sequence is successful, the link returns to normal operation. Even if the link returns to its normal operation, the 1-second or 1.5-second time at zero Tx credit causes severe backwards congestion in the SAN. This backward congestion can work its way back all the way to the source of the frames. Severs or initiators typically see that a large amount of IO errors recorded due to many timeout drops that occur.

  When the link first initializes an LR and LRR, sequence occurs normally and does not indicate a level 3 slow drain condition.

  Although, severe congestion can occur in both Fibre Channel and FCoE the Link Credit Reset (LR or LRR) actions only apply to Fibre Channel.

- Level 2—Indicates moderate congestion that is causing frames to drop because the congestion drop timeout threshold has reached. Each frame that is received on an interface is timestamped. If the frame cannot be transmitted to the appropriate egress port within a congestion drop threshold of a switch, the frame is dropped to prevent excessive internal congestion in the switch. This is typically due to the adjacent device on the egress interface withholding credits (Fibre Channel) or sending PFC pauses. Each dropped frame is part of a SCSI (or other protocol) exchange and causes that exchange to fail. Servers or initiators record IO errors and terminate communication when SCSI exchanges fail. When the path between the initiator and target is over shared infrastructure, for example ISLs, other devices that are utilizing the shared infrastructure also sees timeout drops and large delays in their IO completion times. The congestion drop threshold is 500 ms by default and can be set to as low as 200 ms. The congestion drop threshold can be separately set for Fibre Channel and FCoE ports.
- Level 1 and Level 1.5—Indicates that delay occurs when frames cannot be transmitted immediately out of an egress port due to the port being without Tx buffer-to-buffer credits in Fibre Channel or in an Rx Pause state for FCoE. The amount of delay is measured by the TxWait counter and can be calculated as a percentage of time. For example, if a port is unable to transmit for 200 ms (not necessarily continuous)

in a 1-second interval then the TxWait congestion percentage for that 1-second interval is 20% for the specified interval. Level 1.5 indicates a more severe level of delay and is reserved for TxWait greater than or equal to 30%. Level 1 indicates instances when TxWait is less than 30%.

Almost always, higher levels of slow drain include the lower levels. For example, Level 3 slow drain includes level 2, level 1.5, and level 1 because the lack of ability to transmit causes delay and the delay causes timeout dropped frames. Longer delay causes Credit Loss Recovery to be initiated.

The following terms are used in the document:

- Buffer-to-Buffer (BB) credits (Fibre Channel only): BB_credits are a link flow control mechanism that is used in Fibre Channel. A Fibre Channel frame can only be transmitted if the *remaining Tx credit count* is greater than zero. When the frame is transmitted, the *remaining Tx credit count* is decremented by one. When the receiver of the frame processes the frame, it returns a credit that is called Receiver Ready (R_RDY). When an R_RDY is returned, the frame sender increments the *remaining Tx credit count* by one. If the *remaining Tx credit count* hits zero, no further frames can be transmitted until an R_RDY is received.
- R_RDY (Fibre Channel): A Buffer-to-Buffer credit. For more information, see Buffer-to-Buffer (BB) credits (Fibre Channel only).
- ER_RDY (Extended R_RDY): A Virtual Link based Buffer-to-Buffer credit. From Cisco MDS NX-OS 8.1(1), MDS introduced the Congestion-Isolation feature. This feature allows slow-drain devices to be isolated to a slow traffic virtual link (VL2) on an ISL (E port). The ISL must be in the Extended Receiver Ready (ER_RDY) mode for this feature to function. When an ISL is in ER_RDY mode, the link is logically partitioned into four separate virtual links. ER_RDY contains the VL number indicating which VL the BB credit is used for.
- PFC Pause (FCoE only): Priority Flow Control is a class-based flow control mechanism where class-based pause frames are sent to stop the flow of data in one direction for a specific class of service. PFC pause frames contain class bitmap and a value that is called a quanta. The class bitmap specifies which classes, or priorities, the pause frame applies to and the quanta determines how long a class of traffic is paused. There are two types of PFC pause frames: pause frames containing a nonzero quanta and pause frames containing a zero quanta. A PFC pause frame with a nonzero quanta signals the receiver to stop sending frames for the class immediately for a specified amount of time. A PFC pause frame with a zero quanta signals the receiver that it can resume sending frames for the class immediately. A PFC pause frame with a zero quanta can be called an *unpause* or *resume*.
- Transitions to zero (Fibre Channel only): When the *remaining Tx credit count* hits zero, the Tx transition to zero counter is incremented on the Tx side. On the Rx side (the side withholding the BB_credits), the Rx transition to zero counter is incremented. It is important to understand that the amount of time actually at zero *remaining Tx credits* is not represented by this counter. It could be for a short time that does not affect performance or it could be for a longer time that affects performance. Because of this, transitions to zero is not a good measure of congestion.
- TxWait (Fibre Channel and FCoE): TxWait is a measure of time when a port cannot transmit when it has frames queued in it. A port cannot transmit if it is at zero *remaining Tx credit count* (Fibre Channel) or if it has received a PFC pause frame. Each time TxWait increments, the port (or class) is unable to transmit for 2.5 microseconds. TxWait value can be converted to seconds by multiplying it by 2.5 and then dividing by 1,000,000.
- RxWait (FCoE only): RxWait is a measure of time where a port cannot receive frames. A port cannot receive frames if it has transmitted a PFC pause frame (FCoE). Each time RxWait increments, the port (or class) is unable to receive for 2.5 microseconds. RxWait can be converted to seconds by multiplying it by 2.5 and then dividing by 1,000,000.

- Tx Credit not Available (Fibre Channel only): Tx Credit not Available is a software counter that increments by one when the port is at zero *remaining Tx credits* continuously for 100 ms.

  Timeout-drop (Fibre Channel and FCoE): A frame is dropped as a timeout drop when a received frame is unable to be transmitted out of the egress interface in the configured congestion-drop threshold time. This condition is typically due to congestion at the egress interface that is caused by a lack of Tx BB_credits (Fibre Channel) or in an Rx Pause state (FCoE). The default timeout drop value is 500 ms for both Fibre Channel and FCoE but can be configured to a value as low as 200 ms. Also, the frames that are dropped when the no-credit-drop (Fibre Channel) or pause-drop threshold is reached are also marked as timeout drops.

- Credit Loss Recovery (Fibre Channel only): Credit Loss Recovery occurs when a port is at zero *remaining Tx credits* continuously for 1 second (F or NP port) or 1.5 seconds (E port). When this condition occurs a Link Credit Reset (LR) Fibre Channel primitive is sent to reinitialize the credits (both directions) on the link. If a Link Credit Reset Response (LRR) is returned, all credits are restored and the link resumes to normal operation. If an LRR is not returned, the link fails and must completely reinitialize. For information about reasons for credit-loss-recovery, see Reasons for Credit-Loss-Recovery, on page 14.

- Link Credit Reset (LR) (Fibre Channel only): LR is a Fibre Channel primitive that is used at link initialization, as well as, to reinitialize BB_credits in both directions on an active link when credits are lost.

- Link Credit Reset Response (LRR) (Fibre Channel only): LRR is a Fibre Channel primitive that is a positive response to an LR.

Reasons for Credit-Loss-Recovery

Credit-loss-recovery can occur for the following distinct reasons:

- Frame or R_RDY corruption or loss: As discussed in the section on the BB_SCN feature, frames, and BB_credits (R_RDYs) can be corrupted and lost on the link. If the BB_SCN feature is negotiated between the end-point devices, then corruption or loss of frames can be detected and recovered as long as the number of lost or corrupted frames or BB_credits is less than the total number of credits over the detection window. If the interface completely runs out of transmit BB_credits either because BB_SCN was not negotiated on or the number of lost or corrupted frames or BB_credits was equal to the number of transmit BB_credits, then credit-loss-recovery is initiated. Frame and BB_credits that are lost or corrupted are due to some physical problem in the link. Check and replace SFPs, fiber cables, and patch panels first. Rarely the switch port or HBA could be a fault.

- Severe congestion: This is due to severe congestion in the end device. The reasons for this vary by end device type along with the OS and application so they cannot be described here.

To determine the reason for the credit-loss-recovery:

- Check for invalid CRCs, invalid transmission words, input errors, and any other signs of corrupted data on the interface with the credit-loss recovery. If there are any of these signs, then it is likely that the problem is due to corrupted or lost frames BB_credits. However, if there are no indications of invalid CRCs, invalid transmission words, or input errors, then the problem still could be due to corrupted or lost frames, or BB_credits. This is because a frame or a BB_credit could be corrupted and/or lost after it is transmitted by the MDS. If this is the case, then MDS would not know that has occurred and would not increment any counters indicating a problem. To check for these types of errors use the **show interface fc** *x/y* **counters detailed** command.

- Check for invalid CRCs, invalid transmission words, input errors, and any other signs of corrupted data on the adjacent device's interface or HBA. You can check for errors at the device itself (for example, at the host or target). Also, you can use the **show rdp fcid** *fcid_id* **vsan** *vsan_id* command to query the adjacent device's HBA for errors. Using this command it can be easily determined if there are invalid CRCs, invalid transmission words, or input errors on data being received from MDS. Note that not all HBAs support the **show rdp fcid** *fcid_id* **vsan** *vsan_id* command.
- Check for non-zero BB_SCN counts on the MDS interface. Non-zero BB_SCN counts indicate that BB_SCN is detecting a loss of some BB_credits or frames and is successfully recovering them. This is a good sign of some BB_credit and/or frames being lost or corrupted. To check for BB_SCN recovery occurrences, use the **show interface fc** *x/y* **counters detailed** command and look for the *BB_SCs credit resend actions* and *BB_SCr Tx credit increment actions* lines in the command output.
- Check if credit-loss-recovery is occurring for the same device on both A and B fabrics at the same or similar times. If that is the case, then it is unlikely that there is a similar physical problem with physical components on both links. The problem is most likely severe congestion being reflected back to the MDS switch port. To check for credit loss recovery occurrences use the **show interface fc** *x/y* **counters detailed** command and look for the *Tx Credit loss* line in the command output.
- Check for common or repetitive times of the day or week when this happens. Frames and BB_credits are not usually corrupted and/or lost only at certain times of the day or days of the week. This is a sign of severe congestion and not of BB_credit or frame loss or corruption.
- If the port experiencing Credit-Loss-Recovery is part of a port-channel (either F port-channel or E port-channel/ISL) and there are more than one port in the same port-channel experiencing Credit-Loss-Recovery, then most likely the problem is due to congestion. This is because the MDS load balances across all the members of a port-channel. Consequently, flows for one or more slow devices will be transmitted across all members in the port-channel and will affect all members. If only a single member of the port-channel is experiencing Credit-Loss-Recovery, then most likely the problem is due to physical components in the link.

Information About SAN Congestion Caused by Over Utilization

Small Computer Systems Interface (SCSI) initiator devices request data via various SCSI *read* commands. These SCSI *read* commands contain a data length field, which is the amount of data requested in the specific *read* request. Likewise, SCSI targets request data via the SCSI Xfr_rdy command and the amount of data requested is contained in the burst size. The rate of these *read* or Xfr_rdy requests coupled with the amount of data requested can result in more data flowing to the specific end device than its link can support at a given time. This is compounded by speed mismatches, hosts zoned to multiple targets, and targets zoned to multiple hosts.

The switch infrastructure (SAN) can buffer some of this excess data, but if the rate of requests is continuous then the queues of a switch can fill and Fibre Channel or FCoE back pressure can result. This back pressure is done by withholding BB_credits on Fibre Channel and by sending PFC pauses on FCoE. The resulting effects to the SAN can look identical to slow drain, but the root cause is much different since the end device is not actually withholding buffer-to-buffer credits (or sending PFC Pauses). The main mechanism for detecting congestion caused by over utilization is by monitoring the Tx data rate of the end device ports. Port monitor can be used to detect congestion caused by over utilization.

Information About Congestion Management

Information About Congestion Detection

The following features are used to detect congestion on all slow-drain levels on Cisco MDS switches:

- **All Slow-Drain Levels**

  Display of credits agreed to along with the remaining credits on a port (Fibre Channel only)—The credits that are agreed to in both directions in FLOGI (F ports) and Exchange Link Parameters (ELP) for ISLs are displayed via the **show interface** command. Also, the instantaneous value of the remaining credits is also displayed in the output of the **show interface** command. The credits agreed to is static and unchanging information, at least when the link is up. However, the remaining credit values are constantly changing because each time a frame is transmitted, the Tx remaining count is decremented, and each time a credit is received, the Tx remaining count is incremented. When the remaining credits approach or reach zero, it indicates congestion on that port.

  The following example displays the transmitted and received credits information on an F port:

```
switch# show interface fc9/16
fc9/16 is up
Hardware is Fibre Channel, SFP is short wave laser w/o OFC (SN)
Port mode is F, FCID is 0x0c0100
Transmit B2B Credit is 16
Receive B2B Credit is 32
.
.
.
32 receive B2B credit remaining
16 transmit B2B credit remaining
```

  The following example displays the transmitted and received credits information on an E port that is in R_RDY mode:

```
switch# show interface fc1/5
fc1/5 is trunking (Not all VSANs UP on the trunk)
Hardware is Fibre Channel, SFP is short wave laser w/o OFC (SN)
Transmit B2B Credit is 64
Receive B2B Credit is 500
.
.
.
500 receive B2B credit remaining
64 transmit B2B credit remaining
```

  The following example displays the transmitted and received credits information on an E port that is in ER_RDY mode:

```
switch# show interface fc9/1 | i i fc | credit
fc9/1 is trunking
Transmit B2B Credit for vl0:15 vl1:15 vl2:40 vl3:430
Receive B2B Credit for vl0:15 vl1:15 vl2:40 vl3:430
.
.
```

```
.
Transmit B2B credit remaining for virtual link 0-3: 15,15,40,428
Receive B2B credit remaining for virtual link 0-3: 15,15,40,430
```

- **Level 3**

  Level 3 slow-drain condition is characterized by Fibre Channel BB_credits being unavailable continuously for 1 to 1.5 seconds. This condition causes the Credit Loss Recovery mechanism to be invoked to reinitialize both Tx and Rx credits on a link.

  For links in ER_RDY mode, Credit Loss Recovery link reset will still be initiated if Tx BB_credits are unavailable on virtual links 0, 1, and 3 for 1.5 seconds, and this duration cannot be changed or configured. For VL2, the slow VL, it will be initiated if Tx BB_credits are unavailable for 15 seconds, and this duration cannot be changed or configured.

  Note In the ER_RDY mode, Credit Loss Recovery will reset the credits for all VLs.

  Level 3 slow-drain condition is almost always accompanied by level 2 and level 1 or level 1.5 slow-drain condition.

  Credit Loss Recovery that is initiated by either side of a link can be seen in the following ways:

  The following example displays the count of Credit Loss Recovery being initiated by a switch on an interface:

  Note This command output is applicable for Cisco MDS NX-OS Release 8.4(2) and later releases. The command output varies if you are using Cisco MDS NX-OS Release 8.4(1a) or earlier releases.

```
switch# show interface fc1/4 counters detailed
fc1/4
  Rx 5 min rate bit/sec:                                          0
  Tx 5 min rate bit/sec:                                          0
  Rx 5 min rate bytes/sec:                                        0
  Tx 5 min rate bytes/sec:                                        0
  Rx 5 min rate frames/sec:                                       0
  Tx 5 min rate frames/sec:                                       0

Total Stats:
  Rx total frames:                                                9
  Tx total frames:                                               21
  Rx total bytes:                                               716
  Tx total bytes:                                              1436
  Rx total multicast:                                             0
  Tx total multicast:                                             0
  Rx total broadcast:                                             0
  Tx total broadcast:                                             0
  Rx total unicast:                                               9
  Tx total unicast:                                              21
  Rx total discards:                                              0
  Tx total discards:                                              0
  Rx total errors:                                                0
  Tx total errors:                                                0
```

```
 Rx class-2 frames:                                                0
  Tx class-2 frames:                                               0
  Rx class-2 bytes:                                                0
  Tx class-2 bytes:                                                0
  Rx class-2 frames discards:                                      0
  Rx class-2 port reject frames:                                   0
  Rx class-3 frames:                                               9
  Tx class-3 frames:                                              21
  Rx class-3 bytes:                                              716
  Tx class-3 bytes:                                             1436
  Rx class-3 frames discards:                                      0
  Rx class-f frames:                                               0
  Tx class-f frames:                                               0
  Rx class-f bytes:                                                0
  Tx class-f bytes:                                                0
  Rx class-f frames discards:                                      0

Link Stats:
  Rx Link failures:                                                0
  Rx Sync losses:                                                  0
  Rx Signal losses:                                                0
  Rx Primitive sequence protocol errors:                           0
  Rx Invalid transmission words:                                   0
  Rx Invalid CRCs:                                                 0
  Rx Delimiter errors:                                             0
  Rx fragmented frames:                                            0
  Rx frames with EOF aborts:                                       0
  Rx unknown class frames:                                         0
  Rx Runt frames:                                                  0
  Rx Jabber frames:                                                0
  Rx too long:                                                     0
  Rx too short:                                                    0
  Rx FEC corrected blocks:                                         0
  Rx FEC uncorrected blocks:                                       0
  Rx Link Reset(LR) while link is active:                          0
  Tx Link Reset(LR) while link is active:                          0
  Rx Link Reset Responses(LRR):                                    0
  Tx Link Reset Responses(LRR):                                    1
  Rx Offline Sequences(OLS):                                       0
  Tx Offline Sequences(OLS):                                       1
  Rx Non-Operational Sequences(NOS):                               0
  Tx Non-Operational Sequences(NOS):                               0

Congestion Stats:
  Tx Timeout discards:                                             0
  Tx Credit loss:                                                  0
  BB_SCs credit resend actions:                                    0
  BB_SCr Tx credit increment actions:                              0
  TxWait 2.5us due to lack of transmit credits:                    0
  Percentage TxWait not available for last 1s/1m/1h/72h:  0%/0%/0%/0%
  Rx B2B credit remaining:                                        32
  Tx B2B credit remaining:                                        16
  Tx Low Priority B2B credit remaining:                           16
  Rx B2B credit transitions to zero:                               1
  Tx B2B credit transitions to zero:                               2

Other Stats:
  Zone drops:                                                      0
  FIB drops for ports  1-16:                                       0
  XBAR errors for ports  1-16:                                     0
  Other drop count:                                                0

Last clearing of "show interface" counters :                   never
```

The following example displays instances of Credit Loss Recovery being initiated by a switch in OBFL error-stats:

Note The other slow drain indications displayed that accompany Credit Loss Recovery.

```
switch# show logging onboard error-stats

----------------------------
 Show Clock
----------------------------
2018-08-22 12:59:20

---------------------------------
 Module: 1 error-stats
---------------------------------


--------------------------------------------------------------------------
 ERROR STATISTICS INFORMATION FOR DEVICE DEVICE: FCMAC
--------------------------------------------------------------------------
Interface|                                 |         |    Time Stamp
  Range  |   Error Stat Counter Name       |  Count  |MM/DD/YY HH:MM:SS
         |                                 |         |
--------------------------------------------------------------------------------
fc1/1    |F16_TMM_TOLB_TIMEOUT_DROP_CNT    |14713116 |08/22/18 10:25:15
fc1/1    |FCP_SW_CNTR_TX_WT_AVG_B2B_ZERO   |1781669  |08/22/18 10:25:15
fc1/1    |FCP_SW_CNTR_CREDIT_LOSS          |18       |08/22/18 10:25:15
fc1/1    |F16_TMM_TOLB_TIMEOUT_DROP_CNT    |13338566 |08/22/18 10:24:55
fc1/1    |FCP_SW_CNTR_TX_WT_AVG_B2B_ZERO   |1781544  |08/22/18 10:24:55
fc1/1    |FCP_SW_CNTR_CREDIT_LOSS          |10       |08/22/18 10:24:55
fc1/1    |F16_TMM_TOLB_TIMEOUT_DROP_CNT    |11929676 |08/22/18 10:24:35
fc1/1    |FCP_SW_CNTR_TX_WT_AVG_B2B_ZERO   |1781418  |08/22/18 10:24:35
fc1/1    |F16_TMM_TOLB_TIMEOUT_DROP_CNT    |11881213 |08/22/18 10:24:15
fc1/1    |FCP_SW_CNTR_TX_WT_AVG_B2B_ZERO   |1781307  |08/22/18 10:24:15
```

The following example displays instances of Credit Loss Recovery failing due to the adjacent device not returning an LR. This causes a link failure:

```
switch# show logging log | i i timeout
…
2018 Aug 17 12:54:59 MDS9710 %PORT-5-IF_DOWN_LINK_FAILURE: %$VSAN 1%$ Interface fc1/2
is down (Link failure Link reset failed due to timeout) port-channel228
2018 Aug 17 13:42:01 MDS9710 %PORT-5-IF_DOWN_LINK_FAILURE: %$VSAN 1%$ Interface fc1/2
is down (Link failure Link reset failed due to timeout)
```

The following example displays LRR received on a port:

```
switch# show interface fc1/1 counters detailed
fc1/1
      27651428465 frames, 59174056872960 bytes received
```

```
…
      0 link reset received while link is active              <<<<< Credit Loss Recovery
 initiated from the adjacent device
…
       18 link reset responses received                          <<<<< LRRs received
       0 link reset responses transmitted                        <<<<< LRRs transmitted
```

The following example displays a received LR failing due to severe ingress congestion on that interface:

```
switch# show log last 20
…
2018 Aug 22 10:21:44 MDS9710 %PORT-5-IF_DOWN_LINK_FAILURE: %$VSAN 237%$ Interface fc1/13
 is down (Link failure Link Reset failed nonempty recv queue)
```

- **Level 2**

  Level 2 slow-drain condition indicates that the links are so congested that the received frames that are destined for the congested links cannot be transmitted within the congestion-drop threshold. When this condition occurs, these frames are discarded or dropped as timeout-drops. These dropped frames cause SCSI exchanges to fail at the end hosts. Timeout discards would normally be accompanied by level 1 or level 1.5 congestion.

  Timeout-drops are displayed in the following ways:

  - Count of timeout-drops on an interface

    ```
    switch# show interface fc1/1 counters | i fc | discard
    fc1/13
          0 discards, 0 errors, 0 CRC/FCS
          14713116 discards, 0 errors    <<<<< total drops/discards
        14713116 timeout discards, 18 credit loss      <<<<< timeout drops/discards
    ```

    Discards—Specifies the total output discards or dropped frames. Discards are also known as frame drops.

    Timeout discards—Specifies the total output frames discarded due to congestion-drop or no-credit-drop threshold being reached.

  - Instances of timeout-drops in OBFL error-stats

    ```
    switch# show logging onboard module 1 error-stats

    ----------------------------
     Show Clock
    ----------------------------
    2018-08-22 17:15:32

    ---------------------------------
     Module: 1 error-stats
    ---------------------------------


    -------------------------------------------------------------------------------
     ERROR STATISTICS INFORMATION FOR DEVICE DEVICE: FCMAC
    -------------------------------------------------------------------------------
    ```

```
Interface  |                                |         |    Time Stamp
  Range    |   Error Stat Counter Name      |  Count  |MM/DD/YY HH:MM:SS
           |                                |         |
-------------------------------------------------------------------------
fc1/1      |F16_TMM_TOLB_TIMEOUT_DROP_CNT   |14713116 |08/22/18 10:25:15
fc1/1      |FCP_SW_CNTR_TX_WT_AVG_B2B_ZERO  |1781669  |08/22/18 10:25:15
fc1/1      |FCP_SW_CNTR_CREDIT_LOSS         |18       |08/22/18 10:25:15
fc1/1      |F16_TMM_TOLB_TIMEOUT_DROP_CNT   |13338566 |08/22/18 10:24:55
fc1/1      |FCP_SW_CNTR_TX_WT_AVG_B2B_ZERO  |1781544  |08/22/18 10:24:55
fc1/1      |FCP_SW_CNTR_CREDIT_LOSS         |10       |08/22/18 10:24:55
```

- Instances of timeout-drops in OBFL flow-control timeout-drops

```
switch# show logging onboard flow-control timeout-drops

---------------------------------
 Module: 1 flow-control timeout-drops
---------------------------------

-----------------------------
 Show Clock
-----------------------------
2018-08-22 17:16:57


-------------------------------------------------------------------------
 ERROR STATISTICS INFORMATION FOR DEVICE DEVICE: FCMAC
-------------------------------------------------------------------------
 Interface |                                |         |    Time Stamp
   Range   |    Error Stat Counter Name     |  Count  |MM/DD/YY HH:MM:SS
           |                                |         |
-------------------------------------------------------------------------
fc1/1      |F16_TMM_TOLB_TIMEOUT_DROP_CNT   |14713116 |08/22/18 10:25:15
fc1/1      |F16_TMM_TOLB_TIMEOUT_DROP_CNT   |13338566 |08/22/18 10:24:55
fc1/1      |F16_TMM_TOLB_TIMEOUT_DROP_CNT   |11929676 |08/22/18 10:24:35
fc1/1      |F16_TMM_TOLB_TIMEOUT_DROP_CNT   |11881213 |08/22/18 10:24:15
fc1/1      |F16_TMM_TOLB_TIMEOUT_DROP_CNT   |11771790 |08/22/18 10:23:55
```

- **Level 1 or Level 1.5**

  Level 1 or level 1.5 slow-drain condition indicates that the interface is without transmit BB_credits at times. The interface can track the exact amount of time an interface is at zero transmit credits, in Fibre Channel and the exact amount of time FCoE class is paused in both directions. When an FCoE interface receives a PFC pause, it cannot transmit in a similar fashion to a Fibre Channel interface when the Fibre Channel interface is at zero transmit credits. This duration of time when an interface cannot transmit credits is called TxWait and is counted in 2.5 micro-second intervals. An FCoE interface transmitting a PFC pause (to prevent the other side from transmitting) is like a Fibre Channel interface not returning BB_credits. This duration of time when an interface cannot receive credits is called RxWait and is also counted in 2.5-micro intervals. Currently, RxWait is only measured for FCoE. In Fibre Channel, this duration of time an interface cannot receive credits is only measured by a software process. It is measured only when the interface is at zero Rx credits remaining for a continuous 100 ms amount of time.

  - Display of credit transitions to zero on a port (Fibre Channel only)—Whenever a port hits zero transmit or receive BB_credits, the transmit (Tx) or receive (Rx) BB_credits transitions to zero is incremented. When the transmit BB_credit transitions to zero is incremented, it indicates that the adjacent device has withheld BB_credits or BB_credits are lost. When the receive BB_credit

transitions to zero is incremented, it indicates that the switchport is withholding BB_credit from an adjacent device. These interface counters can increment occasionally under normal conditions. These interface counters do not give any indication of the amount of time the interface was at zero credits. Therefore, these counters are not a preferred indication of congestion on a port. See the TxWait and RxWait counters for a better indication of Tx and Rx congestion on a port.

```
switch# show interface fc1/13 counters
fc1/13
    5 minutes input rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
    5 minutes output rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
    0 frames input, 0 bytes
      0 class-2 frames, 0 bytes
      0 class-3 frames, 0 bytes
      0 class-f frames, 0 bytes
      0 discards, 0 errors, 0 CRC/FCS
      0 unknown class, 0 too long, 0 too short
    0 frames output, 0 bytes
      0 class-2 frames, 0 bytes
      0 class-3 frames, 0 bytes
      0 class-f frames, 0 bytes
      0 discards, 0 errors
    0 timeout discards, 0 credit loss
    0 input OLS, 0 LRR, 0 NOS, 0 loop inits
    0 output OLS, 0 LRR, 0 NOS, 0 loop inits
    0 link failures, 0 sync losses, 0 signal losses
     0 Transmit B2B credit transitions to zero
     0 Receive B2B credit transitions to zero
      0 2.5us TxWait due to lack of transmit credits
      Percentage Tx credits not available for last 1s/1m/1h/72h: 0%/0%/0%/0%
      32 receive B2B credit remaining
      31 transmit B2B credit remaining
      31 low priority transmit B2B credit remaining
    Last clearing of "show interface" counters: 2d00h
```

Transmit B2B credit transitions to zero – Count of times the interface was at zero Tx B2B credits remaining and unable to transmit. This could be because the adjacent device withheld B2B credits from this interface, credits (or frames which should have generated credits) were lost, or because there were insufficient credits for the speed, average frame size, and distance of the link.
Receive B2B credit transitions to zero – Count of times the interface was at zero Rx B2B credits remaining. This is due to this interface withholding B2B credits.

- Display of the total amount of TxWait and RxWait on an interface. Each increment represents 2.5 microseconds of time an interface was at zero Tx or Rx credits. This can be displayed using the **show interface counters** and **show interface counters detailed** commands.

```
switch# show interface fc1/1 counters
fc1/1
    5 minutes input rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
    5 minutes output rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
    27651428465 frames input, 59174056872960 bytes
      0 class-2 frames, 0 bytes
      0 class-3 frames, 59174056872960 bytes
      0 class-f frames, 0 bytes
      0 discards, 0 errors, 0 CRC/FCS
      0 unknown class, 0 too long, 0 too short
    907817 frames output, 1942720200 bytes
      0 class-2 frames, 0 bytes
```

```
      907817 class-3 frames, 1942720200 bytes
      0 class-f frames, 0 bytes
      14713116 discards, 0 errors
    14713116 timeout discards, 18 credit loss
    0 input OLS, 18 LRR, 0 NOS, 0 loop inits
    0 output OLS, 0 LRR, 0 NOS, 0 loop inits
    0 link failures, 0 sync losses, 0 signal losses
     903218 Transmit B2B credit transitions to zero
     743093 Receive B2B credit transitions to zero
      108369199104 2.5us TxWait due to lack of transmit credits
      Percentage Tx credits not available for last 1s/1m/1h/72h: 0%/0%/0%/0%
      32 receive B2B credit remaining
      128 transmit B2B credit remaining
    Last clearing of "show interface" counters: 6w 4d
```

2.5us TxWait due to lack of transmit credits – Count of TxWait ticks in 2.5us since the interface counters have been cleared last. In this example, 108369199104 * 2.5 / 1000000 = 270922.99776 seconds of time the interface has not been able to transmit in the past 6 weeks and 4 days.

Percentage Tx credits not available for last 1s/1m/1h/72h: 0%/0%/0%/0% - Percentage of TxWait as calculated in the last 1 second, 1 minute, 1 hour, and 72 hour intervals.

- Display of TxWait, RxWait, and percentage Tx and Rx credits not available for the last 1 second, 1 minute, 1 hour, and 72 hour—This can be displayed using the **show interface counters detailed** command.

```
switch# show interface fc1/1 counters
fc1/1
    5 minutes input rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
    5 minutes output rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
    27651428465 frames input, 59174056872960 bytes
      0 class-2 frames, 0 bytes
      0 class-3 frames, 59174056872960 bytes
      0 class-f frames, 0 bytes
      0 discards, 0 errors, 0 CRC/FCS
      0 unknown class, 0 too long, 0 too short
    907817 frames output, 1942720200 bytes
      0 class-2 frames, 0 bytes
      907817 class-3 frames, 1942720200 bytes
      0 class-f frames, 0 bytes
      14713116 discards, 0 errors
    14713116 timeout discards, 18 credit loss
    0 input OLS, 18 LRR, 0 NOS, 0 loop inits
    0 output OLS, 0 LRR, 0 NOS, 0 loop inits
    0 link failures, 0 sync losses, 0 signal losses
     903218 Transmit B2B credit transitions to zero
     743093 Receive B2B credit transitions to zero
      108369199104 2.5us TxWait due to lack of transmit credits
      Percentage Tx credits not available for last 1s/1m/1h/72h: 0%/0%/0%/0%
      32 receive B2B credit remaining
      128 transmit B2B credit remaining
    Last clearing of "show interface" counters: 6w 4d
```

2.5us TxWait due to lack of transmit credits – Count of TxWait ticks in 2.5us since the interface counters have been cleared last. In this example, 108369199104 * 2.5 / 1000000 = 270922.99776 seconds of time the interface has not been able to transmit in the past 6 weeks and 4 days.

Percentage Tx credits not available for last 1s/1m/1h/72h: 0%/0%/0%/0% - Percentage of TxWait as calculated in the last 1 second, 1 minute, 1 hour, and 72 hour intervals.

```
switch# show interface vfc1/3 counters
```

```
vfc1/3
    3166 fcoe in packets
    460532 fcoe in octets
    3166 fcoe out packets
    1005564 fcoe out octets
    0 2.5 us TxWait due to pause frames for VL3
    0 2.5 us RxWait due to pause frames for VL3
    0 Tx frames with pause opcode for VL3
    0 Rx frames with pause opcode for VL3
    Percentage pause in TxWait per VL3 for last 1s/1m/1h/72h: 0%/0%/0%/0%
    Percentage pause in RxWait per VL3 for last 1s/1m/1h/72h: 0%/0%/0%/0%
```

- Display of histograms showing Tx credit unavailability TxWait (Fibre Channel) and PFC pause (TxWait and RxWait) for the last 60 seconds, 60 minutes, and 72 hours—You can display this information using the **show process creditmon txwait-history** (Fibre Channel) and **show system** {**txwait-history** | **rxwait-history**} (FCoE) commands.

Note From Cisco MDS NX-OS Release 8.4(1), the **show process creditmon txwait-history** and **show hardware internal** {**txwait-history** | **rxwait-history**} command has changed to the **show interface** [*interface-range*] {**txwait-history** | **rxwait-history**} command.

TxWait (or credit unavailability) increments because of lack of transmit BB_credits (Fibre Channel) or because of receiving PFC pause frames (FCoE).

RxWait (currently FCoE only) increments when the interface transmits PFC Pause frames.

There are three graphs for each command and each graph has the most recent second, minute, or hour unit on the X axis:

1. Seconds scale—Indicates the past 60 seconds, where each column represents a second of time. Above the histogram are the amounts of time, in milliseconds, that the ports were unable to transmit and is represented in a vertical format. In the first graph shown, 8 seconds before the command being run, there were 857 ms of TxWait (credit unavailability) in the 1-second interval. The most current second is displayed on the left.
2. Minutes scale—Indicates the past 60 minutes, where each column represents a minute of time. Above the histogram are the amounts of time, in seconds, that the ports were unable to transmit and is represented in a vertical format. In the second graph shown, a minute before the command being run, there was a 22.7 second of TxWait (credit unavailability) in the 1-minute interval. The most current minute is displayed on the left.
3. Hours scale—Indicates the past 72 hours, where each column represents an hour of time. Above the histogram are the amounts of time, in seconds, that the ports were unable to transmit and is represented in a vertical format. In the third graph shown, 24 hours before the command being run, there was a 342 seconds of TxWait (credit unavailability) in the 1-minute interval. And, 52 hours prior, there was a 220 seconds of TxWait in that hour. The most current hour is displayed on the left.

```
switch#  show interface fc1/1 txwait-history | no-more

TxWait history for port fc1/1:
```

```
==============================
            8999994              299999999999999999997
            5888883              188879888888888889999998
     000000076362570000000000000066468354635464357888708700000000
1000         #####                 ########################
 900        ######                 ########################
 800        ######                 #########################
 700        ######                 #########################
 600        ######                 #########################
 500        ######                 #########################
 400        #######                #########################
 300        #######                #########################
 200        #######               ##########################
 100        #######               ##########################
    0....5....1....1....2....2....3....3....4....4....5....5....6
              0    5    0    5    0    5    0    5    0    5    0

          Tx Credit Not Available per second (last 60 seconds)
                 # = TxWait (ms)

   245 45         1
   239121000000000800000000000000000000000000000000000000000000000000
   ...............................................................
   772006000000000800000000000000000000000000000000000000000000000000
60   #
54   #  #
48   #  #
42  ## ##
36  ## ##
30  ## ##
24 ### ##
18 ### ##         #
12 ### ##         #
 6 ### ##         #
  0....5....1....1....2....2....3....3....4....4....5....5....6
            0    5    0    5    0    5    0    5    0    5    0

          Tx Credit Not Available per minute (last 60 minutes)
                 # = TxWait (secs)


                            3                           2
                            4                           2
     000000000000000000000002000000000000000000000000000000000000000000000000000
3600
3240
2880
2520
2160
1800
1440
1080
 720
 360                        #                           #
    0....5....1....1....2....2....3....3....4....4....5....5....6....6....7.7
              0    5    0    5    0    5    0    5    0    5    0    5    0 2

          Tx Credit Not Available per hour (last 72 hours)
                 # = TxWait (secs)
```

The **show interface** [*interface-range*] **rxwait-history** (FCoE) command is similar to the TxWait history commands except that it shows the amount of time the port was receiving PFC pause; therefore, preventing the adjacent device from transmitting.

```
switch# show interface e1/47 rxwait-history

RxWait history for port Eth1/47:
  ==============================

       7887777787777778777777788778778777787778777777777777877777778
       9009999909999990999999900990990999909909999999999199999990
  1000
   900
   800
   700
   600
   500
   400
   300
   200
   100 ############################################################
      0....5....1....1....2....2....3....3....4....4....5....5....6
                0    5    0    5    0    5    0    5    0    5    0

            RxWait per second (last 60 seconds)
                   # = RxWait (ms)

              1
     444444444557000000000000000000000000000000000000000000000000
     ............................................................
     777777777587000000000000000000000000000000000000000000000000
  60
  54
  48
  42
  36
  30
  24
  18            #
  12            #
   6 ############
    0....5....1....1....2....2....3....3....4....4....5....5....6
              0    5    0    5    0    5    0    5    0    5    0

            RxWait per minute (last 60 minutes)
                   # = RxWait (secs)


       2                 1                           1
       7                2 5                          9
       000000000000000006060002000000000000000000000009000000000000000000000001
  3600
  3240
  2880
  2520
  2160
  1800
  1440
  1080
   720
   360 #                                             #
      0....5....1....1....2....2....3....3....4....4....5....5....6....6....7.7
                0    5    0    5    0    5    0    5    0    5    0    5    0 2
```

```
  RxWait per hour (last 72 hours)
              # = RxWait (secs)
```

- Display of delta TxWait and RxWait values in 20-second intervals where the delta TxWait is greater than or equal to 100 ms—You can use the **show logging onboard txwait** (Fibre Channel and FCoE) or **show logging onboard rxwait** (FCoE) commands to display the delta TxWait and RxWait values.

  TxWait and RxWait are logged to the persistent log (logging onboard or OBFL) whenever a port accumulates 100 ms or more of TxWait or RxWait in a 20-second interval. If a port accumulates less than 100 ms of TxWait or RxWait, nothing is logged for that 20-second interval.

  The following information is displayed in the logging onboard TxWait and RxWait:

  - Delta TxWait or RxWait ticks—Each tick represents 2.5 microseconds. Because the minimum value logged is the equivalent of 100 ms, the minimum value that is displayed in the output is 40,000.
  - Delta TxWait or RxWait in seconds—TxWait value that is multiplied by 2.5 and then divided by 1,000,000 results in the TxWait value, in seconds. The TxWait value is displayed as an integer in the output. Therefore, TxWait value less than 1 second is displayed as 0.
  - Congestion Percentage (%)—TxWait or RxWait value that is divided by 20 results in TxWait or RxWait, in seconds. This value gives a quick way of seeing how the congestion was in the 20-second interval.
  - Timestamp—Indicates the date and time at the end of a 20-second interval when the delta TxWait was determined.

```
switch# show logging onboard txwait module 2

--------------------------------
  Module: 2 txwait count
--------------------------------

----------------------------
  Show Clock
----------------------------
2019-04-08 13:56:52
Notes:
      - Sampling period is 20 seconds
      - Only txwait delta >= 100 ms are logged

-------------------------------------------------------------------------
| Interface | Delta TxWait Time     |Congestion| Timestamp               |
|           | 2.5us ticks | seconds |          |                         |
-------------------------------------------------------------------------
|Eth2/2(VL3)|   882562    |    2    |   11%    | Tue Sep 11 08:52:34 2018|
|Eth2/1(VL3)|  4647274    |   11    |   58%    | Tue Sep 11 08:52:14 2018|
|Eth2/2(VL3)|  7529479    |   18    |   94%    | Tue Sep 11 08:52:14 2018|
|Eth2/1(VL3)|  7829159    |   19    |   97%    | Tue Sep 11 08:51:54 2018|
|Eth2/2(VL3)|  7923544    |   19    |   99%    | Tue Sep 11 08:51:54 2018|
|Eth2/1(VL3)|  5299754    |   13    |   66%    | Tue Sep 11 08:50:34 2018|
|Eth2/2(VL3)|   362484    |    0    |    4%    | Tue Sep 11 08:50:34 2018|
|Eth2/1(VL3)|  7924925    |   19    |   99%    | Tue Sep 11 08:50:14 2018|
|Eth2/2(VL3)|  2566450    |    6    |   32%    | Tue Sep 11 08:50:14 2018|
|Eth2/1(VL3)|  7935558    |   19    |   99%    | Tue Sep 11 08:49:54 2018|
|Eth2/2(VL3)|  6762560    |   16    |   84%    | Tue Sep 11 08:49:54 2018|
|Eth2/1(VL3)|  7908259    |   19    |   98%    | Tue Sep 11 08:49:34 2018|
```

```
|Eth2/2(VL3)|  5264976    |   13    |   65%    | Tue Sep 11 08:49:34 2018|
|Eth2/1(VL3)|  7925639    |   19    |   99%    | Tue Sep 11 08:49:14 2018|

switch# show logging onboard rxwait module 2

---------------------------------
  Module: 2 rxwait count
---------------------------------

----------------------------
  Show Clock
----------------------------
2019-04-08 13:58:03
Notes:
      - Sampling period is 20 seconds
      - Only rxwait delta >= 100 ms are logged

-------------------------------------------------------------------------------
| Interface | Delta RxWait Time     | Congestion | Timestamp                |
|           | 2.5us ticks | seconds |            |                          |
-------------------------------------------------------------------------------
|Eth2/1(VL7)|  6568902    |   16    |      82%   | Thu Aug  2 14:29:54 2018|
|Eth2/1(VL6)|  6568927    |   16    |      82%   | Thu Aug  2 14:29:54 2018|
|Eth2/1(VL5)|  6568951    |   16    |      82%   | Thu Aug  2 14:29:54 2018|
|Eth2/1(VL4)|  6568975    |   16    |      82%   | Thu Aug  2 14:29:54 2018|
|Eth2/1(VL3)|  6569000    |   16    |      82%   | Thu Aug  2 14:29:54 2018|
|Eth2/1(VL2)|  6569024    |   16    |      82%   | Thu Aug  2 14:29:54 2018|
|Eth2/1(VL1)|  6569050    |   16    |      82%   | Thu Aug  2 14:29:54 2018|
|Eth2/1(VL0)|  6569075    |   16    |      82%   | Thu Aug  2 14:29:54 2018|
|Eth2/2(VL7)|  7523430    |   18    |      94%   | Thu Aug  2 14:29:54 2018|
|Eth2/2(VL6)|  7523455    |   18    |      94%   | Thu Aug  2 14:29:54 2018|
|Eth2/2(VL5)|  7523479    |   18    |      94%   | Thu Aug  2 14:29:54 2018|
|Eth2/2(VL4)|  7523504    |   18    |      94%   | Thu Aug  2 14:29:54 2018|
|Eth2/2(VL3)|  7523528    |   18    |      94%   | Thu Aug  2 14:29:54 2018|
|Eth2/2(VL2)|  7523552    |   18    |      94%   | Thu Aug  2 14:29:54 2018|
|Eth2/2(VL1)|  7523578    |   18    |      94%   | Thu Aug  2 14:29:54 2018|
```

- Display of average Tx credit not available in 100-ms intervals—Cisco MDS switches have a software process that runs every 100 ms to check for ports that are in continuous state of 0 Tx credits remaining. The ports that are in the continuous state of 0 Tx credits are displayed in the output of the **show system internal snmp credit-not-available** [**module** *module*] and **show logging onboard error-stats** commands. These commands display 100 ms, 200 ms, or more of continuous state of 0 Tx credits remaining.

  The **show system internal snmp credit-not-available** [**module** *module*] command shows the Tx Credit Not Available alerts from port monitor. The alerts are in 100-ms intervals, as a percentage, of the configured port-monitor polling interval. If the Tx Credit Not Available (tx-credit-not-available) port-monitor counter is not configured in the active policy, no events are displayed.

  The *Duration of time not available* column is the percentage of polling interval where Tx credits were at zero and unavailable. In the command output, for the Event Time, Tue Aug 18 19:41:34 2018, the *Duration of time not available* is 10% and indicates 100 ms (10% of the polling interval of 1 second is 100 ms). At Tue Aug 18 19:52:52 2018, the port-monitor policy was changed so that the tx-credit-not-available counter's polling interval was 10 seconds and the rising-threshold was 20%. The *Duration of time not available* column shows 49% and indicates that almost 5 of the 10 seconds of Tx credits were at zero.

```
switch# show system internal snmp credit-not-available
Module: 1      Number of events logged: 20
```

```
--------------------------------------------------------------------------------
Port     Threshold Rising Interval(s) Event Time        Type      Duration of time

         /Falling                                                 not available
--------------------------------------------------------------------------------
fc1/94   10/0(%)            1          Tue Aug 18       Rising    10%
                                       19:41:34 2018
fc1/94   10/0(%)            1          Tue Aug 18       Falling   0%
                                       19:42:14 2018
fc1/94   10/0(%)            1          Tue Aug 18       Rising    10%
                                       19:42:15 2018
fc1/94   10/0(%)            1          Tue Aug 18       Falling   0%
                                       19:42:55 2018
fc1/94   10/0(%)            1          Tue Aug 18       Rising    10%
                                       19:42:56 2018
fc1/94   10/0(%)            1          Tue Aug 18       Falling   0%
                                       19:44:34 2018
fc1/94   10/0(%)            1          Tue Aug 18       Rising    10%
                                       19:44:35 2018
fc1/94   10/0(%)            1          Tue Aug 18       Falling   0%
                                       19:48:50 2018
fc1/94   10/0(%)            1          Tue Aug 18       Rising    20%
                                       19:48:51 2018
fc1/94   10/0(%)            1          Tue Aug 18       Falling   0%
                                       19:49:31 2018
fc1/94   10/0(%)            1          Tue Aug 18       Rising    20%
                                       19:49:32 2018
fc1/94   10/0(%)            1          Tue Aug 18       Falling   0%
                                       19:51:42 2018
fc1/94   10/0(%)            1          Tue Aug 18       Rising    10%
                                       19:51:43 2018
fc1/94   10/0(%)            1          Tue Aug 18       Falling   0%
                                       19:52:51 2018
fc1/94   10/0(%)            1          Tue Aug 18       Rising    10%
                                       19:52:52 2018
fc1/94   10/0(%)            1          Tue Aug 18       Falling   0%
                                       19:53:14 2018
fc1/94   10/0(%)            1          Tue Aug 18       Rising    20%
                                       19:53:15 2018
fc1/94   10/0(%)            1          Tue Aug 18       Falling   0%
                                       19:58:36 2018
fc1/94   20/0(%)            10         Tue Aug 18       Rising    49%
                                       20:20:02 2018
fc1/94   20/0(%)            10         Tue Aug 18       Falling   0%
                                       20:21:45 2018
```

- Display of average Tx credit not available in logging onboard error-stats—The **show logging onboard error-stats** command displays the average Tx credit not available in 100-ms intervals as indicated by the FCP_SW_CNTR_TX_WT_AVG_B2B_ZERO counter. This counter increments by 1 for every 100 ms that an interface is in a continuous state of 0 Tx credits remaining. The increments are recorded in the command output every 20 seconds. Information about other counters is also included in the command output.

```
switch# show logging onboard error-stats

----------------------------
    Module:  1
----------------------------

----------------------------
```

```
 Show Clock
----------------------------
2018-08-28 12:28:15

---------------------------------
 Module: 1 error-stats
---------------------------------


--------------------------------------------------------------------------------
 ERROR STATISTICS INFORMATION FOR DEVICE: FCMAC
--------------------------------------------------------------------------------
    Interface|                                    |         |    Time Stamp
      Range  |     Error Stat Counter Name        |  Count  |MM/DD/YY HH:MM:SS
             |                                    |         |
--------------------------------------------------------------------------------
fc7/2        |IP_FCMAC_CNT_STATS_ERRORS_RX_BAD_   |35806503 |03/17/19 11:32:44
              WORDS_FROM_DECODER
fc7/2        |FCP_SW_CNTR_TX_WT_AVG_B2B_ZERO      |2        |03/17/19 11:32:44
fc7/1        |FCP_SW_CNTR_TX_WT_AVG_B2B_ZERO      |1        |03/17/19 11:32:44
fc7/15       |FCP_SW_CNTR_RX_WT_AVG_B2B_ZERO      |1        |03/15/19 22:10:25
fc7/15       |FCP_SW_CNTR_RX_WT_AVG_B2B_ZERO      |16       |03/15/19 18:32:44
fc7/15       |F16_TMM_TOLB_TIMEOUT_DROP_CNT       |443      |03/15/19 15:39:42
fc7/15       |FCP_SW_CNTR_RX_WT_AVG_B2B_ZERO      |12       |03/15/19 13:37:59
fc7/15       |FCP_SW_CNTR_RX_WT_AVG_B2B_ZERO      |8        |03/15/19 13:29:59
fc7/15       |FCP_SW_CNTR_RX_WT_AVG_B2B_ZERO      |4        |03/15/19 13:26:19
fc7/15       |FCP_SW_CNTR_RX_WT_AVG_B2B_ZERO      |3        |01/01/17 13:12:14
fc7/15       |FCP_SW_CNTR_RX_WT_AVG_B2B_ZERO      |25       |03/14/19 21:13:34
fc7/15       |FCP_SW_CNTR_RX_WT_AVG_B2B_ZERO      |21       |03/14/19 21:06:34
fc7/15       |FCP_SW_CNTR_RX_WT_AVG_B2B_ZERO      |17       |03/14/19 20:58:34
```

- Display of Tx and Rx transitions to zero (Fibre Channel only)—When an interface reaches zero remaining credits in either direction, the *transitions to zero* counter is incremented. This incrementation of the counter indicates that a port is running out of credits, but does not indicate the duration that the port was at zero credits. The port could have been at zero credits momentarily or for a longer time. TxWait provides a better view of the impact of credits running out because it gives the actual duration that the port was at 0 Tx credits remaining. Transitions to zero are shown in the **show interface counters** and **show interface counters detailed** commands.

  The following example displays the *transition to zero* counts of the transmit and receive credits:

```
switch# show interface fc1/1 counters | i fc | transitions
fc1/1
0 Transmit B2B credit transitions to zero
0 Receive B2B credit transitions to zero
```

- Priority-flow-control pauses (FCoE only)—Provides a count of PFC pause frames that are sent and received on an interface. PFC pause is a count and includes both PFC pauses with a nonzero quanta (actual pause frames) and PFC pauses with a zero quanta (unpause or resume frames). This count does not give any indication of the amount of time the port is paused. The port could have been paused momentarily or for a longer time. TxWait and RxWait give a better view of the impact of these pause frames because they provide the actual amount of time the port was paused in each direction. PFC pauses can be displayed via the **show interface** and **show interface priority-flow-control** commands.

  The following example displays the *pause* counts in the transmit and receive direction:

```
switch# show interface eth3/1
Ethernet3/1 is up
admin state is up, Dedicated Interface
Belongs to Epo540
...snip
RX
555195 unicast packets 105457 multicast packets 0 broadcast packets
...snip
230870335 Rx pause
TX
326283313 unicast packets 105258 multicast packets 0 broadcast packets
...snip
0 Tx pause
```

The following example displays the RxPause, TxPause counts and the corresponding RxWait, and TxWait for Ethernet ports used for FCoE:

```
switch# show interface priority-flow-control
RxPause: No. of pause frames received
TxPause: No. of pause frames transmitted
TxWait: Time in 2.5uSec a link is not transmitting data[received pause]
RxWait: Time in 2.5uSec a link is not receiving data[transmitted pause]
================================================================================
Interface Admin Oper (VL bmap) VL RxPause    TxPause RxWait-    TxWait-
                                                     2.5us(sec) 2.5us(sec)
================================================================================

Epo540    Auto   NA    (8)     3   456200000  0       0(0)       152866694355(382166)
Eth2/1    Auto   On    (8)     3   4481929    0       0(0)       5930346153(14825)
...snip
Eth2/48   Auto   Off
Eth3/1    Auto   On    (8)     3   0          0       0(0)       0(0)
...snip
Eth3/6    Auto   Off
Eth3/7    Auto   On    (8)     3   0          0       0(0)       0(0)
```

- Slowport monitor (Fibre Channel only)—A threshold value of slowport monitor is specified to detect ports that are at zero transmit credits for a specified continuous duration. When a port is at zero Tx credits continuously for the specified threshold value, the switch records an entry in the slowport-monitor log and in logging onboard. This entry is shown in the **show process creditmon slowport-monitor-events** and **show logging onboard slowport-monitor-events** commands. The entry that is shown in the outputs of these commands is identical, except that the slowport-monitor log only holds the last ten events per port. However, the logging onboard holds the events in chronological order and can hold more events when compared to the slowport-monitor log.

  Events are recorded at a maximum frequency of 100 ms. When the count goes up, operational delay is displayed in the command output. Operational delay indicates the length of time when the port was at zero Tx credits. If the count goes up by more than one from the previous entry, then the operational delay is the average operational delay from multiple events in the 100 ms interval.

  In the following example, at 02/02/18 18:12:37.308 the slowport detection count was 276 and the previous value was 273. This example indicates that there were three intervals of time in the previous 100 ms where the port was at zero Tx credits for 1 ms or more. The average time the port was at zero credits is shown in the *oper delay* column (4 ms). Oper delay of 4 ms indicates that there was a total of 12 ms of time when the port was at zero Tx credits in the previous 100 ms. The 12-ms duration was in three separate intervals.

Port monitor can also generate a slowport-monitor alert by using port monitor. By default, slowport-monitor alert is set to off. Slowport-monitor must be configured to get the port-monitor slowport-monitor alerts.

The **show process creditmon slowport-monitor-events** [**module** *number*] [**port** *number*] command shows the last ten events per port.

```
switch# show process creditmon slowport-monitor-events

        Module: 01      Slowport Detected: NO


        Module: 09      Slowport Detected: YES
================================================================================
 Interface = fc9/2
--------------------------------------------------------------------
| admin  | slowport  | oper  |            Timestamp            |
| delay  | detection | delay |                                 |
| (ms)   | count     | (ms)  |                                 |
--------------------------------------------------------------------
|   1    |       289 |    2  | 1. 02/02/18 21:33:20.853        |
|   1    |       279 |   10  | 2. 02/02/18 21:33:20.749        |
|   1    |       279 |   19  | 3. 02/02/18 21:33:20.645        |
|   1    |       276 |    4  | 4. 02/02/18 18:12:37.308        |
|   1    |       273 |    3  | 5. 02/02/18 17:07:44.395        |
|   1    |       258 |    2  | 6. 02/02/18 13:33:08.451        |
|   1    |       254 |    1  | 7. 02/02/18 12:49:01.899        |
|   1    |       253 |   14  | 8. 02/02/18 12:49:01.794        |
|   1    |       242 |    1  | 9. 02/02/18 10:07:33.594        |
|   1    |       242 |    3  |10. 02/02/18 10:07:32.865        |
```

The **show logging onboard slowport-monitor-events** command shows all slowport monitor events per module.

```
switch# show logging onboard slowport-monitor-events module 9

---------------------------------
 Module: 9 slowport-monitor-events
---------------------------------

----------------------------
 Show Clock
----------------------------
2018-02-03 12:27:45

---------------------------------
 Module: 9 slowport-monitor-events
---------------------------------

-------------------------------------------------------------------------------
| admin  | slowport  | oper  |            Timestamp         | Interface
| delay  | detection | delay |                              |
| (ms)   | count     | (ms)  |                              |
-------------------------------------------------------------------------------
|   1    |       289 |    2  | 02/02/18 21:33:20.853        | fc9/2
|   1    |       279 |   10  | 02/02/18 21:33:20.749        | fc9/2
|   1    |       277 |   19  | 02/02/18 21:33:20.645        | fc9/2
|   1    |       276 |    4  | 02/02/18 18:12:37.308        | fc9/2
…snip
```

- RxWait (FCoE only)—It is a measure of time that a port is in a transmit PFC pause state that is preventing the adjacent device from transmitting to the port. RxWait increments by 1 every 2.5 microseconds that a port is unable to receive.

  RxWait is shown in the following ways:

  - Cumulative count—Indicates the time the interface counters were last cleared, using the **show interface counters**, **show interface counters detailed** , and **show interface priority-flow-control** commands.
  - Count in percent—Indicates when the credits are unable to transmit for the last 1 second, 1 minute, 1 hour, and 72 hours, using the **show interface counters** and **show interface counters detailed** commands.
  - Graphical representation of the count for the last 60 seconds, 60 minutes, and 72 hours—In FCoE, the count is displayed using the **show interface** [*interface-range*] **rxwait-history** command.
  - On-Board Failure Log (OBFL)—An entry in the OBFL when a port accumulates 100 ms or more RxWait in a 20-second interval. This entry is displayed using the **show logging onboard rxwait** command.

In the following example, the **show interface counters** command output displays the data "1104349910 2.5 us TxWait due to pause frames (VL3)." This data is cumulative from the time when the counters were last cleared or from the time when the module first came up. In this example, TxWait is incremented 1104349910 times. This data converted to seconds is (1104349910 * 2.5) / 1000000 = 2760.874 seconds. The VFC port channel was unable to transmit for 2760.874 seconds.

In the following example, the **show interface counters** command output shows the data "205484298144 2.5 us RxWait due to PFC Pause frames (VL3)." This data is cumulative from the time the counters were last cleared or from the time when the module first came up. In the example, RxWait is incremented 205484298144 times. This data converted to seconds is (205484298144 * 2.5) / 1000000 = 513710.745 seconds. The VFC port channel was unable to receive for 513710.745 seconds.

The following example also shows the percentage of time that the VFC was paused in each direction over the last 1 second, 1 minute, 1 hour, and 72 hours. For TxWait, this is the percentage of time that the VFC received PFC pauses. For RxWait, this is the percentage of time that the VFC was sending pause frames preventing the other side from transmitting. In this example, in the last one minute the VFC was prevented from transmitting (TxWait) 33% of the time (20 seconds).

Note When the interface displayed is a VFC port channel or a VFC bound to an Ethernet port-channel, all values are cumulative for all members in the Ethernet port-channel.

```
switch# show interface vfc-po540 counters

vfc-po540
    1571394073 fcoe in packets
    3322884900540 fcoe in octets
    79445277 fcoe out packets
    69006091691 fcoe out octets
    1104349910 2.5 us TxWait due to pause frames (VL3)
    205484298144 2.5 us RxWait due to pause frames (VL3)
    0 Tx frames with pause opcode (VL3)
    3302000 Rx frames with pause opcode (VL3)
```

```
Percentage pause in TxWait per VL3  for last 1s/1m/1h/72h: 0%/33%/0%/0%
Percentage pause in RxWait per VL3  for last  1s/1m/1h/72h: 0%/0%/0%/30%
```

The **show logging onboard error-stats** command has several different counters that pertain to SAN congestion. Most of these counters are module or switch dependent. For information about tx-credit-not-available or rx-credit-not-available, the following counters are used:

- FCP_SW_CNTR_TX_WT_AVG_B2B_ZERO[5, 50i,48S,96S]
- F32_MAC_KLM_CNTR_TX_WT_AVG_B2B_ZERO[6]
- Count of the number of times that an interface was at zero Tx BB_credits for 100 ms. This count typically indicates congestion at the device that is attached to an interface.
- FCP_SW_CNTR_RX_WT_AVG_B2B_ZERO[5,50i,48S,96S]
- F32_MAC_KLM_CNTR_RX_WT_AVG_B2B_ZERO[6]
- Count of the number of times an interface was at zero Rx BB_credits for 100 ms. This count typically indicates that the switch is withholding R_RDY primitive to a device attached to an interface of a switch due to congestion in the path to devices with which it is communicating.

  Also, port monitor can generate tx-credit-not-available alerts (Fibre Channel only). See the Port Monitor section.

- Overutilization—Configuring port monitor with the Tx datarate and Rx datarate counters allow the MDS to issue alerts, syslog entries, and record entries in the output of the **logging onboard datarate** command. In an all MDS environment, only Tx datarate is required to determine overutilization. In mixed environments, with other types of switches that do not support Tx datarate, configuring Rx datarate can help to determine the ingress rate from a non-MDS switch.

  Tx datarate and Rx datarate must be configured as follows and included in the active port-monitor policy:

```
counter tx-datarate poll-interval 10 delta rising-threshold 80 event 4 falling-threshold
 79 event 4
counter rx-datarate poll-interval 10 delta rising-threshold 80 event 4 falling-threshold
 79 event 4
```

In the **show logging log** and **show logging onboard datarate** commands, the time an interface was running at a high Tx utilization is the time from rising threshold to falling threshold.

```
switch# show logging log
2018 Aug 24 13:09:07 %PMON-SLOT1-3-RISING_THRESHOLD_REACHED: TX Datarate has reached
the rising threshold (port=fc1/4 [0x1003000], value=820766704) .
2018 Aug 24 13:09:09 %PMON-SLOT12-5-FALLING_THRESHOLD_REACHED: TX Datarate has reached
 the falling threshold (port=fc12/11 [0x158a000], value=34050354) .
2018 Aug 24 13:09:18 %PMON-SLOT1-5-FALLING_THRESHOLD_REACHED: TX Datarate has reached
the falling threshold (port=fc1/4 [0x1003000], value=233513787) .
2018 Aug 24 13:09:42 %PMON-SLOT12-3-RISING_THRESHOLD_REACHED: TX Datarate has reached
the rising threshold (port=fc12/11 [0x158a000], value=878848923) .
2018 Aug 24 13:10:45 %PMON-SLOT12-5-FALLING_THRESHOLD_REACHED: TX Datarate has reached
 the falling threshold (port=fc12/11 [0x158a000], value=387111312) .
```

```
switch# show logging onboard datarate
```

```
----------------------------
    Module:  1
----------------------------

---------------------------------
 Module: 1 datarate
---------------------------------

----------------------------
 Show Clock
----------------------------
2018-08-28 15:43:33

---------------------------------
 Module: 1 datarate
---------------------------------
      - DATARATE INFORMATION FROM FCMAC

 ------------------------------------------------------------------------------------
| Interface |  Speed  |     Alarm-types      |   Rate  |         Timestamp          |
 ------------------------------------------------------------------------------------
|  fc1/94   |    4G   |  TX_DATARATE_FALLING |    57%  | Tue Aug 28 15:42:52 2018 |
|  fc1/94   |    4G   |  TX_DATARATE_RISING  |    86%  | Tue Aug 28 15:38:54 2018 |
|  fc1/94   |    4G   |  TX_DATARATE_FALLING |     8%  | Tue Aug 28 15:38:33 2018 |
|  fc1/94   |    4G   |  TX_DATARATE_RISING  |    85%  | Tue Aug 28 15:37:42 2018 |
```

Port Monitor

- Port monitor (Fibre Channel only)—Port monitor can generate alerts for various congestion-related counters. Port monitor has two thresholds that are called rising threshold and falling threshold. The rising threshold is when the counter of a port reaches or exceeds the configured threshold value. The falling threshold is when the counter of a port reaches or falls below a configured value. For each event, an alert is generated. The time that the port was between the rising threshold and falling threshold is when the event was occurring. These alerts are recorded in the RMON log in all releases.
- Port monitor does not have any effect on logging of the various congestion counters except in the case of tx-datarate and rx-datarate. From Cisco MDS NX-OS 8.2(1) and later releases, the alerts are logged in OBFL and are displayed in the **show logging onboard datarate** command. See the "Overutilization" section for the optimal tx-datarate and rx-datarate counter configuration to detect overutilization.

Table 1: Features to Detect Slow Drain, on page 35 describes the features that help detect the slow-drain condition:

Table 1: Features to Detect Slow Drain

Feature Name	Description
Port monitor's credit-loss-reco counter	Credit-loss-reco counter resets a link when there is not enough transmit credits available for 1 second for edge ports and 1.5 seconds for core ports.
Port monitor's invalid-crc counter	Invalid-crc counter represents the total number of CRC errors that a port receives.
Port monitor's invalid-words counter	Invalid-words counter represents the total number of invalid words that a port receives.

Feature Name	Description
Port monitor's link-loss counter	Link-loss counter represents the total number of link failures that a port encounters.
Port monitor's lr-rx counter	Lr-rx counter represents the total number of link reset primitive sequences that a port receives.
Port monitor's lr-tx counter	Lr-tx counter represents the total number of link reset primitive sequences that a port transmits.
Port monitor's rx-datarate counter	Rx-datarate counter receives frame rates in bytes per seconds.
Port monitor's signal-loss counter	Signal-loss counter represents the number of times a port encountered laser or signal loss.
Port monitor's state-change counter	State-change counter represents the number of times a port has transitioned to an operational up state.
Port monitor's sync-loss counter	Sync-loss counter represents the number of times a port experienced loss of synchronization in Rx.
Port monitor's tx-credit-not-available counter	Tx-credit-not-available counter increments by one if there are no transmit buffer-to-buffer credits available for a duration of 100 ms.
Port monitor's timeout-discards counter	Timeout-discards counter represents the total number of frames that are dropped at egress due to congestion timeout or no-credit-drop timeout.
Port monitor's tx-datarate counter	Tx-datarate counter represents the transmit frame rate in bytes per seconds.
Port monitor's tx-discards counter	Tx-discards counter represents the total number of frames that are dropped at egress due to timeout, abort, offline, and so on.
Port monitor's tx-slowport-count counter	Tx-slowport-count counter represents the number of times slow port events were detected by a port for the configured slowport-monitor timeout. This counter is applicable only for Generation 3 modules.
Port monitor's tx-slowport-oper-delay counter	Tx-slowport-oper-delay counter captures average credit delay (or R_RDY delay) experienced by a port. The value is in milliseconds.
Port monitor's txwait counter	TxWait counter is an aggregate time-counter that counts transmit wait time of a port. Transmit wait is a condition when a port experiences no transmit credit available (Tx B2B = 0) and frames are waiting for transmission.
Port monitor's tx-datarate-burst counter	Tx-datarate-burst counter monitors the number of times the datarate crosses the configured threshold datarate in 1 second intervals.
Port monitor's rx-datarate-burst counter	Rx-datarate-burst counter monitors the number of times the datarate crosses the configured threshold datarate in 1 second intervals.

Information About Congestion Avoidance

Congestion avoidance focuses on minimizing or completely avoiding the congestion that results from frames being queued to congested ports.

Cisco MDS switches have multiple features designed to void congestion in SAN:

- Congestion-drop timeout threshold (Fibre Channel and FCoE): The congestion-drop timeout threshold determines the amount of time a queued Fibre Channel or FCoE frame will stay in the switch awaiting transmission. Once the threshold is reached the frame is discarded as a *timeout drop*. The lower the value the quicker these queued frames are dropped and the result buffer freed. This can relieve some back pressure in the switch, especially on ISLs. By default it is 500 ms but can be configured as low as 200 ms in 1 ms increments. It is configured using the **system timeout congestion-drop** (Fibre Channel) and **system timeout fcoe congestion-drop** (FCoE) commands.
- No-credit-drop timeout threshold (Fibre Channel only): No-credit-drop timeout threshold is used to time when a Fibre Channel port is at zero Tx credits. Once a Fibre Channel port hits zero Tx credits the timer is started. If the configured threshold is reached then all frames queued to that port will be dropped regardless of their actual age in the switch. Furthermore, as long as the port remains at zero Tx credits, all newly arriving frames are immediately dropped. This can have a dramatic effect on reliving congestion especially on upstream ISLs. This allows unrelated flows to move continuously. This is off by default. If configured, it should be set to a value that is lower than the configured (or defaulted) Fibre Channel congestion-drop timeout. It is configured via the **system timeout no-credit-drop** command. The no-credit timeout functionality is only used for edge ports because these ports are directly connected to the slow-drain devices.
- Pause-drop timeout threshold (FCoE only): Pause-drop timeout threshold is used to time when a FCoE port is in a continuous state of Rx pause (unable to transmit). After an FCoE port receives a PFC pause with a non-zero quanta, the timer is started. If the port continues to receive PFC pauses with a non-zero quanta such that it remains in the Rx pause state continuously for the pause-drop threshold, then all frames queued to that port will be dropped regardless of their actual age in the switch. Furthermore, as long as the port remains in a Rx pause state, all newly arriving frames are immediately dropped. This can have a dramatic effect on relieving congestion especially on the upstream ISLs. This allows unrelated flows to move continuously. This is on by default with a value of 500 ms. If configured, it should be set to a value that is lower than the configured (or defaulted) FCoE congestion-drop timeout. It is configured via the **system timeout fcoe pause-drop** commands (available from Cisco MDS NX-OS Release 8.2(1) onwards). The FCoE pause-drop timeout functionality is only used for edge ports, because these ports are directly connected to the slow-drain devices.
- Port monitor with portguard actions of flap and error disable: For more information, see the Port Monitor section.

Information About Congestion Isolation

The Congestion Isolation feature can detect a slow-drain device via port monitor or manual configuration and isolate the slow-drain device from other normally performing devices on an ISL. After the traffic to the slow-drain device is isolated, the traffic to the rest of the normally behaving devices will be unaffected. Traffic isolation is accomplished using the following three features:

- Extended Receiver Ready—This feature allows each ISL between supporting switches to be split into four separate virtual links, with each virtual link assigned its own buffer-to-buffer credits. Virtual link 0 used to carry control traffic, virtual link 1 is used to carry high-priority traffic, virtual link 2 is used to carry slow devices, and virtual link 3 is used to carry normal traffic.

- Congestion Isolation—This feature allows devices to be categorized as slow by either configuration command or by port monitor.
- Port monitor portguard action for Congestion Isolation—Port monitor has a new portguard option to allow the categorization of a device as slow so that it can have all traffic flowing to the device routed to the slow virtual link.

Extended Receiver Ready

Note Extended Receiver Ready (ER_RDY) feature functions only on Fibre Channel Inter-Switch Links (ISL) and only between switches that support this feature.

ER_RDY primitives are used as an alternative to Receiver Ready (R_RDY). ER_RDY primitives virtualize a physical link into multiple virtual links (VLs) that are assigned individual buffer-to-buffer credits, thereby controlling the flow to the physical link. The ER_RDY feature is used by Congestion Isolation to route slow flows to a specific VL, called a low-priority VL (VL2), so that all the normal flows are unaffected. ER_RDY supports up to four VLs.

Figure 1: Traffic Flow Using Virtual Links, on page 38 shows VLs managing the good flow and slow flow. VL0 (red link) is used for control traffic, VL1 (orange link) is used for high-priority traffic, VL2 (blue link) is used for slow traffic, and VL3 (green link) is used for normal-data traffic. Slow flow detected at Host H2 is automatically assigned to VL2, which prevents the congestion of the link and allows the good flow from Host H1 to use either the VL1 or VL3 depending on the flow priority.

Figure 1: Traffic Flow Using Virtual Links

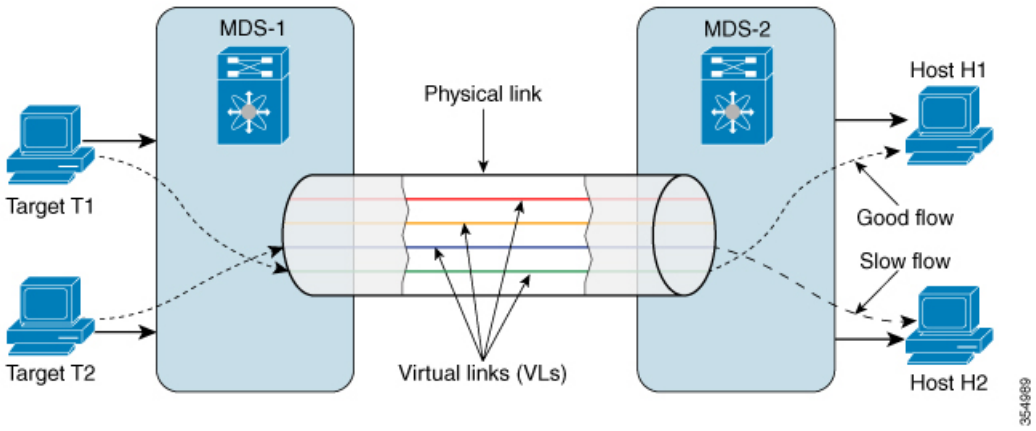


Table 2: Virtual Link-to-QoS Priority Mapping, on page 38 provides VL-to-QoS priority mapping information. Use this information while setting a zone QoS priority in a zone where Congestion Isolation is enabled in order to avoid QoS priority flow from being treated as slow flow.

Table 2: Virtual Link-to-QoS Priority Mapping

Virtual Link	QoS Priority
VL0 (control traffic)	7

VL1 (not used for any traffic)	5, 6
VL2 (slow traffic)	2, 3, 4
VL3 (normal traffic)	0, 1

Congestion Isolation

The Congestion Isolation feature uses VL capabilities to isolate the flows to the congested devices on an ISL to a low-priority VL that has less buffer-to-buffer credits than the buffer-to-buffer credits used for the normal traffic VL. Traffic in the direction of the congested device is routed to a low-priority VL. Normal devices continue to use the normal VL that has more buffer-to-buffer credits. Congested devices can be marked as slow either via the port monitor or manually.

Note Prior to Cisco MDS NX-OS Release 8.5(1), when a device is manually marked as a congested device or automatically detected as a congested device via the port monitor, the Fibre Channel Name Server (FCNS) database registers the congested-device attribute (slow-dev) for the device and distributes the information to the entire fabric. For more information, see Configuring Congestion Isolation, on page 68.

From Cisco MDS NX-OS Release 8.5(1), when a device is manually marked as a congested device or automatically detected as a congested device via the port monitor, the information about the congested device will be displayed in the FPM database and FPM distributes this information to the entire fabric. For more information, see Configuring Congestion Isolation, on page 68.

You must ensure that the following requirements are met before enabling the Congestion Isolation feature:

- Flows must traverse ISLs because Congestion Isolation functions only across Fibre Channel ISLs.
- ISLs or port channels must be in ER_RDY flow-control mode.
- If you want the port monitor to automatically detect the slow devices, the port-monitor policies must be configured to use the congestion isolation port-guard action (cong-isolate).

  Optionally, devices can be configured manually as congested devices.

Port-Monitor Portguard Action for Congestion Isolation

The cong-isolate port-monitor portguard action automatically isolates a port after a given event rising-threshold is reached.

Note Absolute counters do not support portguard actions. However, the tx-slowport-oper-delay absolute counter supports Congestion Isolation portguard action (cong-isolate).

The following is the list of counters that you can use to trigger the Congestion Isolation port-monitor portguard action (cong-isolate):

- credit-loss-reco
- tx-credit-not-available

- tx-slowport-oper-delay
- txwait

Congestion Isolation Recovery

Prior to Cisco MDS NX-OS Release 8.5(1) when a slow device was detected, the flows to the congested device were automatically moved to the low-priority VL using the Congestion Isolation feature. After the congested device recovered from congestion, the flows had to be manually moved flows from the low-priority VL to normal VL.

From Cisco MDS NX-OS Release 8.5(1), the Congestion Isolation Recovery feature automatically recovers the traffic to congested device from a low-priority VL to the normal VL. This recovery is done without any user intervention unlike the Congestion Isolation feature where user had to manually recover the traffic flowing to the congested device from low-priority VL to normal VL after the device had recovered from congestion.

The *cong-isolate-recover* portguard action is available in the port monitor policy for the supported slowdrain counters.

The recovery process uses a **recovery-interval** to check if the traffic going to congested device in the low-priority VL can be moved back to the normal VL. The following is the process that is used for recovery:

1. A device is identified as a congested device when the port monitor counter detects a rising threshold. After the device is identified as a congested device, the traffic destined to the congested device is moved to the low-priority VL.
2. When port monitor detects a falling threshold for the congested device, a recovery interval (15 minutes by default) is started. During this interval, if the port monitor counter stays at or below the falling threshold continuously, then the device is no longer marked as congested device and the traffic destined to the device is moved from the low-priority VL to normal VL.

   However, if port monitor detects an event threshold that is more than the falling threshold before the expiry of the recovery interval, the interval is discarded and the device remains classified as a congested device. The recovery timer is again started when the next falling threshold is detected by the port monitor. The recovery interval can be configured. For more information, see Configuring Congestion Isolation Recovery, on page 71.
3. Also, the Congestion Isolation Recovery feature allows you to determine the number of times the traffic destined to a congested device can toggle between congestion isolated and recovered, which is known as the *number of occurrences*. If the traffic destined to a congested device toggles between congestion isolated and recovered for the specified number of occurrences within a specified duration called **isolate-duration**, then on the last occurrence the device would be marked as a congestion isolated device and would not be allowed to recover until the isolate-duration expires. Isolate duration is a recurring interval and starts when a port monitor policy is activated.

For example, let us consider a device P1 that is detected as a congested device. The traffic destined to the device is moved to low-priority VL and has recovered back after sometime. The traffic destined to the device P1 keeps being detected as slow and then recovering. In such cases, you can configure the number of such transitions or occurrences known as *number of occurrences* for a specified **isolate-duration**. Suppose you have chosen this value to be 3 and the isolate duration to be 24 hours. When an event threshold that is more the falling threshold is detected for P1 for the third time in the first 2 hours of activating the isolate duration, P1 is marked as congested device. The flows will be moved to low-priority VL for the remaining 22 hours and any subsequent falling threshold detected is ignored. The device P1 would remain as congested device until the end of 22 hours after which the device is recovered and monitored for an event threshold that is more

than the falling threshold again. However, you can manually recover flows from low-priority VL to normal VL. For more information, see Configuring Excluded List of Congested Devices, on page 69.

Note The **isolate-duration** starts only after the corresponding port monitor policy is activated.

The following is the list of counters that you can use to trigger the Congestion Isolation Recovery port-monitor portguard action (cong-isolate-recover):

- credit-loss-reco
- tx-credit-not-available
- tx-slowport-oper-delay
- txwait

Fabric Notifications—FPIN and Congestion Signal

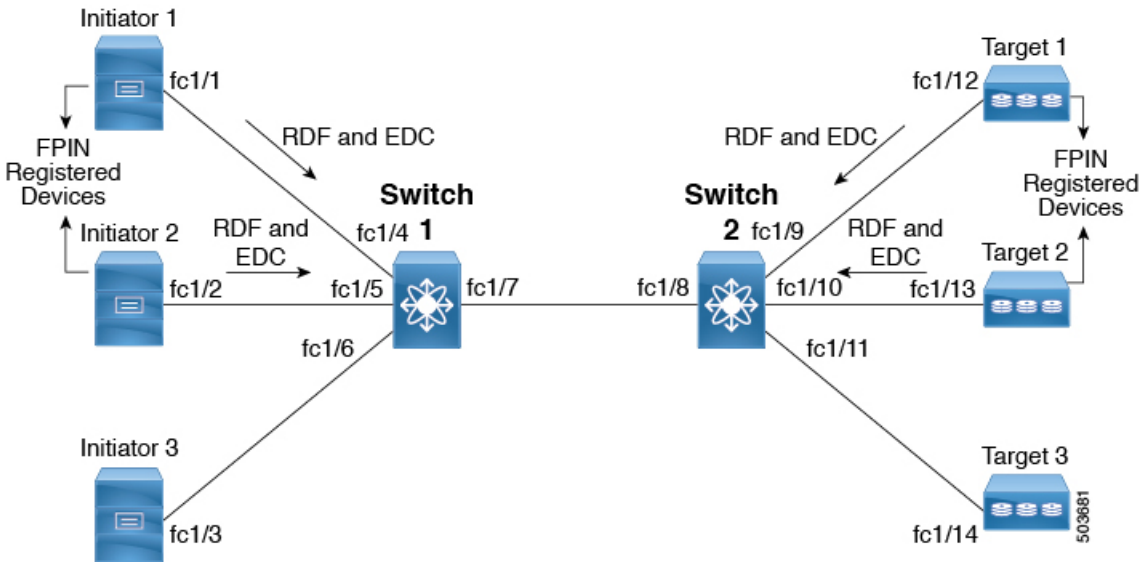
Fabric Notifications are used to notify end devices of performance impacting conditions and behaviors that affect the normal flow of IO such as link integrity degradation and congestion. The information provided may be used by the end devices to modify their behavior to address the reported conditions. The functions include notifications in the form of ELS (Extended Link Service) primitives and Signals primitives.

The following capabilities for operations supporting fabric notifications are added in Fabric Performance Monitor (FPM):

- Registration: Register Diagnostic Functions (RDF) and Exchange Diagnostic Capabilities (EDC) ELS exchange between end device and a switch registering for fabric notifications RDF requests FPM to register the port on the end device that wants to receive Fabric Performance Impact Notifications (FPIN) ELS when link integrity degradation and congestion are detected in the fabric. EDC requests FPM to register the port on end device that wants to receive congestion signal primitives on detecting congestion events on the attached port.

  Figure 2: RDF and EDC ELS Exchange, on page 42 displays a sample topology where Initiator 1, Initiator 2, Target 1, and Target 2 are registered for FPIN via RDF and EDC. Initiator 3 and Target 3 are not registered for FPIN.

Figure 2: RDF and EDC ELS Exchange



- Notifications: FPIN ELS alerts registered end devices about occurrences that impact performance and contains the descriptions of the event occurrences.

  The following are the types of events for which FPIN is generated:

  - Congestion: A congestion condition that is detected at an F port will be notified to the connected end device.
  - Peer congestion: A congestion condition that is detected at an F port will be notified to all the devices communicating via the port. The information that is notified includes the type of slowdrain condition and the list of impacted devices.
  - Link integrity: A condition that checks for port integrity The information that is notified includes the reason, such as, link failure, loss of signal, and so on, and a threshold value that was exceeded.

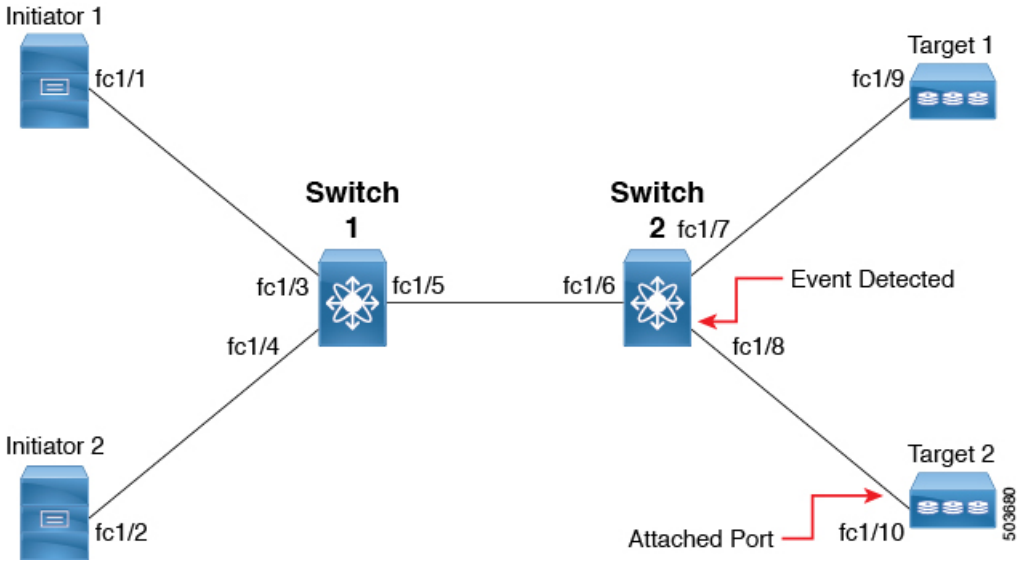
    The following is the list of counters that you can use to trigger the link integrity events:

    - link-loss
    - sync-loss
    - signal-loss
    - invalid-words
    - invalid-crc

Note The Congestion Isolation Recovery feature is not supported on these counters. For more information, see Congestion Isolation Recovery, on page 40.

Figure 3: FPIN Events, on page 43 displays a sample topology where all the devices are configured in a single zone. An event is detected at port fc1/8 and Target 2 is the attached port or peer port.

Figure 3: FPIN Events



The following provides how the information is shared between the devices when events are detected:

- Congestion: When a congestion event is detected at port fc1/8, an FPIN congestion descriptor is sent to Target 2.
- Peer congestion: When a congestion event is detected at port fc1/8 and FPIN peer congestion event is sent to Initiator 1, Initiator 2, and Target 1 containing the pWWN list of Target 2.
- Link integrity: When a link integrity event is detected at port fc1/8, FPIN link integrity is sent to Initiator 1, Initiator 2, and Target 1 with the pWWN list of Target 2 and FPIN link integrity is also sent to Target 2 with the pWWN list of Initiator 1, Initiator 2, and Target 1.

Note Cisco MDS port does not handle FPINs received from adjacent devices. Instead, they are discarded.

- Signals: Congestion signal primitives sent to a receiving port of an end device by an attached switch port indicating TxWait conditions on the ports that have exceeded a threshold. End devices register with switches for receiving congestion signal primitives at specific interval. This interval is negotiated by the end device with the switch and cannot be configured. You can check this interval using the **show fpm registration congestion-signal** command. Depending on the type of event detected, port monitor sends warning or alarm signal primitives at the specified interval.

  The following types of congestion signal primitives are supported and are configurable in the port monitor policy for the TxWait counter:

  - Warning congestion signal: This signal is sent when the TxWait condition on a port exceeds warning threshold.
  - Alarm congestion signal: This signal is sent when the TxWait condition on a port exceeds alarm threshold.

FPM receives notifications about link integrity degradation and congestion from port monitor when its counters detect a configured rising threshold.

The following port monitor counters support FPIN portguard actions to check the link integrity degradation:

- LinkFailures
- SyncLoss
- SigLoss
- Invalid TxWords
- InvalidCRCs

The TxWait port monitor counter supports FPIN portguard action to check congestion. TxWait also supports configuring congestion signal.

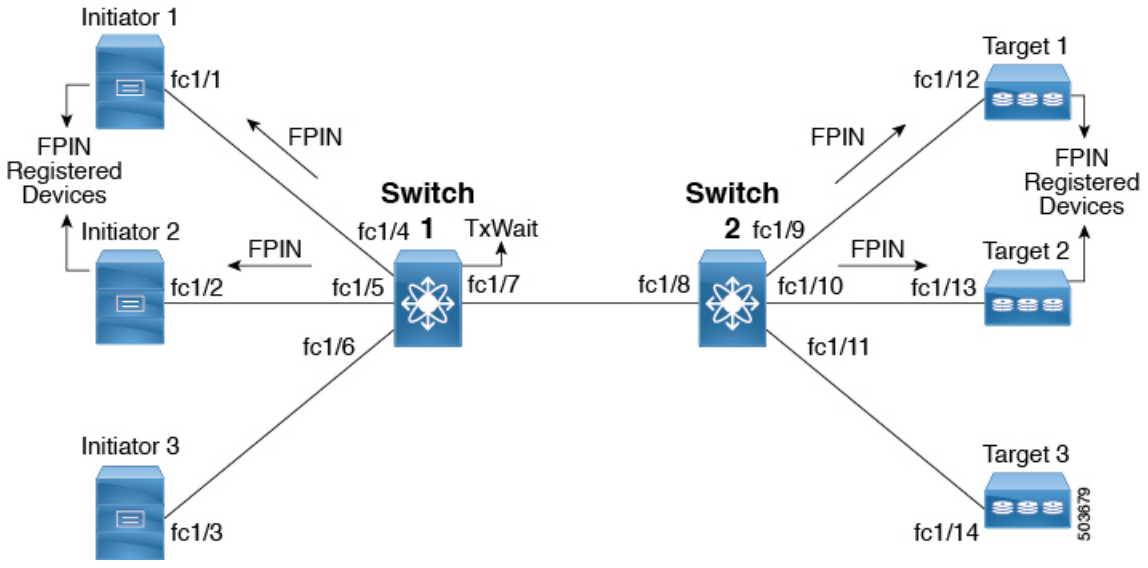
Recovery of congestion events is also notified through FPIN to the end device. Recovery of congestion events is notified from port monitor when a counter value remains below the falling threshold for the **recovery-interval**. For information about configuring recovery interval for FPIN, see Configuring the Port-Monitor Portguard Action for FPIN, on page 73.

For configuring FPIN and congestion signal fabric notifications, see Configuring EDC Congestion Signal, on page 75.

FPM can manually classify a device as congested and also exclude a device from detection of link integrity degradation and congestion. For more information, see Configuring Fabric Notifications, on page 73.

Figure 4: Fabric Notifications, on page 44 displays a sample topology where end devices Initiator 1, Initiator 2, Target 1, and Target 2 are registered for FPIN via RDF and EDC. Initiator 3 and Target 3 are not registered for FPIN. When Initiator 1 becomes slow and TxWait is seen on fc1/4, FPIN is sent to all zoned end devices of Initiator 1 that are registered for FPIN and not to devices that are not registered for FPIN.

Figure 4: Fabric Notifications

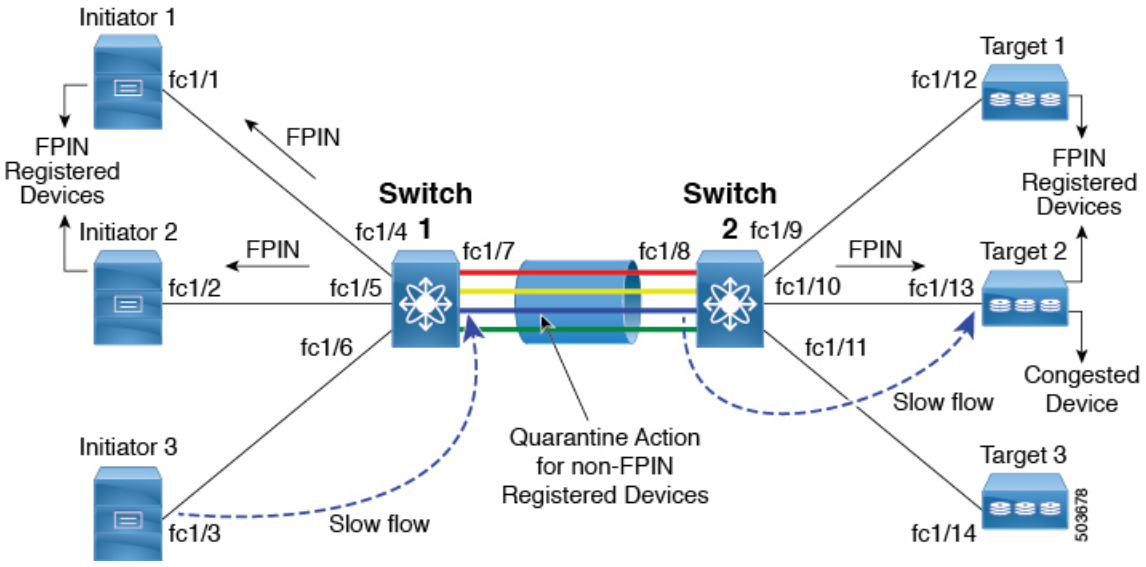


FPIN and ER_RDY

FPIN can also work in conjunction with the ER_RDY feature to isolate flows to the low-priority VL if end devices have not registered themselves with RDF for fabric notifications. The recovery of flows from low-priority VL to normal VL happens when port monitor notifies FPM about the recovery. For FPIN to work with the ER_RDY feature, you need to enable the ER_RDY feature. For more information, see Enabling Extended Receiver Ready, on page 67.

Figure 5: FPIN and ER_RDY, on page 45 displays a sample topology where Initiator 1, Initiator 2, Target 1, and Target 2 are registered for FPIN through RDF. Also, Initiator 1 is zoned with Target 1 and Target 2, Initiator 2 is zoned with Target 2 and Target 3, and Initiator 3 is zoned to Target 2 and Target 3. Initiator 3 and Target 3 are not registered for FPIN. Congestion is detected at Target 2 and all the zoned devices of Target 2 that are registered for FPIN are notified about the congested device. Initiator 3 is not registered for FPIN and as we have ER_RDY enabled, the flow from Initiator 3 to Target 2 uses the low-priority VL.

Figure 5: FPIN and ER_RDY



Dynamic Ingress Rate Limiting

Dynamic Ingress Rate Limiting (DIRL) is used to automatically limit the rate of ingress commands and other traffic to reduce or eliminate the congestion that is occurring in the egress direction. DIRL does this by reducing the rate of IO solicitations such that the data generated by these IO solicitations matches the ability of the end device to actually process the data without causing any congestion. As the device's ability to handle the amount of solicited data changes, DIRL, will dynamically adjust seeking to supply it the maximum amount of data possible without the end device causing congestion. After the end device recovers from congestion, DIRL will automatically stop limiting the traffic that is sent to the switch port.

In case of slow drain and over utilization, the assumption is that if the rate of IO solicitation requests is reduced then this will make a corresponding reduction in the amount of data solicited and being sent to the end device. By reducing the amount of data this will resolve both the slow drain and over utilization cases.

DIRL is comprised of two functions and can perform equally well on congestion caused both slow drain and over utilization:

- Port monitor: Detects slow drain and over utilization conditions and if the portguard action is set as **DIRL**, it notifies FPM. Port monitor portguard action **DIRL** can be configured on the following counters:
  - txwait: Use for detection of slow drain.
  - tx-datarate: Used for detection of over utilization.
  - tx-datarate-burst: Use for detection of over utilization.
- FPM: DIRL actions are taken by FPM as notified by port monitor. On detecting a rising threshold from port monitor, FPM does rate reduction causing the rate of ingress traffic to be reduced. On detecting the value of a counter being below the falling threshold continuously for the DIRL recovery interval, FPM does rate recovery.

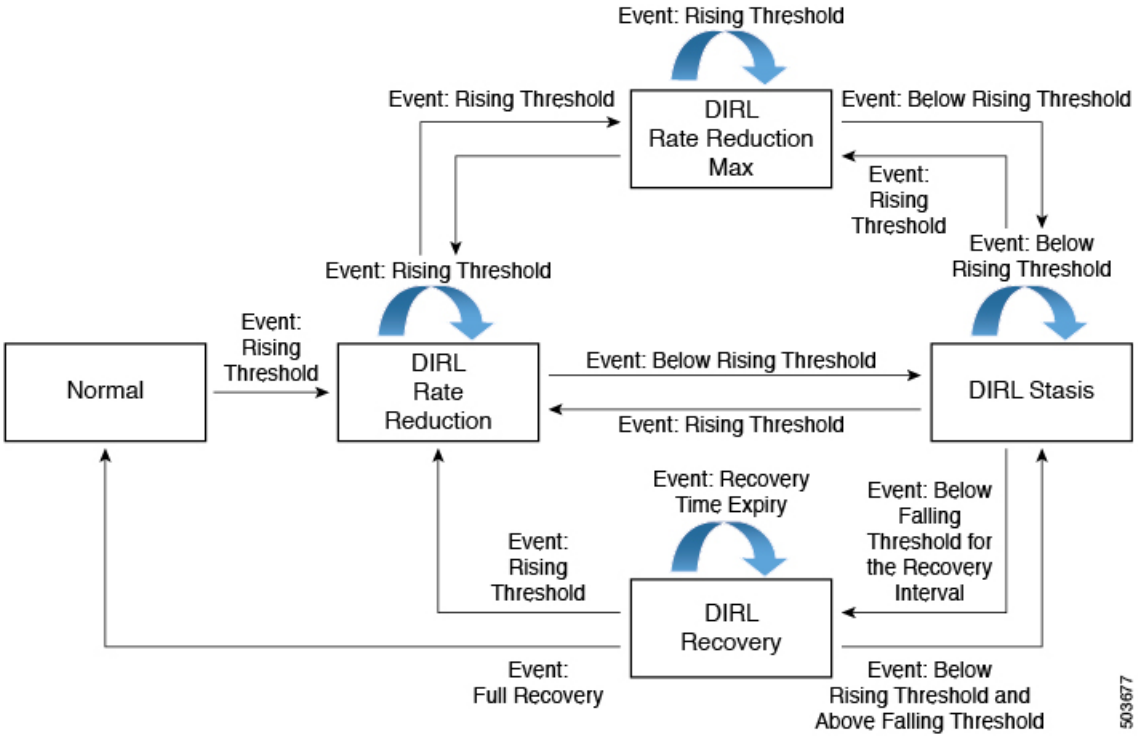
After the port monitor policy is configured with the DIRL portguard action and activated, all non-default F ports are monitored by default and FPM is notified if congestion is detected on any of these ports. However, you can manually exclude certain interface from being monitored. For more information, see Configuring Excluded List of Congested Devices, on page 69.

Note If an interface is configured using static ingress rate limit by using the **switchport ingress-rate** *limit* command, then DIRL will not function for that port. However, a port that is subject to DIRL can be overridden by static ingress rate limit.

The following are the different transition states of DIRL:

- Normal: The state in which a port is functioning normally and state before it enters DIRL Rate Reduction. After full recovery, port returns to the Normal state.
- DIRL Rate Reduction: The state in which an event rising threshold triggers the DIRL rate reduction process.
- DIRL Rate Reduction Maximum: The state in which the DIRL rate reduction has reached its maximum value and more rising thresholds events are detected.
- DIRL Stasis: The state in which an event below rising threshold and above falling threshold is detected. This state will transition to DIRL Recovery state when an event below falling threshold is detected for the configured **recovery-interval**.
- DIRL Rate Recovery: The state in which the DIRL rate recovery happens on detecting an event below falling threshold for the configured **recovery-interval**. This state will transition to the Normal state after the port recovers completely from DIRL. This state is a recurring state and there will be multiple rate recoveries before the ports are completely recovered from DIRL. This state will transition to the DIRL Stasis state when an event below rising threshold and above falling threshold is detected.

Figure 6: Different States of DIRL



Let us consider the following example where the DIRL rate recovery process has started on port fc4/12 after detecting an event rising threshold:

```
switch# show fpm ingress-rate-limit events interface fc4/12
------------------------------------------------------------------------------------------------------------------------
Interface|Counter|  Event |       Action  |Operating | Input  |Output|Current|Applied|
Time
         |       |        |               |port-speed| rate   | rate |  rate |  rate |

                                         |   Mbps   | Mhps   | Mbps |limit %|limit %|
------------------------------------------------------------------------------------------------------------------------
fc4/12    txwait  rising    rate-reduction  16000.00   8853.37 8853.10 77.010  31.563  Mon
 Jan 18 22:34:44 2021
fc4/12    txwait  recovery  rate-recovery   16000.00   8369.35 8369.35 61.608  77.010  Mon
 Jan 18 22:34:37 2021
fc4/12    txwait  recovery  rate-recovery   16000.00   6697.13 6697.16 49.287  61.608  Mon
 Jan 18 22:33:37 2021
fc4/12    txwait  recovery  rate-recovery   16000.00   5359.97 5359.95 39.429  49.287  Mon
 Jan 18 22:32:36 2021
fc4/12    txwait  recovery  rate-recovery   16000.00   4288.87 4288.86 31.543  39.429  Mon
 Jan 18 22:31:36 2021
fc4/12    txwait  rising    rate-reduction  16000.00   8847.91 8848.01 100.000 31.543  Mon
 Jan 18 22:30:24 2021
```

The following are the actions that are initiated by DIRL depending on the type of event detected on the port:

Note The events are listed in reverse chronological order with the most current event at the top.

1. An event rising threshold is detected on the port and DIRL is initiated for the port. The port ingress traffic rate is reduced to 50% of its current rate.
2. In the next polling interval, the **recovery-interval** expires without detecting a rising threshold. The port ingress traffic is increased by 25% of its current capacity.
3. In the next polling interval, the **recovery-interval** expires without detecting a rising threshold. The port ingress traffic is increased by 25% of its current capacity.
4. In the next polling interval, the **recovery-interval** expires without detecting a rising threshold. The port ingress traffic is increased by 25% of its current capacity.
5. In the next polling interval, the **recovery-interval** expires without detecting a rising threshold. The port ingress traffic is increased by 25% of its current capacity.
6. In the next polling interval, an event rising threshold is detected on the port and DIRL is initiated for the port. The port ingress traffic is reduced again to 50% of its current rate.

Static Ingress Port Rate Limiting

A static port rate limiting feature helps control the bandwidth for individual Fibre Channel ports using the **switchport ingress-rate** *limit* command. Port rate limiting is also referred to as ingress rate limiting because it controls ingress traffic into a Fibre Channel port. The feature controls traffic flow by slowing the rate of B2B credits transmitted from the FC port to the adjacent device. Port rate limiting works on all Fibre Channel ports. Prior to Cisco MDS NX-OS Release 8.5(1), the rate limit ranges from 1 to 100%. From Cisco MDS NX-OS Release 8.5(1), the limit ranges from 0.0126 to 100%. The default rate limit is 100%.

Starting from Cisco MDS NX-OS Release 8.5(1), the FPM feature needs to be configured before configuring the dynamic or static ingress port rate limiting feature on all Cisco MDS switches except Cisco MDS 9250i and MDS 9148S switches. Prior to Cisco MDS NX-OS Release 8.5(1) or on Cisco MDS 9250i and MDS 9148S switches, static ingress port rate limiting can be configured on all Cisco MDS switches and modules only if the QoS feature is enabled.

Guidelines and Limitations for Congestion Management

Guidelines and Limitations for Congestion Detection

The **show tech-support slowdrain** command contains all the congestion detection indications, counters, and log messages as well as other commands that allow an understanding of the switches, MDS NX-OS versions, and topology. Since, congestion can propagate from one switch to another, the **show tech-support slowdrain** command should be gathered from all the switches at approximately the same time to have the best view of where the congestion started and how it spread. This can be easily done via the DCNM SAN client using the **Tools**-> **Run CLI** feature. This feature will issue a command or commands to all the switches in the fabric and consolidates the individual switch output files into a single fabric zip file.

Some commands display simple counters such as the **show interface counters** command, whereas some commands display counter information with accompanying date and time stamps. The commands that display counters with accompanying date and time stamps are mostly the **show logging onboard** commands.

There are various *sections* of show logging onboard that contain information pertaining to slow drain and over utilization. Most *sections* will update periodically and include counters only when they actually change in the prior interval. Different sections have different update periods. They are:

- Error-stats—Includes many error counters accompanying date and time stamps
- Txwait—Includes interfaces that record 100 ms or more of TxWait in a 20-second interval. The values displayed are not the current value of TxWait, but only deltas from the previous 20-second interval. If TxWait incremented by the equivalent of less than 100 ms there is no entry.
- Rxwait—Includes interfaces that record 100 ms or more of RxWait in a 20-second interval. The values displayed are not the current value of RxWait, but only deltas from the previous 20-second interval. If RxWait incremented by the equivalent of less than 100 ms there is no entry.

When a counter increments in the interval the current value of the counter is displayed along with the date and time when the counter was checked. To determine the amount the counter incremented, the delta value, in the interval the current value must be subtracted from the previously recorded value.

For example, the following show logging onboard error-stats output shows that when the counter was checked at 01/12/18 11:37:55 the timeout-drop counter, F16_TMM_TOLB_TIMEOUT_DROP_CNT, for port fc1/8 was a value of 743. The previous time it incremented was 12/20/17 06:31:47 and it was a value of 626. This means that since error-stats interval is 20 seconds, between at 01/12/18 11:37:35 and at 01/12/18 11:37:55 the counter incremented by 743 – 626 = 117 frames. There were 117 frames discarded at timeout-drops during that 20-second interval ending at 01/12/18 11:37:55.

```
switch# show logging onboard error-stats

----------------------------
 Show Clock
----------------------------
2018-01-24 15:01:35

---------------------------------
 Module: 1 error-stats
---------------------------------


----------------------------------------------------------------------------------------------------
```

```
 ERROR STATISTICS INFORMATION FOR DEVICE DEVICE: FCMAC
-------------------------------------------------------------------------------------------------

    Interface      |                                          |         |    Time Stamp
      Range        |                Error Stat Counter Name   |  Count  |MM/DD/YY HH:MM:SS
                   |                                          |         |
-------------------------------------------------------------------------------------------------
fc1/8              |F16_TMM_TOLB_TIMEOUT_DROP_CNT             |743      |01/12/18 11:37:55
fc1/8              |F16_TMM_TOLB_TIMEOUT_DROP_CNT             |626      |12/20/17 06:31:47
fc1/5              |F16_TMM_TOLB_TIMEOUT_DROP_CNT             |627      |12/20/17 06:31:47
fc1/3              |F16_TMM_TOLB_TIMEOUT_DROP_CNT             |556      |12/20/17 06:31:47
fc1/8              |F16_TMM_TOLB_TIMEOUT_DROP_CNT             |623      |12/20/17 04:05:05
```

Guidelines and Limitations for Congestion Avoidance

The default value for system timeout congestion-drop is 500 ms. This value can be safely reduced to 200 ms.

System timeout no-credit-drop is disabled by default. When configured, this feature reduces the effects of slow drain in the fabric. However, if it is configured to a value that is too low, it can cause disruption. The disruption is caused because many frames are discarded when a device withholds credits for even a short duration. The lower the value, the quicker it can start discarding frames that are queued from an upstream ISL to this (slow) port. This will relieve the back pressure or congestion on that ISL and allow other normally performing devices to continue their operation. The actual value chosen is fabric and implementation dependent.

Following are some guidelines for choosing the system timeout congestion-drop value:

- 200 ms—Safe value for most fabrics
- 100 ms—Aggressive value
- 50 ms—Very aggressive value

Generally, before configuring the no-credit-drop value, the switches should be checked for the presence of large amounts of continuous time at zero Tx credits. The **show logging onboard start time mm/dd/yy-hh:mm:ss error-stats** command can be run looking for instances of the FCP_SW_CNTR_TX_WT_AVG_B2B_ZERO counter indicating 100ms intervals at zero credits. Additionally, the **port-monitor tx-credit-not-available** and the **show system internal snmp credit-not-available** commands will show similar information. Only when the fabric only shows very limited amounts of 100ms at zero Tx credits should no-credit-drop be considered. If there are large amounts of ports with 100ms at zero Tx credits, then the problems with those end devices should be investigated and resolved prior to configuring no-credit-drop.

Note No-credit-drop can only be configured for ports that are classified *logical-type edge*. These are typically F ports.

Slowport-monitor, if configured, must have a value lower than no-credit-drop since it will only indicate a slow port if the port has no credits for at least the amount of time configured and there are frames queued for transmit. Since no-credit-drop will drops any frames queued for transmit, if no-credit-drop is configured for a value equal to or less than slowport-monitor, there will be no frames queued for transmit and slowport-monitor will not detect the slow port.

Guidelines and Limitations for Congestion Isolation

Extended Receiver Ready

- ER_RDY is supported only on Fibre Channel ports on Cisco MDS 9700 Series with Cisco MDS 9700 16-Gbps Fibre Channel Switching Module (DS-X9448-768K9), Cisco MDS 9000 Series 24/10 SAN Extension Module (DS-X9334-K9) (Fibre Channel ports only), Cisco MDS 9700 48-Port 32-Gbps Fibre Channel Switching Module (DS-X9648-1536K9), MDS 9396S, MDS 9132T, MDS 9148T, MDS 9220i, and MDS 9396T switches. In a fabric consisting of supported and unsupported switches (mixed fabric), this feature may not work effectively. In a mixed fabric, ER_RDY flow-control mode is used only between supported switches and R_RDY flow-control mode is used between unsupported switches.
- Trunking must be enabled on all ISLs in the topology for ER_RDY flow-control mode to work.
- After the **system fc flow-control er_rdy** command is configured on both the local switch and its adjacent switch, the ISLs connecting the switches should be flapped to put the ISLs in ER_RDY flow-control mode. In port channels, these links can be flapped one at a time, preventing loss of connectivity.
- For migration purposes, port channels can have their member links in both R_RDY and ER_RDY flow-control modes. This is to facilitate nondisruptive conversion from R_RDY to ER_RDY flow-control mode. Do not allow this inconsistent state to persist longer that it takes to perform the conversion from R_RDY to ER_RDY flow-control mode.
- VL1 is reserved for host bus adapter (HBA) and zone quality of service (QoS).
- Inter VSAN Routing (IVR), Fibre Channel Redirect (FCR), Fibre Channel Over TCP/IP (FCIP), and Fibre Channel over Ethernet (FCoE) are not supported in ER_RDY flow-control mode.
- From Cisco MDS NX-OS Release 8.5(1), use IOD only if your environment cannot support out-of-order frame delivery. To achieve In-Order Delivery (IOD), enable IOD using the **in-order-guarantee vsan** *id*. When a flow moves from normal VL to slow VL or vice versa, to achieve IOD functionality traffic disruption may be seen. Lossless IOD is not guaranteed.

  Prior to Cisco MDS NX-OS Release 8.5(1), In-Order Delivery (IOD) may get affected when the flow-control mode is initially set to ER_RDY and when the device's flows are moved from one VL to another VL.
- Switches running releases prior to Cisco MDS NX-OS Release 8.1(1) in a fabric are unaware of slow devices. Upon upgrading to Cisco MDS NX-OS Release 8.1(x) or later, these switches become aware of the slow devices.
- If you have configured the buffer-to-buffer credits using the **switchport fcrxbbcredit** *value* command in the Cisco MDS NX-OS Release 7.3(x) or earlier, upgraded to Cisco MDS NX-OS Release 8.1(1), and set flow-control mode to ER_RDY, the buffer-to-buffer credits that are already configured get distributed to the VLs in the following manner:
  - If the buffer-to-buffer credits value that is configured is 50, the default buffer-to-buffer credit values 5, 1, 4, and 40 are allocated to VL0, VL1, VL2, and VL3 respectively.
  - If the buffer-to-buffer credits value that is configured is more than 34 and less than 50, the buffer-to-buffer credits get distributed in the ratio 5:1:4:40.
  - If the buffer-to-buffer credits value that is configured is more than 50, the default values 5, 1, 4, and 40 are allocated to VL0, VL1, VL2, and VL3 respectively. The remaining buffer-to-buffer credits get distributed in the ratio 15:15:40:430 (VL0:VL1:VL2:VL3).

- If you are upgrading or if you are in the Cisco MDS NX-OS Release 8.1(1), if ER_RDY was enabled, and if the buffer-to-buffer credits value that is configured is less than 34, the VLs are stuck in the initialization state because the control lane (VL0) is allocated 0 credits. To recover from this situation, shutdown the link and allocate more than 34 buffer-to-buffer credits using the **switchport fcrxbbcredit** *value* or allocate at least one buffer-to-buffer credit to VL0, using the **switchport vl-credit vl0** *value* **vl1** *value* **vl2** *value* **vl3** *value* command.

Note The sum of the buffer-to-buffer credits configured for VLs cannot exceed 500.

- If you had configured the buffer-to-buffer credits using the **switchport fcrxbbcredit** *value* **mode E** command, and used the **switchport vl-credit vl0** *value* **vl1** *value* **vl2** *value* **vl3** *value* command to set the new buffer-to-buffer credits values for the VLs, the sum of the configured buffer-to-buffer credits for VLs are pushed to the **switchport fcrxbbcredit** *value* **mode E** command.
- Use the **no switchport fcrxbbcredit** *value* or **switchport vl-credit default** command to set the default buffer-to-buffer credits value for the VLs.
- If you have configured the extended buffer-to-buffer credits using the **switchport fcrxbbcredit extended** *value* in the Cisco MDS NX-OS Release 7.3(x) or earlier, upgraded to Cisco MDS NX-OS Release 8.1(1), and set the flow-control mode to ER_RDY, the extended buffer-to-buffer credits that are already configured are distributed to the VLs in the following manner:
  - If the buffer-to-buffer credits value that is configured is less than 50, the minimum values 5, 1, 4, and 40 are allocated to VL0, VL1, VL2, and VL3 respectively.
  - If the buffer-to-buffer credits value that is configured is more than 34 and less than 50, the buffer-to-buffer credits get distributed in the ratio 5:1:4:40.
  - If the buffer-to-buffer credits value that is configured is more than 50, the minimum values 15, 15, 4, and 430 are allocated to VL0, VL1, VL2, and VL3 respectively. The remaining buffer-to-buffer credits are distributed in the ratio 30:30:100:3935 (VL0:VL1:VL2:VL3).
  - If you are upgrading to or if you are in the Cisco MDS NX-OS Release 8.1(1), ER_RDY is enabled, and the buffer-to-buffer credits value configured is less than 34, the VLs are stuck in the initialization state because the control lane (VL0) is allocated 0 credits. To recover from this situation, shutdown the link and allocate more than 34 buffer-to-buffer credits using the **switchport fcrxbbcredit** *value* or allocate at least one buffer-to-buffer credit to VL0, using the **switchport vl-credit vl0** *value* **vl1** *value* **vl2** *value* **vl3** *value* command.

Note The sum of the extended buffer-to-buffer credits configured for VLs cannot exceed 4095 on a Cisco MDS 9700 16-Gbps Fibre Channel Switching Module, and 8191 on a Cisco MDS 9700 48-Port 32-Gbps Fibre Channel Switching Module, MDS 9132T, MDS 9148T, MDS 9220i, and MDS 9396T switches.

- You cannot configure regular buffer-to-buffer credits after you configure the extended buffer-to-buffer credits. You must first disable the extended buffer-to-buffer credits using the **no fcrxbbcredit extended enable** command and then configure the regular buffer-to-buffer credits.

- You cannot disable the extended buffer-to-buffer credits configuration even if one link is running in the extended buffer-to-buffer credits mode.
- ER_RDY is not supported on interfaces whose speed is set to 10-Gbps.

Congestion Isolation

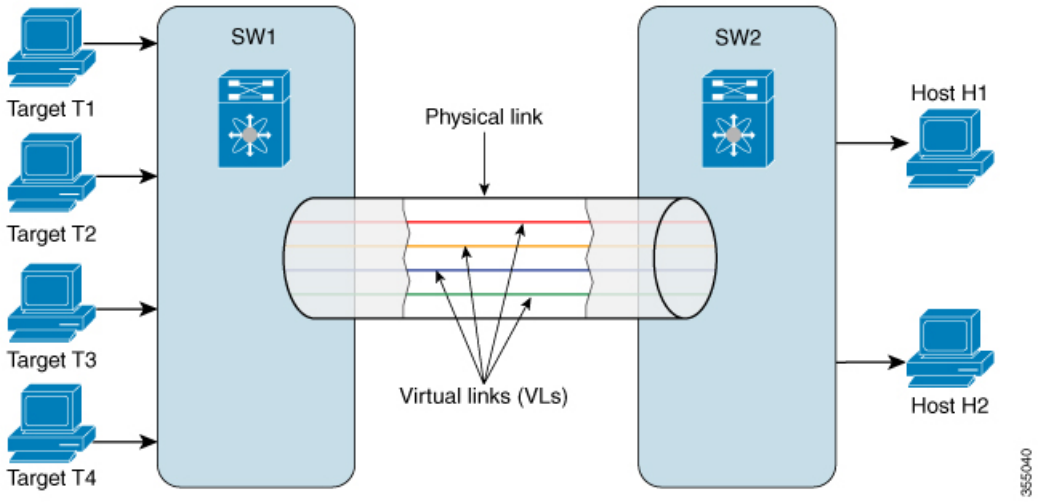
- Congestion Isolation is disabled by default.
- The port monitor portguard action for Congestion Isolation is not supported on E (core) ports. Consequently, it should only be configured on a *logical-type edge* port-monitor policy.

  If you are upgrading to Cisco MDS NX-OS Release 8.5(1) or later release and if you have the cong-isolate portguard action configured on a *logical-type core* policy, then you must remove this policy before upgrading.
- Congestion Isolation and its configurations are applicable only to the switch being configured, and not to the entire fabric.
- If you enable the ER_RDY and Congestion Isolation features on a supported switch before adding it to a fabric that is using ER_RDY flow-control mode, the ISLs that are connected between the supported switch and its adjacent switch are automatically in the ER_RDY flow-control mode and you need not flap the links on the switch for the links to use the ER_RDY flow-control mode.
- In a fabric consisting of supported and unsupported switches, Congestion Isolation functions as desired only between supported switches. Congestion Isolation functionality between unsupported devices is unpredictable.
- After a device is detected as slow, only the traffic moving in the direction of the slow device is routed to a low-priority VL (VL2). Traffic in the reverse direction is not classified as slow, and is unaffected.
- Prior to Cisco MDS NX-OS Release 8.5(1), when a slow device is detected or a device is configured as slow, the switch sends an FCNS notification to all the other switches that are capable of supporting the Congestion Isolation feature and also to the switches that may not have this feature enabled. If the switch is capable of supporting this feature but does not have it enabled, then the FCNS notification is rejected and the following messages are displayed at the originating switch:
  - %FCNS-2-CONGESTION_ISOLATION_FAILURE: %$VSAN vsan-id%$ SWILS-RJT received from domain domain-id for congestion-isolation. Issue includes CLI/FCNS DB refresh on the remote domain.
  - %FCNS-2-CONGESTION_ISOLATION_INT_ERROR: %$VSAN 237%$ Error reason: Congestion-Isolation disabled on the remote domain. Please enable the feature on the remote domain.

  If the Congestion Isolation feature is configured on all the intended switches, these messages do not have any negative effect and can be ignored. For example, if a Cisco MDS switch is connected via FCoE ISLs then the Congestion Isolation feature does not apply to this switch and these messages can be ignored. However, ER_RDY and Congestion Isolation features can be configured on an FCoE connected switch preventing the messages from being displayed.
- Figure 7: Traffic Flow When Multiple Targets are Connected shows a fabric that has multiple targets connected to switch SW1 and two hosts (Host H1 and Host H2) connected to switch SW2. Both hosts H1 and H2 are zoned with all four targets T1 to T4. Host H2 is detected as a slow device. The traffic from the targets to host H2 is marked as slow and is routed to VL2. Since VL2 has fewer buffer-to-buffer credits and because host H2 is itself withholding buffer-to-buffer credits from SW2, traffic on VL2 from

SW1 to SW2 will be constrained by what host H2 can receive. This results in switch SW1 withholding buffer-to-buffer credits from all four targets T1 to T4. This will affect all traffic being sent by the targets to any destination. Consequently, other hosts zoned with the targets, like host H1, will also see their traffic affected. This is an expected behavior. In such a situation, resolve the slow-drain condition for the traffic to flow normally.

Figure 7: Traffic Flow When Multiple Targets are Connected



- If in a zone, the zone QoS priority is set to medium and Congestion Isolation is enabled on the switches in the zone, the traffic with zone QoS priority medium are treated as slow, and Congestion Isolation routes the traffic to the low-priority VL (VL2). To avoid this situation, set the zone QoS priority to low or high.
- When a link to a Cisco NPV switch carrying multiple fabric logins (FLOGIs) is detected as a slow device, all the devices connected to the Cisco NPV switch are marked as slow devices.
- Downgrading from a supported release to an unsupported release is disabled after the Congestion Isolation and Congestion Isolation Recovery features are enabled. To downgrade to an unsupported release:

  1. If **cong-isolate** or **cong-isolate-recover** port monitor portguard action is configured in a port monitor policy, remove the action from the policy.
  2. Remove any devices that are manually included or excluded as slow-drain devices.
  3. Disable the Congestion Isolation feature.
  4. Reset the flow-control mode to R_RDY.
  5. Flap all the ISLs.
  6. Display the ISLs currently functioning in R_RDY mode.
  7. Display the ISLs currently functioning in ER_RDY mode.

Note The port monitor detects slow devices when a given rising-threshold is reached and triggers the congestion isolation feature in the switch to move traffic to that slow device into the slow Virtual Link (VL2). The switch does not automatically remove any devices from congestion isolation. This must be done manually once the problem with the slow device is identified and resolved.

Guidelines and Limitations for Fabric Notifications

- Fabric Notifications is supported only on Fibre Channel ports.
- Fabric Notifications is supported only on Cisco MDS 9132T, MDS 9148T, MDS 9220i, MDS 9396S, MDS 9396T, MDS 9706, MDS 9710, and MDS 9718 switches.
- Fabric Notifications is not supported on Cisco MDS 9250i and MDS 9148S switches.
- Fabric Notifications is supported on MDS 9706, MDS 9710, and MDS 9718 switches using 48-port 32-Gbps Fibre Channel Switch module.
- In Cisco MDS NX-OS Release 8.5(1), Fabric Notifications is not supported on switches that are operating in the Cisco NPV mode.
- Devices that are configured with FPIN must register with RDF and EDC for using the Fabric Notifications capabilities.
- Fabric Notifications does not monitor devices that are behind vfc interfaces.
- Fabric Notifications supports only Tx of congestion signals and not Rx.
- Fabric Notifications supports following FPIN capabilities:
  - FPIN Link Integrity:
    - Link Failure
    - Loss-of-Synchronization
    - Loss-of-Signal
    - Invalid Transmission Word
    - Invalid CRC
  - FPIN Congestion:
    - Credit Stall
  - FPIN Peer Congestion:
    - Credit Stall
    - Priority Update Notification
- Fabric Notifications do not support following FPIN capabilities:
  - FPIN Link Integrity:

    - Primitive Sequence Protocol Error
  - FPIN Congestion:
    - Oversubscription
    - Lost Credit
  - FPIN Peer Congestion:
    - Oversubscription
    - Lost Credit
  - FPIN Delivery:
    - Timeout
    - Unable to Route

- If the logical type of the port is changed using the **switchport logical-type** command after the device is marked as congested, the device will not be marked as normal automatically. The device needs to be recovered using the **fpm congested-device recover pwwn** *pwwn* **vsan** *id* command.
- For devices that are not registered for FPIN, all the flows destined to slow devices are moved to the low-priority VL. After the slow devices recover from congestion, the flows are moved back to the normal VL.
- Ensure that the portguard action that is configured for slow drain counters is consistent across switches in a fabric.
- Portguard action is initiated from the switch where congestion is detected.
- Port monitor does not take action on devices that part of the excluded list. For more information, see Configuring Excluded List of Congested Devices, on page 69.
- FPIN is not supported on devices that are part of Inter VSAN Routing (IVR) zoneset.
- If you are upgrading to Cisco MDS NX-OS Release 8.5(1) or later release and if the Congestion Isolation feature is enabled, ensure that you disable the Congestion Isolation feature and then enable FPM after upgrading. After upgrading, the port monitor configurations are cleared and it starts detecting events afresh. For enabling the Congestion Isolation feature, see Configuring Congestion Isolation, on page 66.

Guidelines and Limitations for DIRL

- DIRL is supported on Cisco MDS 9132T, MDS 9148T, MDS 9220i, and MDS 9396T switches.
- DIRL is not supported on Cisco MDS 9250i, MDS 9148S, and MDS 9396S switches.
- DIRL is supported on MDS 9706, MDS 9710, and MDS 9718 switches using 48-port 32-Gbps Fibre Channel Switch module.
- DIRL is not supported on Cisco MDS 9700 48-Port 16-Gbps Fibre Channel Switching Module and Cisco MDS 9700 24/10-Port SAN Extension Module.
- DIRL is not supported on switches operating in Cisco NPV mode.

- DIRL is supported only on F ports.
- If you are upgrading to Cisco MDS NX-OS Release 8.5(1) or later release and if you have configured the port ingress rate limiting on one or more interfaces, any static ingress rate limiting must be removed using the **no switchport ingress-rate** prior to upgrading to Cisco MDS NX-OS Release 8.5(1) or later release.

  After upgrading to Cisco MDS NX-OS Release 8.5(1) or later, static ingress rate limiting can once again be configured on any interface, if necessary. However, if static ingress rate limiting is configured for an interface, then this interface will not be subject to DIRL.
- The following table shows the maximum (lowest) ingress rate limits set by DIRL for each link speed.

Table 3: Minimum Ingress Rates by Hardware Type and Operational Speed

Switches/Modules	Operational Link Speed	Maximum (lowest) Ingress Rate Limit
• Cisco MDS 9132T • Cisco MDS 9396T • Cisco MDS 9148T • Cisco MDS 9220i • Cisco MDS 9700 48-Port 32-Gbps Fibre Channel Switching Module	32 Gbps	0.01250% (0.4 Gbps)
	16 Gbps	0.02435% (0.4 Gbps)
	8 Gbps	0.04870% (0.4 Gbps)
	4 Gbps	0.09741% (0.4 Gbps)

Configuring Congestion Management

Configuring Congestion Detection

Most of the features used for congestion detection are enabled by default and do not require any additional configuration. These features include txwait, rxwait, interface priority flow control, OBFL error stats, and tx-credit-not-available. The following congestion detection features are configurable.

Modules and switches included in “Module and Switch Support” section of Table 20.

- 16-Gbps modules or switches:
  - Cisco MDS 9700 Series 16-Gbps Fibre Channel Module (DS-X9448-768K9)
  - Cisco MDS 9000 Series 24/10 SAN Extension Module (DS-X9334-K9)
  - Cisco MDS 9250i Fabric Switch
  - Cisco MDS 9148S Fabric Switch
  - Cisco MDS 9396S Fabric Switch
- 32-Gbps modules or switches:
  - Cisco MDS 9000 Series 32-Gbps Fibre Channel Module (DS-X9648-1536K9)
  - Cisco MDS 9132T Fibre Channel Switch
- 10-Gbps FCoE module:
  - Cisco MDS 9700 48-Port 10-Gbps Fibre Channel over Ethernet (DS-X9848-480K9)
- 40-Gbps FCoE module:
  - Cisco MDS 9700 40-Gbps 24-Port Fibre Channel over Ethernet Module (DS-X9824-960K9)

Table 4: Slow Port Monitor Support on Fibre Channel and FCoE Switching Modules, on page 58 displays the congestion detection features supported on different Fibre Channel and FCoE switching modules for the Cisco MDS NX-OS Release 8.x.

Table 4: Slow Port Monitor Support on Fibre Channel and FCoE Switching Modules

Function	Module and Switch Support	
	16 Gbps and 32 Gbps Fibre Channel	10 Gbps and 40 Gbps FCoE
Txwait OBFL logging	Yes	Yes, from Cisco MDS NX-OS Release 8.2(1) onwards.
Txwait port monitor counter	Yes	No
Txwait interface counter	Yes	Yes, from Cisco MDS NX-OS Release 8.2(1) onwards.

Function	Module and Switch Support	
Txwait interface unable to transmit for the last 1 second, 1 minute, 1 hour, and 72 hours	Yes	Yes, from Cisco MDS NX-OS Release 8.2(1) onwards.
A graphical representation of txwait for the last 60 seconds, 60 minutes, and 72 hours	Yes	Yes, from Cisco MDS NX-OS Release 8.2(1) onwards.
Rxwait OBFL logging	No	Yes, from Cisco MDS NX-OS Release 8.2(1) onwards.
Rxwait interface counter	No	Yes, from Cisco MDS NX-OS Release 8.2(1) onwards.
Rxwait interface unable to receive for the last 1 second, 1 minute, 1 hour, and 72 hours	No	Yes, from Cisco MDS NX-OS Release 8.2(1) onwards.
A graphical representation of rxwait for the last 60 seconds, 60 minutes, and 72 hours	No	Yes, from Cisco MDS NX-OS Release 8.2(1) onwards.
Port monitor slow-port counter	Yes	No
OBFL error stats	Yes	Yes, from Cisco MDS NX-OS Release 8.2(1) onwards.
Interface priority flow control	No	Yes, from Cisco MDS NX-OS Release 8.2(1) onwards.

Configuring the Slow-Port Monitor Timeout Value for Fibre Channel

The slow-port monitor functionality is similar to the no-credit frame timeout and drop functionality, except that it does not drop frames; it only logs qualifying events. When a Fibre Channel egress port has no transmit credits continuously for the slow-port monitor timeout period, the event is logged. No frames are dropped unless the no-credit frame timeout period is reached and no-credit frame timeout drop is enabled. If the no-credit frame timeout drop is not enabled, no frames are dropped until the congestion frame timeout period is reached.

Slow-port monitoring is implemented in the hardware, with the slow-port monitor functionality being slightly different in each generation of hardware. The 16-Gbps and 32-Gbps modules and switches can detect each instance of the slow-port monitor threshold being crossed. The slow-port monitor log is updated at 100-ms intervals. A log for a slow-port event on a 16-Gbps and 32-Gbps module or system increments the exact number of times the threshold is reached.

Slow port monitor can also generate an alert and syslog message via port monitor.

To configure the slow-port monitor timeout value, perform these steps:

Step 1 Enter configuration mode:

switch# **configure terminal**

Step 2 Specify the slow-port monitor timeout value:

switch(config)# **system timeout slowport-monitor** *milliseconds* **logical-type** {**core | edge**}

Valid values for the slow-port monitor timeout are:

- 32-Gbps and 16-Gbps modules or switches—1 to 500 ms in 1-ms increments.

Note For 32-Gbps modules, ISLs (E ports) and trunking F and NP ports (TF and TNP ports) will use the core timeout value and non-trunking F ports (F and NP ports) or edge ports will use the edge timeout value.

(Optional) Revert to the default slow-port monitor timeout value (50 ms) for the specified port type:

switch(config)# **system timeout slowport-monitor default logical-type** {**core | edge**}

(Optional) Disable the slow-port monitor:

switch(config)# **no system timeout slowport-monitor default logical-type** {**core | edge**}

Configuring Slow Port Monitor for Port Monitor

Slow port monitor can be configured in port monitor via the tx-slowport-oper-delay counter. The **system timeout slowport-monitor** command also must be configured with a value that is less than or equal to the tx-slowport-oper-delay rising threshold. The port monitor logical type must also match the **system timeout slowport-monitor logical-type** command. Failure to do this results in no port monitor alerts being generated for tx-slowport-oper-delay.

Configuring the Transmit Average Credit-Not-Available Duration Threshold and Action in Port Monitor

Cisco MDS monitors its ports that are at zero transmit credits for 100 ms or more. This is called transmit average credit-not-available duration. The Port Monitor feature can monitor this using the TX Credit Not Available counter. When the transmit average credit-not-available duration exceeds the threshold set in the port monitor policy, an SNMP trap with interface details is sent, indicating the transmit average credit not available duration event along with a syslog message. Additionally, the following events may be configured:

- A warning message is displayed.
- The port can be error disabled.
- The port can be flapped.

The Port Monitor feature provides the CLI to configure the thresholds and actions. The threshold configuration is configured as a percentage of the interval. The thresholds can be 0 to 100 percent in multiples of 10, and the interval can be 1 second to 1 hour. The default is 10 percent of a 1-second interval and generates a SNMP trap and syslog message when the transmit-average-credit-not-available duration hits 100 ms.

The following edge port monitor policy is active by default. No port monitor policy is enabled for core ports by default.

```
switch# show port-monitor slowdrain

Policy Name  : slowdrain
Admin status : Not Active
Oper status  : Not Active
```

```
Port type    : All Edge Ports
--------------------------------------------------------------------------------------------------------------
|         Counter          | Threshold  | Interval |      Warning       |     Thresholds
   |             Rising/Falling actions              | Congestion-signal |
|                          |    Type    |  (Secs)
|--------------------------|------------|-------------------------------------------|------------------|
|                          |            |          | Threshold | Alerts |  Rising  | Falling
  | Event |     Alerts       |     PortGuard        | Warning | Alarm   |
--------------------------------------------------------------------------------------------------------------
| Credit Loss Reco         | Delta      | 1        | none      | n/a    | 1        | 0
   | 4     | syslog,rmon      | none                 | n/a     | n/a     |
| TX Credit Not Available | Delta      | 1        | none      | n/a    | 10%      | 0%
   | 4     | syslog,rmon      | none                 | n/a     | n/a     |
| TX Datarate              | Delta      | 10       | none      | n/a    | 80%      | 70%
   | 4     | syslog,rmon      | none                 | n/a     | n/a     |
--------------------------------------------------------------------------------------------------------------
```

The following example shows how to configure a new policy similar to the slowdrain policy with the tx-credit not available threshold set to 200 ms:

Note The default *slowdrain* port monitor policy cannot be modified; hence, a new policy needs to be configured.

```
switch# configure
switch(config)# port-monitor name slowdrain_tx200ms
switch(config-port-monitor)# logical-type edge
switch(config-port-monitor)# no monitor counter all
switch(config-port-monitor)# monitor counter credit-loss-reco
switch(config-port-monitor)# monitor counter tx-credit-not-available
switch(config-port-monitor)# counter tx-credit-not-available poll-interval 1 delta
rising-threshold 20 event 4 falling-threshold 0
switch(config-port-monitor)# no port-monitor activate slowdrain
switch(config)# port-monitor activate slowdrain_tx200ms
switch(config)# end

switch# show port-monitor active
Policy Name  : slowdrain_tx200ms
Admin status : Not Active
Oper status  : Not Active
Port type    : All Edge Ports
--------------------------------------------------------------------------------------------------------------
|         Counter          | Threshold  | Interval |      Warning       |     Thresholds
   |             Rising/Falling actions              | Congestion-signal |
|                          |    Type    |  (Secs)
|--------------------------|------------|-------------------------------------------|------------------|
|                          |            |          | Threshold | Alerts |  Rising  | Falling
  | Event |     Alerts       |     PortGuard        | Warning | Alarm   |
--------------------------------------------------------------------------------------------------------------
| Credit Loss Reco         | Delta      | 1        | none      | n/a    | 1        | 0
   | 4     | syslog,rmon      | none                 | n/a     | n/a     |
| TX Credit Not Available | Delta      | 1        | none      | n/a    | 20%      | 0%
   | 4     | syslog,rmon      | none                 | n/a     | n/a     |
--------------------------------------------------------------------------------------------------------------
```

Configuring Other Congestion Related Port Monitor Counters

The following port-monitor counters related to SAN congestion can be configured:

Table 5: Port-Monitor Counters

Counter Name	Description
invalid-words	Represents the total number of invalid words received by a port.
link-loss	Represents the total number of link failures encountered by a port.
lr-rx	Represents the total number link reset primitive sequence received by a port.
lr-tx	Represents the total number of link reset primitive sequence transmitted by a port.
rx-datarate	Receives frame rate in bytes per seconds.
signal-loss	Represents the number of times a port encountered laser or signal loss.
state-change	Represents the number of times a port has transitioned to an operational up state.
sync-loss	Represents the number of times a port experienced loss of synchronization in RX.
tx-credit-not-available	Increments by one if there is no transmit buffer-to-buffer credits available for a duration of 100 ms.
timeout-discards	Represents the total number of frames dropped at egress due to congestion timeout or no-credit-drop timeout.
tx-datarate	Represents the transmit frame rate in bytes per seconds.
tx-discards	Represents the total number of frames dropped at egress due to timeout, abort, offline, and so on.
tx-slowport-count	Represents the number of times slow port events were detected by a port for the configured slowport-monitor timeout. This is applicable only for generation 3 modules.
tx-slowport-oper-delay	Captures average credit delay (or R_RDY delay) experienced by a port. The value is in milliseconds.

Configuring Congestion Avoidance

The following features can be configured for congestion avoidance:

- Congestion-drop
- No-credit-drop
- Pause-drop
- Port-monitor portguard action for congestion avoidance

Configuring the Congestion Drop Timeout Value for FCoE

When an FCoE frame takes longer than the congestion drop timeout period to be transmitted by the egress port, the frame is dropped. This dropping of frames is useful in controlling the effect of slow egress ports that are paused almost continuously (long enough to cause congestion), but not long enough to trigger the pause timeout drop. Frames dropped due to the congestion drop threshold are counted as egress discards against the egress port. Egress discards release buffers in the upstream ingress ports of a switch, allowing the unrelated flows to move continuously through the ports.

The congestion drop timeout value is 500 ms by default for all port types. We recommend that you retain the default timeout for core ports, and consider configuring a lower value for edge ports. The congestion drop timeout value should be equal to or greater than the pause drop timeout value for that port type.

To configure the congestion drop timeout value for FCoE, perform these steps:

Step 1 Enter configuration mode:

switch# **configure terminal**

Step 2 Depending on the Cisco MDS NX-OS release version you are using, use one of the following commands to configure the system-wide FCoE congestion drop timeout, in milliseconds, for core or edge ports:

- Cisco MDS NX-OS Release 8.1(1) and earlier releases

  switch(config)# **system default interface congestion timeout** *milliseconds* **mode {core | edge}**

  The FCoE congestion drop timeout range is from 100 to 1000 ms.

  Note To prevent premature packet drops, the minimum value recommended for FCoE congestion drop timeout is 200 ms.

- Cisco MDS NX-OS Release 8.2(1) and later releases

  switch(config)# **system timeout fcoe congestion-drop** {*milliseconds* | **default**} **mode {core | edge}**

  The FCoE congestion drop timeout range is from 200 to 500 ms.

  Note In Cisco MDS NX-OS Release 8.1(1) and earlier releases, the FCoE congestion drop timeout value could be configured to as low as 100 ms. However, under certain circumstances configuring a congestion drop timeout value of 100 ms led to premature packet drops. In Cisco MDS NX-OS 8.2(1) and later releases, the minimum congestion drop timeout value was set to 200 ms to prevent premature packet drops. Therefore, we do not recommended that you specify a congestion drop timeout value of less than 200 ms in Cisco MDS NX-OS Release 8.1(1) and earlier releases.

(Optional) Depending on the Cisco MDS NX-OS release version you are using, use one of the following commands to revert to the default FCoE congestion drop timeout value of 500 milliseconds:

- Cisco MDS NX-OS Release 8.1(1) and earlier releases

  switch(config)# **no system default interface congestion timeout** *milliseconds* **mode {core | edge}**

- Cisco MDS NX-OS Release 8.2(1) and later releases

  switch(config)# **no system timeout fcoe congestion-drop** {*milliseconds* | **default**} **mode {core | edge}**

Configuring Pause Drop Timeout for FCoE

When an FCoE port is in a state of continuous pause during the FCoE pause drop timeout period, all the frames that are queued to that port are dropped immediately. As long as the port continues to remain in the pause state, the newly arriving frames destined for the port are dropped immediately. These drops are counted as egress discards on the egress port, and free up buffers in the upstream ingress ports of the switch, allowing unrelated flows to continue moving through them.

To reduce the effect of a slow-drain device on unrelated traffic flows, configure a lower-pause drop timeout value than the congestion frame timeout value, for edge ports. This causes the frames that are destined for a slow port to be dropped immediately after the FCoE pause drop timeout period has occurred, rather than waiting for the congestion timeout period to drop them.

By default, the FCoE pause drop timeout is enabled on all ports and the value is set to 500 ms. We recommend that you retain the default timeout core ports and consider configuring a lower value for edge ports.

To configure the FCoE pause drop timeout value, perform these steps:

Step 1 Enter configuration mode:

switch# **configure terminal**

Step 2 Depending on the Cisco MDS NX-OS release version that you are using, use one of the following commands to configure the system-wide FCoE pause drop timeout value, in milliseconds, for edge or core ports:

- Cisco MDS NX-OS Release 8.1(1) and earlier releases

  switch(config)# **system default interface pause timeout** *milliseconds* **mode {core | edge}**

- Cisco MDS NX-OS Release 8.2(1) and later releases

  switch(config)# **system timeout fcoe pause-drop** {*milliseconds* | **default**} **mode {core | edge}**

The range is from 100 to 500 milliseconds.

(Optional) Depending on the Cisco MDS NX-OS release version that you are using, use one of the following commands to enable the FCoE pause drop timeout to the default value of 500 milliseconds for edge or core ports:

- Cisco MDS NX-OS Release 8.1(1) and earlier releases

  switch(config)# **system default interface pause mode {core | edge}**

- Cisco MDS NX-OS Release 8.2(1) and later releases

  switch(config)# **system timeout fcoe pause-drop default mode {core | edge}**

(Optional) Depending on the Cisco MDS NX-OS release version that you are using, use one of the following commands to disable the FCoE pause drop timeout for edge or core ports:

- Cisco MDS NX-OS Release 8.1(1) and earlier releases

  switch(config)# **no system default interface pause mode {core | edge}**

- Cisco MDS NX-OS Release 8.2(1) and later releases

  switch(config)# **no system timeout fcoe pause-drop default mode {core | edge}**

Configuring the Congestion Drop Timeout Value for Fibre Channel

When a Fibre Channel frame takes longer than the congestion timeout period to be transmitted by the egress port, the frame is dropped. This option of the frames being dropped is useful for controlling the effect of slow egress ports that lack transmit credits almost continuously; long enough to cause congestion, but not long enough to trigger the no-credit timeout drop. These drops are counted as egress discards on the egress port, and release buffers into the upstream ingress ports of the switch, allowing unrelated flows to continue moving through them.

By default, the congestion timeout value is 500 ms for all port types. We recommend that you retain the default timeout for core ports and configure a lower value (not less than 200 ms) for edge ports. The congestion timeout value should be equal to or greater than the no-credit frame timeout value for that port type.

To configure the congestion frame timeout value for the Fibre Channel, perform these steps:

Step 1 Enter configuration mode:

switch# **configure terminal**

Step 2 Configure the Fibre Channel congestion drop timeout value, in milliseconds, for the specified port type:

switch(config)# **system timeout congestion-drop** *milliseconds* **logical-type** {**core | edge**}

The range is 200-500 ms in multiples of 10.

Step 3 (Optional) Revert to the default value for the congestion timeout for the specified port type:

switch(config)# **no system timeout congestion-drop default logical-type** {**core | edge**}

Configuring the No-Credit-Drop Frame Timeout Value for Fibre Channel

When a Fibre Channel egress port has no transmit credits continuously for the no-credit timeout period, all the frames that are already queued to that port are dropped immediately. As long as the port remains in this condition, newly arriving frames destined for that port are dropped immediately. These drops are counted as egress discards on the egress port, and release buffers in the upstream ingress ports of the switch, allowing unrelated flows to continue moving through them.

No-credit dropping can be enabled or disabled. By default, frame dropping is disabled and the frame timeout value is 500 ms for all port types. We recommend that you retain the default frame timeout for core ports and configure a lower value (300 ms) for edge ports. If the slow-drain events continue to affect unrelated traffic flows, the frame timeout value for the edge ports can be lowered to drop the previous slow-drain frames. This

frees the ingress buffers for frames of unrelated flows, thus reducing the latency of the frames through the switch.

Note

- The no-credit frame timeout value should always be less than the congestion frame timeout for the same port type, and the edge port frame timeout values should always be lower than the core port frame timeout values.
- The slow-port monitor delay value should always be less than the no-credit frame timeout value for the same port type.

On 16-Gbps and later modules and systems, the no-credit timeout value can be 1 to 500 ms in multiples of 1 ms. Dropping starts immediately after the no-credit condition comes into existence for the configured timeout value.

To configure the no-credit timeout value, perform these steps:

Step 1 Enter configuration mode:

switch# **configure terminal**

Step 2 Specify the no-credit timeout value:

switch(config)# **system timeout no-credit-drop** *milliseconds* **logical-type edge**

(Optional) Revert to the default no-credit timeout value (500 ms):

switch(config)# **system timeout no-credit-drop default logical-type edge**

(Optional) Disable the no-credit drop timeout value:

switch(config)# **no system timeout no-credit-drop logical-type edge**

Configuring Congestion Isolation

The Congestion Isolation feature allows slow devices to be put into their own virtual link automatically as the port monitor detects the slow-drain condition.

The following port-monitor counters are used to detect slow drain and isolate the devices on an interface.

- credit-loss-reco
- tx-credit-not-available
- tx-slowport-oper-delay
- txwait

Configure the Slow-Drain Device Detection and Congestion Isolation feature in the following sequence:

1. Configure the Extended Receiver Ready feature. For more information, see Enabling Extended Receiver Ready, on page 67.

2. Configure the Congestion Isolation feature. For more information, see Configuring Congestion Isolation, on page 68.
3. Configure a port-monitor policy with one or more counters containing the portguard action *cong-isolate*. For more information, see Configuring Congestion Isolation.

Configuring Extended Receiver Ready

Enabling Extended Receiver Ready

To enable Extended Receiver Ready (ER_RDY) on a switch, perform these steps:

Before you begin

You must enable ER_RDY flow-control mode using the **system fc flow-control er_rdy** command on the local and adjacent switches, and then flap the ISLs connecting the local and adjacent switches to enable ER_RDY flow-control mode on the ISLs.

Step 1 Enter configuration mode:

switch# **configure terminal**

Step 2 Enable ER_RDY flow-control mode:

switch(config)# **system fc flow-control er_rdy**

Note Enable the ER_RDY flow-control mode on both the connected switches for an existing Inter-Switch Link (ISL) before proceeding to step 3.

Step 3 Select a Fibre Channel interface and enter interface configuration submode:

switch(config-if)# **interface fc** *slot/port*

Step 4 Gracefully shut down the interface and administratively disable traffic flow:

switch(config-if)# **shutdown**

Step 5 Enable traffic flow on the interface:

switch(config-if)# **no shutdown**

Step 6 Return to privileged executive mode:

switch(config-if)# **end**

Step 7 Verify if the link is in ER_RDY flow-control mode:

switch# **show flow-control er_rdy**

Disabling Extended Receiver Ready

To disable Extended Receiver Ready (ER_RDY) on a switch, perform these steps:

Before you begin

1. Remove the congestion-isolation portguard action for the links in the port-monitor policy. For more information, see Configuring Congestion Isolation.
2. Disable the Congestion Isolation feature. For more information, see Configuring Congestion Isolation, on page 68.

Step 1 Enter configuration mode:

switch# **configure terminal**

Step 2 Disable ER_RDY flow-control mode:

switch(config)# **no system fc flow-control**

Step 3 Select a Fibre Channel interface and enter interface configuration submode:

switch(config-if)# **interface fc** *slot/port*

Step 4 Gracefully shut down the interface and administratively disable traffic flow:

switch(config-if)# **shutdown**

Step 5 Enable traffic flow on the interface:

switch(config-if)# **no shutdown**

Step 6 Return to privileged executive mode:

switch(config-if)# **end**

Step 7 Verify if the link is in R_RDY flow-control mode:

switch# **show flow-control r_rdy**

Configuring Congestion Isolation

To configure Congestion Isolation, perform these steps:

Before you begin

Configure Extended Receiver Ready. For more information, see Enabling Extended Receiver Ready, on page 67.

Step 1 Enter configuration mode:

switch# **configure terminal**

Step 2 Enable Congestion Isolation:

Prior to Cisco MDS NX-OS Release 8.5(1)

switch(config)# **feature congestion-isolation**

Cisco MDS NX-OS Release 8.5(1) and later releases

switch(config)# **feature fpm**

Step 3 Specify the counter parameters for the portguard to take Congestion Isolation action on a port:

Prior to Cisco MDS NX-OS Release 8.5(1)

switch(config-port-monitor)# **counter** {**credit-loss-reco** | **tx-credit-not-available** | **tx-slowport-oper-delay** | **txwait**} **poll-interval** *seconds* {**absolute** | **delta**} **rising-threshold** *count1* **event** *event-id* **warning-threshold** *count2* **falling-threshold** *count3* **event** *event-id* **portguard cong-isolate**

switch(config-port-monitor)# **exit**

Cisco MDS NX-OS Release 8.5(1) and later releases

switch(config-port-monitor)# **counter** {**credit-loss-reco** | **tx-credit-not-available** | **tx-slowport-oper-delay** | **txwait**} **poll-interval** *seconds* {**absolute** | **delta**} **rising-threshold** *count1* **event** *event-id* **warning-threshold** *count2* **falling-threshold** *count3* **portguard cong-isolate**

switch(config-port-monitor)# **exit**

Note Absolute counters do not support portguard actions. However, the tx-slowport-oper-delay absolute counter supports Congestion Isolation portguard action.

Step 4 Activate the specified port-monitor policy:

switch(config)# **port-monitor activate** *policyname*

From Cisco MDS NX-OS Release 8.5(1)

Configuring Excluded List of Congested Devices

To explicitly exclude a device from congestion actions, perform these steps:

Before you begin

Enable FPM. For more information, see Enabling FPM, on page 73.

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Enter congested device exclude mode:

switch(config)# **fpm congested-device exclude list**

Step 3 Exclude a device from congestion actions:

switch(config-congested-dev-exc)# **member pwwn** *pwwn* **vsan** *id*

Configuring Static List of Congested Devices

To explicitly configure a device as congested, perform these steps:

Before you begin

Enable FPM. For more information, see Enabling FPM, on page 73.

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Enter congested device static mode:

switch(config)# **fpm congested-device static list**

Step 3 Configure a device as congested:

switch(config-congested-dev-static)# **member pwwn** *pwwn* **vsan** *id* **credit-stall**

Recovering a Congested Device

Use this procedure to recover congested device that was detected by port monitor.

To recover a device from congestion, perform this step:

Recover a device from congestion:

switch# **fpm congested-device recover pwwn** *pwwn* **vsan** *id*

Prior to Cisco MDS NX-OS Release 8.5(1)

Including or Excluding a Congested Device

To explicitly include a device as congested such that it is identified as a congested device by the port monitor, or exclude a device that is identified as a congested device by the port monitor, perform these steps:

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Explicitly include a device as congested or exclude a device from being detected as congested:

switch# **congestion-isolation {exclude | include} pwwn** *pwwn* **vsan** *vsan-id*

Removing an Interface

Port monitor detects slow devices when a given threshold is reached and triggers the congestion isolation feature in the switch to move traffic to that slow device into the slow Virtual Link (VL2). The switch does not automatically remove any devices from congestion isolation. This must be done manually once the problem with the slow device is identified and resolved.

To manually remove an interface from being detected as slow, perform these steps:

Remove an interface from being detected as slow by the port monitor:

switch#: **congestion-isolation remove interface** *slot/port*

Configuring Congestion Isolation Recovery

To configure the Congestion Isolation Recovery feature, perform these steps:

Before you begin

Enable Extended Receiver Ready. For more information, see Enabling Extended Receiver Ready, on page 67.

Step 1 Enter configuration mode:

switch# **configure terminal**

Step 2 Enable FPM:

switch(config)# **feature fpm**

Step 3 Specify the policy name and enter port monitoring policy configuration mode:

switch(config)# **port-monitor name** *policyname*

Step 4 Specify the counter parameters for the portguard to take Congestion Isolation Recovery action on a port:

switch(config-port-monitor)# **counter** {**credit-loss-reco** | **tx-credit-not-available** | **tx-slowport-oper-delay** | **txwait**} **poll-interval** *seconds* {**absolute** | **delta**} **rising-threshold** *count1* **event** *event-id* **warning-threshold** *count2* **falling-threshold** *count3* **event** *event-id* **portguard cong-isolate-recover**

Note Absolute counters do not support portguard actions. However, the tx-slowport-oper-delay absolute counter supports Congestion Isolation Recovery portguard actions.

Step 5 Return to configuration mode:

switch(config-port-monitor)# **exit**

Step 6 (Optional) Change recovery-interval:

switch(config)# **port-monitor cong-isolation-recover recovery-interval** *seconds*

Step 7 (Optional) Specify isolate-duration:

switch(config)# **port-monitor cong-isolation-recover isolate-duration** *hours* **num-occurrence** *number*

Step 8 Activate the specified port-monitor policy:

switch(config)# **port-monitor activate** *policyname*

Step 9 (Optional) You can manually exclude a device to be detected as a slow device.

See Configuring Excluded List of Congested Devices, on page 69.

Configuring Static List of Congested Devices

To explicitly configure a device as congested, perform these steps:

Before you begin

Enable FPM. For more information, see Enabling FPM, on page 73.

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Enter congested device static mode:

switch(config)# **fpm congested-device static list**

Step 3 Configure a device as congested:

switch(config-congested-dev-static)# **member pwwn** *pwwn* **vsan** *id* **credit-stall**

Configuring Excluded List of Congested Devices

To explicitly exclude a device from congestion actions, perform these steps:

Before you begin

Enable FPM. For more information, see Enabling FPM, on page 73.

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Enter congested device exclude mode:

switch(config)# **fpm congested-device exclude list**

Step 3 Exclude a device from congestion actions:

switch(config-congested-dev-exc)# **member pwwn** *pwwn* **vsan** *id*

Recovering a Congested Device

Use this procedure to recover congested device that was detected by port monitor.

To recover a device from congestion, perform this step:

Recover a device from congestion:

switch# **fpm congested-device recover pwwn** *pwwn* **vsan** *id*

Configuring Fabric Notifications

Enabling FPM

To enable FPM, perform these steps:

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Enable FPM:

switch# **feature fpm**

Disabling FPM

To disable FPM, perform these steps:

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Disable FPM:

switch# **no feature fpm**

Configuring the Port-Monitor Portguard Action for FPIN

To configure the port-monitor portguard action for FPIN, perform these steps:

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Enable FPM:

switch(config)# **feature fpm**

Step 3 Specify the policy name and enter port monitoring policy configuration mode:

switch(config)# **port-monitor name** *policyname*

Step 4 Specify the counter parameters for the portguard for FPIN:

switch(config-port-monitor)# **counter** {**invalid-crc** | **invalid-words** | **link-loss** | **signal-loss** | **sync-loss** | **txwait**} **poll-interval** *seconds* {**absolute** | **delta**} **rising-threshold** *count1* **event** *event-id* **warning-threshold** *count2* **falling-threshold** *count3* **portguard FPIN**

Step 5 Return to configuration mode:

switch(config-port-monitor)# **exit**

Step 6 Activate the specified port-monitor policy:

switch(config)# **port-monitor activate** *policyname*

Step 7 (Optional) Specify the recovery interval. By default, the recovery interval is set to 900 seconds (15 minutes).

switch(config)# **port-monitor fpin recovery-interval** *seconds*

Step 8 (Optional) Specify the isolate duration:

switch(config)# **port-monitor fpin isolate-duration** *hours* **num-occurrence** *number*

Configuring Static List of Congested Devices

To explicitly configure a device as congested, perform these steps:

Before you begin

Enable FPM. For more information, see Enabling FPM, on page 73.

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Enter congested device static mode:

switch(config)# **fpm congested-device static list**

Step 3 Configure a device as congested:

switch(config-congested-dev-static)# **member pwwn** *pwwn* **vsan** *id* **credit-stall**

Configuring Excluded List of Congested Devices

To explicitly exclude a device from congestion actions, perform these steps:

Before you begin

Enable FPM. For more information, see Enabling FPM, on page 73.

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Enter congested device exclude mode:

switch(config)# **fpm congested-device exclude list**

Step 3 Exclude a device from congestion actions:

switch(config-congested-dev-exc)# **member pwwn** *pwwn* **vsan** *id*

Recovering a Congested Device

Use this procedure to recover congested device that was detected by port monitor.

To recover a device from congestion, perform this step:

Recover a device from congestion:

switch# **fpm congested-device recover pwwn** *pwwn* **vsan** *id*

Configuring FPIN Notification Interval

To change the default FPIN notification interval, perform these steps:

Before you begin

Enable FPM. For more information, see Enabling FPM, on page 73.

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Change the FPIN notification interval:

switch(config)# **fpm fpin period** *seconds*

Configuring EDC Congestion Signal

To configure the EDC interval for sending congestion signal, perform these steps:

Before you begin

Enable FPM. For more information, see Enabling FPM, on page 73.

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Specify the policy name and enter port monitoring policy configuration mode:

switch(config)# **port-monitor name** *policyname*

Step 3 Specify the counter parameters for congestion signals:

switch(config-port-monitor)# **counter txwait warning-signal-threshold** *count1* **alarm-signal-threshold** *count2* **portguard congestion-signals**

Step 4 (Optional) Exit the port monitor configuration mode:

switch(config-port-monitor)# **exit**

Step 5 (Optional) Specify the EDC switch-side period for sending congestion signal. By default, the switch-side congestion signal period is set to 1 second.

switch(config)# **fpm congestion-signal period** *seconds*

Configuring DIRL

Before You Begin

Enable FPM. For more information, see Enabling FPM, on page 73.

Configuring the Port-Monitor Portguard Action for DIRL

To configure the port-monitor portguard action for DIRL, perform these steps:

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Enable FPM:

switch(config)# **feature fpm**

Step 3 Specify the policy name and enter port monitoring policy configuration mode:

switch(config)# **port-monitor name** *policyname*

Step 4 Specify the counter parameters for the portguard for DIRL:

switch(config-port-monitor)# **counter** {**tx-datarate** | **tx-datarate-burst** | **txwait**} **poll-interval** *seconds* {**absolute** | **delta**} **rising-threshold** *count1* **event** *event-id* **warning-threshold** *count2* **falling-threshold** *count3* **portguard DIRL**

Step 5 Return to configuration mode:

switch(config-port-monitor)# **exit**

Step 6 Activate the specified port-monitor policy:

switch(config)# **port-monitor activate** *policyname*

Step 7 (Optional) Specify the recovery interval. By default, the recovery interval is set to 60 seconds .

switch(config)# **port-monitor dirl recovery-interval** *seconds*

Configuring DIRL Rate Reduction and Recovery Percentages

To configure the DIRL rate reduction percentages, perform these steps:

Step 1 Enter configuration mode:

switch# **configure**

Step 2 (Optional) Specify the ingress rate reduction and recovery percentages:

switch(config)# **fpm dirl reduction** *percentage* **recovery** *percentage*

What to do next

To configure ingress port rate limit, see Configuring Static Ingress Port Rate Limiting, on page 78.

Excluding Interfaces from DIRL Rate Reduction

To exclude an interface from DIRL rate reduction, perform these steps:

Note Interfaces having devices with FC4-feature as *init* are monitored by default. If other interfaces need to be monitored, use the **no member fc4-feature target** command.

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Enter DIRL exclude list mode:

switch(config)# **fpm dirl exclude list**

Step 3 Specify an interface:

switch(config-dirl-excl)# **member interface fc** *slot/port*

Step 4 Specify the interface to be excluded from DIRL rate reduction:

switch(config-dirl-excl)# **member** {**fc4-feature target** | **interface fc** *slot/port*}

Recovering Interfaces from DIRL Rate Reduction

To recover interface from DIRL rate reduction, perform these steps:

Recover interface from DIRL rate reduction:

switch# **fpm dirl recover interface fc** *slot/port*

Configuring Static Ingress Port Rate Limiting

To configure the static port rate limiting value, follow these steps

Before you begin

From Cisco MDS NX-OS Release 8.5(1), you need to enable FPM before configuring the port rate limiting value. For more information, see Enabling FPM, on page 73.

Step 1 Enter configuration mode:

switch# **configure**

Step 2 Select the interface to specify the static ingress port rate limit:

switch(config)# **interface fc** *slot/port*

Step 3 Configure the static port rate limit for the selected interface:

switch(config-if)# **switchport ingress-rate** *limit*

Configuration Examples for Congestion Management

Configuration Examples for Congestion Detection

This example shows how to configure the FCoE congestion drop timeout to the default value of 500 milliseconds for a core port type in Cisco MDS NX-OS Release 8.1(1) and earlier releases:

```
switch# configure terminal
switch(config)# system default interface congestion timeout 500 mode core
```

This example shows how to configure the FCoE congestion drop timeout to the default value of 500 milliseconds for a core port type in Cisco MDS NX-OS Release 8.2(1) and later releases:

```
switch# configure terminal
switch(config)# system timeout fcoe congestion-drop default mode core
```

This example shows how to configure the FCoE congestion drop timeout to the default value of 500 milliseconds for an edge port type in Cisco MDS NX-OS Release 8.1(1) and earlier releases:

```
switch# configure terminal
switch(config)# system default interface congestion timeout 500 mode edge
```

This example shows how to configure the FCoE congestion drop timeout to the default value of 500 milliseconds for an edge port type in Cisco MDS NX-OS Release 8.2(1) and later releases:

```
switch# configure terminal
switch(config)# system timeout fcoe congestion-drop default mode edge
```

This example shows how to configure the FCoE congestion drop timeout to the value of 200 milliseconds for a core port type in Cisco MDS NX-OS Release 8.1(1) and earlier releases:

```
switch# configure terminal
switch(config)# system default interface congestion timeout 200 mode core
```

This example shows how to configure the FCoE congestion drop timeout to the value of 200 milliseconds for a core port type in Cisco MDS NX-OS Release 8.2(1) and later releases:

```
switch# configure terminal
switch(config)# system timeout fcoe congestion-drop 200 mode core
```

This example shows how to configure the FCoE congestion drop timeout to the value of 200 milliseconds for an edge port type in Cisco MDS NX-OS Release 8.1(1) and earlier releases:

```
switch# configure terminal
switch(config)# system default interface congestion timeout 200 mode edge
```

This example shows how to configure the FCoE congestion drop timeout to the value of 200 milliseconds for an edge port type in Cisco MDS NX-OS Release 8.2(1) and later releases:

```
switch# configure terminal
switch(config)# system timeout fcoe congestion-drop 200 mode edge
```

This example shows how to configure the FCoE pause drop timeout value of 100 milliseconds for a core port type in Cisco MDS NX-OS Release 8.1(1) and earlier releases:

```
switch# configure terminal
switch(config)# system default interface pause timeout 100 mode core
```

This example shows how to configure the FCoE pause drop timeout value of 200 milliseconds for a core port type in Cisco MDS NX-OS Release 8.2(1) and later releases:

```
switch# configure terminal
switch(config)# system timeout fcoe pause-drop 200 mode core
```

This example shows how to configure the FCoE pause drop timeout value of 100 milliseconds for an edge port type in Cisco MDS NX-OS Release 8.1(1) and earlier releases:

```
switch# configure terminal
switch(config)# system default interface pause timeout 100 mode edge
```

This example shows how to configure the FCoE pause drop timeout value of 200 milliseconds for a edge port type in Cisco MDS NX-OS Release 8.2(1) and later releases:

```
switch# configure terminal
switch(config)# system timeout fcoe pause-drop 200 mode edge
```

This example shows how to configure the FCoE pause drop timeout to the default of 500 milliseconds for the core port type in Cisco MDS NX-OS Release 8.1(1) and earlier releases:

```
switch# configure terminal
switch(config)# system default interface pause mode core
```

This example shows how to configure the FCoE pause drop timeout to the default of 500 milliseconds for the core port type in Cisco MDS NX-OS Release 8.2(1) and later releases:

```
switch# configure terminal
switch(config)# system timeout fcoe pause-drop default mode core
```

This example shows how to configure the FCoE pause drop timeout to the default of 500 milliseconds for the edge port type in Cisco MDS NX-OS Release 8.1(1) and earlier releases:

```
switch# configure terminal
switch(config)# system default interface pause mode edge
```

This example shows how to configure the FCoE pause drop timeout to the default value of 500 milliseconds for an edge port type in Cisco MDS NX-OS Release 8.2(1) and later releases:

```
switch# configure terminal
switch(config)# system timeout fcoe pause-drop default mode edge
```

This example shows how to disable the FCoE pause drop timeout for a core port type in Cisco MDS NX-OS Release 8.1(1) and earlier releases:

```
switch# configure terminal
switch(config)# no system default interface pause mode core
```

This example shows how to disable the FCoE pause drop timeout for a core port type in Cisco MDS NX-OS Release 8.2(1) and later releases:

```
switch# configure terminal
switch(config)# no system timeout fcoe pause-drop default mode core
```

This example shows how to disable the FCoE pause drop timeout for an edge port type in Cisco MDS NX-OS Release 8.1(1) and earlier releases:

```
switch# configure terminal
switch(config)# no system default interface pause mode edge
```

This example shows how to disable the FCoE pause drop timeout for an edge port type in Cisco MDS NX-OS Release 8.2(1) and later releases:

```
switch# configure terminal
switch(config)# no system timeout fcoe pause-drop default mode edge
```

Configuration Examples for Congestion Avoidance

Note

- From Cisco MDS NX-OS Release 8.1(1), mode E is treated as logical-type core and mode F is treated as logical-type edge.
- The port *Logical type* is displayed as the *Port type*.

This example shows hot to check the currently active port-monitor policy:

```
switch# show port-monitor active
Policy Name  : sample
Admin status : Active
Oper status  : Active
Port type    : All Ports
------------------------------------------------------------------------------------------
Counter     Threshold  Interval Rising     event Falling   event Warning      PMON
                                Threshold        Threshold       Threshold    Portguard
------------------------------------------------------------------------------------------
```

```
Link
Loss          Delta      10        6          4     5          4   Not enabled  Flap
Sync
Loss          Delta      60        5          4     1          4   Not enabled  Not enabled
Signal
Loss          Delta      60        5          4     1          4   Not enabled   Not enabled
Invalid
Words         Delta      60        1          4     0          4   Not enabled   Not enabled
Invalid
CRC's         Delta      30        20         2     10         2   Not enabled   Not enabled
State
Change        Delta      60        5          4     0          4   Not enabled   Not enabled
TX
Discards      Delta      60        200        4     10         4   Not enabled   Not enabled
LR RX         Delta      60        5          4     1          4   Not enabled   Not enabled
LR TX         Delta      60        5          4     1          4   Not enabled   Not enabled
Timeout
Discards      Delta      60        200        4     10         4   Not enabled   Not enabled
Credit
Loss Reco     Delta      1         1          4     0          4   Not enabled   Not enabled
TX Credit
Not Available Delta      3         40%        4     2%         4   Not enabled   Not enabled
RX Datarate   Delta      60        80%        4     20%        4   Not enabled   Not enabled
TX Datarate   Delta      60        80%        4     20%        4   Not enabled   Not enabled
ASIC Error
Pkt to xbar   Delta      300       5          4     0          4   Not enabled   Not enabled
-------------------------------------------------------------------------------------------
```

This example shows how to configure the Fibre Channel congestion drop timeout value of 210 milliseconds for logical type core:

```
switch# configure terminal
switch(config)# system timeout congestion-drop 210 logical-type core
```

This example shows how to configure the Fibre Channel congestion drop timeout to the default value of 200 milliseconds for logical type core:

```
switch# configure terminal
switch(config)# system timeout congestion-drop default logical-type core
```

This example shows how to configure the Fibre Channel no-credit drop timeout value of 100 milliseconds for logical type edge:

```
switch# configure terminal
switch(config)# system timeout no-credit-drop 100 logical-type edge
```

This example shows how to configure the Fibre Channel no-credit drop timeout to the default value of 500 milliseconds for logical type edge:

Note The no-credit drop timeout value is disabled by default.

```
switch# configure terminal
switch(config)# system timeout no-credit-drop default logical-type edge
```

This example shows how to disable the Fibre Channel no-credit drop timeout for logical type edge when it is enabled:

```
switch# configure terminal
switch(config)# no system timeout no-credit-drop logical-type edge
```

This example shows how to configure the Fibre Channel hardware slowport monitoring value of 10 milliseconds for logical type edge:

```
switch# configure terminal
switch(config)# system timeout slowport-monitor 10 logical-type edge
```

This example shows how to configure the Fibre Channel hardware slowport monitoring to the default value of 50 milliseconds for logical type edge:

Note The slowport monitoring value is disabled by default.

```
switch# configure terminal
switch(config)# system timeout slowport-monitor default logical-type edge
```

This example shows how to disable the Fibre Channel hardware slowport monitoring for logical type edge when it is enabled:

```
switch# configure terminal
switch(config)# no system timeout slowport-monitor logical-type edge
```

Configuration Examples for Congestion Isolation

This example shows how to enable ER_RDY flow-control mode:

```
switch# configure terminal
switch(config)# system fc flow-control er_rdy
Flap the ISLs to activate ER_RDY mode on E ports.Use the CLI show flow-control r_rdy to
list the ports that are still in R_RDY mode
```

This example shows how to disable ER_RDY flow-control mode:

Note You need to disable the Congestion Isolation feature before disabling the ER_RDY flow-control mode.

```
switch# configure terminal
switch(config)# no feature congestion-isolation
switch(config)# no system fc flow-control
```

This example shows how to enable Congestion Isolation for releases prior to Cisco MDS NX-OS Release 8.5(1):

```
switch# configure terminal
switch(config)# feature congestion-isolation
Flap the ISLs to activate ER_RDY mode on E ports. Use the CLI show flow-control r_rdy to
list the ports that are still in R_RDY mode
```

This example shows how to enable Congestion Isolation in Cisco MDS NX-OS Release 8.5(1) and later release:

```
switch# configure terminal
switch(config)# feature fpm
```

This example shows how to disable Congestion Isolation for releases prior to Cisco MDS NX-OS Release 8.5(1):

```
switch# configure terminal
switch(config)# no feature congestion-isolation
Flap the ISLs to activate ER_RDY mode on E ports.Use the CLI show flow-control r_rdy to
list the ports that are still in R_RDY mode
```

This example shows how to disable Congestion Isolation in Cisco MDS NX-OS Release 8.5(1) and later releases:

```
switch# configure terminal
switch(config)# no feature fpm
```

This example shows how to manually configure a device as a congested device for releases prior to Cisco MDS NX-OS Release 8.5(1). The configured device will be permanently treated as a congested device until it is removed from congestion isolation. All traffic to this device traversing the device's ISLs that are in ER_RDY flow-control mode will be routed to the low-priority VL (VL2).

```
switch# configure terminal
switch(config)# congestion-isolation include pwwn 10:00:00:00:c9:f9:16:8d vsan 4
```

This example shows how to manually configure a device as a congested device in Cisco MDS NX-OS Release 8.5(1) and later releases. The configured device will be permanently treated as a congested device until it is removed from congestion isolation. All traffic to this device traversing the device's ISLs that are in ER_RDY flow-control mode will be routed to the low-priority VL (VL2).

```
switch# configure terminal
switch(config)# fpm congested-device static list
switch(config-congested-dev-static)# member pwwn 10:00:00:00:c9:f9:16:8d vsan 4 credit-stall
```

This example shows how to configure a device that is to be excluded from automatic congestion isolation by the port monitor for releases prior to Cisco MDS NX-OS Release 8.5(1). Even when the rising threshold of a port-monitor counter is reached and the portguard action is set to cong-isolate, this device will not be isolated as a congested device, and traffic to this device traversing the device's ISLs that are in ER_RDY flow-control mode will not be routed to the low-priority VL (VL2).

```
switch# configure terminal
```

```
switch(config)# congestion-isolation exclude pwwn 10:00:00:00:c9:f9:16:8d vsan 4
```

This example shows how to configure a device that is to be excluded from automatic congestion isolation by the port monitor in Cisco MDS NX-OS Release 8.5(1) and later releases. Even when the rising threshold of a port-monitor counter is reached and the portguard action is set to cong-isolate, this device will not be isolated as a congested device, and traffic to this device traversing the device's ISLs that are in ER_RDY flow-control mode will not be routed to the low-priority VL (VL2).

```
switch# configure terminal
switch(config)# fpm congested-device exclude list
switch(config-congested-dev-exc)# member pwwn 10:00:00:00:c9:f9:16:8d vsan 4
```

Congested devices can be identified either via the port monitor or manually included or excluded. On removing the exclude configuration, if the device is detected as slow by the port monitor, the device will again be marked as slow. Also, if the exclude configuration is already used for a device marked as slow by the port monitor, the device will no longer behave as a congested device.

This example shows how to manually remove an interface from being detected as slow in the port monitor:

Prior to Cisco MDS NX-OS Release 8.5(1)

1. Identifying the interface that you want to remove from being detected as slow.

```
switch# show congestion-isolation ifindex-list
======================
Ifindex: 1088000(fc2/9)                <<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<< interface
 fc2/9 marked slow
```

2. Identifying the host that is using the interface.

```
switch# show congestion-isolation pmon-list

PMON detected list for vsan 1    : PWWN(FCID)
===============================================

PMON detected list for vsan 2    : PWWN(FCID)
===============================================

PMON detected list for vsan 3    : PWWN(FCID)
===============================================
21:00:00:24:ff:4f:70:46(0x040020)          <<<<<<<<<<<<<<<<<<<<<<host behind interface
 fc2/9 marked slow

PMON detected list for vsan 4    : PWWN(FCID)
===============================================

PMON detected list for vsan 5    : PWWN(FCID)
===============================================
```

3. Remove the interface from being marked as slow.

```
switch# congestion-isolation remove interface fc2/9    <<<<<<<<<<< CLI to remove an
interface from being marked as slow by PMON
```

4. Verifying if the interface is removed from being detected as slow.

```
switch# show congestion-isolation pmon-list

PMON detected list for vsan 1    : PWWN(FCID)
================================================

PMON detected list for vsan 2    : PWWN(FCID)
================================================

PMON detected list for vsan 3    : PWWN(FCID)
================================================

<<<<<<<<<<<<<< host behind interface fc2/9 removed from isolation
PMON detected list for vsan 4    : PWWN(FCID)
================================================

PMON detected list for vsan 5    : PWWN(FCID)
================================================
```

From Cisco MDS NX-OS Release 8.5(1)

1. Identifying the interface that you want to remove from being detected as slow.

```
switch# show fpm congested-device database local
VSAN: 1
----------------------------------
No congested devices found

VSAN: 50
-------------------------------------------------------------------------------------------------
PWWN                     | FCID     | Event type   | Detect type | Detect Time
-------------------------------------------------------------------------------------------------
21:00:f4:e9:d4:54:ac:f8 | 0x7d0000 | credit-stall | local-pmon  | Thu Jan 28 05:08:31
2021
```

2. Remove the interface from being marked as slow.

```
switch# configure
switch(config)# fpm congested-device exclude list
switch(config)# member pwwn 21:00:f4:e9:d4:54:ac:f8 vsan 50
```

3. Verifying if the interface is removed from being detected as slow.

```
switch# show fpm congested-device database local
VSAN: 1
----------------------------------
No congested devices found

VSAN: 50
-------------------------------------------------------------------------------------------------
No congested devices found
```

Configuring Examples for Congestion Isolation Recovery

This example shows how to configure the *isolate-duration* to 24-hours and the number of rising threshold occurrences to be detected in this interval to 3:

```
switch# configure
switch(config)# port-monitor cong-isolation-recover isolate-duration 24 num-occurrence 3
```

This example shows how to configure the *recovery-interval* to 15 minutes:

```
switch# configure
switch(config)# port-monitor cong-isolation-recover recovery-interval 15
```

This example shows how to manually include the device with pWWN 10:00:00:00:c9:f9:16:8d in VSAN 2 as a slow device:

```
switch# configure
switch(config)# fpm congested-device static list
switch(config-congested-dev-static)# member pwwn 10:00:00:00:c9:f9:16:8d vsan 2 credit-stall
```

This example shows how to manually exclude the device with pWWN 10:00:00:00:c9:f9:16:8d in VSAN 2 as a slow device:

```
switch# configure
switch(config)# fpm congested-device exclude list
switch(config-congested-dev-exc)# member pwwn 10:00:00:00:c9:f9:16:8d vsan 2
```

Configuring Examples for Fabric Notifications

This example shows how to enable FPM on a switch:

```
switch# configure
switch(config)# feature fpm
```

This example shows how to disable FPM on a switch:

```
switch# configure
switch(config)# no feature fpm
```

This example shows how to explicitly configure a device with pWWN 10:00:00:00:c9:f9:16:8d in VSAN 2 as congested:

```
switch# configure
switch(config)# fpm congested-device static list
switch(config-congested-dev-static)# member pwwn 10:00:00:00:c9:f9:16:8d vsan 2 credit-stall
```

This example shows how to explicitly exclude a device with pWWN 10:00:00:00:c9:f9:16:8d in VSAN 2 from congestion actions:

```
switch# configure
switch(config)# fpm congested-device exclude list
switch(config-congested-dev-exc)# member pwwn 10:00:00:00:c9:f9:16:8d vsan 2
```

This example shows how to recover the device with pWWN 10:00:00:00:c9:f9:16:8d in VSAN 2 from congestion actions:

```
switch# fpm congested-device recover pwwn 10:00:00:00:c9:f9:16:8d vsan 2
```

This example shows how to configure an FPIN notification interval of 30 seconds:

```
switch# configure
switch(config)# fpm fpin period 30
```

This example shows how to configure the EDC interval for sending congestion signal every 30 seconds:

```
switch# configure
switch(config)# fpm congestion-signal period 30
```

Configuring Examples for DIRL

This example shows how to configure DIRL to specify the ingress reduction rate to 50 percent and ingress recovery rate to 30 percent:

```
switch# configure
switch(config)# fpm dirl reduction 50 recovery 30
```

This example shows how to exclude DIRL based on interface:

```
switch# configure
switch(config)# fpm dirl exclude list
switch(config-dirl-excl)# member interface fc 1/1
switch(config-dirl-excl)# member interface fc 1/1
```

This example shows how to include FC4-type target connected device interface in DIRL:

```
switch# configure
switch(config)# fpm dirl exclude list
switch(config-dirl-excl)# fc4-feature target
```

This example shows how to recover interface fc1/1 which is under DIRL to normal:

```
switch# fpm dirl recover interface fc 1/1
```

Verifying Congestion Management

Verifying Congestion Detection and Avoidance

The following commands display slow-port monitor events:

Note These commands are applicable for both supervisor and module prompts.

Display slow-port monitor events per module:

switch# **show process creditmon slowport-monitor-events [module x [port y]]**

Display the slow-port monitor events on the Onboard Failure Logging (OBFL):

switch# **show logging onboard slowport-monitor-events**

Note The slow-port monitor events are logged periodically into the OBFL.

The following example displays the credit monitor or output of the **creditmon slow-port monitor-events** command for the 16-Gbps and 32-Gbps modules and switches:

```
switch# show process creditmon slowport-monitor-events

        Module: 06      Slowport Detected: YES
================================================================================
Interface = fc6/3
--------------------------------------------------------------------
| admin  | slowport  | oper  |            Timestamp            |
| delay  | detection | delay |                                 |
| (ms)   | count     | (ms)  |                                 |
--------------------------------------------------------------------
|   1    |     46195 |    1  | 1. 10/14/12 21:46:51.615        |
|   1    |     46193 |   50  | 2. 10/14/12 21:46:51.515        |
|   1    |     46191 |   50  | 3. 10/14/12 21:46:51.415        |
|   1    |     46189 |   50  | 4. 10/14/12 21:46:51.315        |
|   1    |     46187 |   50  | 5. 10/14/12 21:46:51.215        |
|   1    |     46185 |   50  | 6. 10/14/12 21:46:51.115        |
|   1    |     46183 |   50  | 7. 10/14/12 21:46:51.015        |
|   1    |     46181 |   50  | 8. 10/14/12 21:46:50.915        |
|   1    |     46179 |   50  | 9. 10/14/12 21:46:50.815        |
|   1    |     46178 |   50  |10. 10/14/12 21:46:50.715        |
--------------------------------------------------------------------
```

TxWait on FCoE or Virtual Fibre Channels (VFC)

Note TxWait on FCoE ethernet or Virtual Fibre Channels (VFC) interfaces is the amount of time a port cannot transmit because of the received Priority Flow Control (PFC) pause frames.

RxWait on FCoE ethernet or VFCs is the amount of time a port cannot receive because of the port transmitting PFC pause frames.

Both TxWait and RxWait are in units of 2.5 microseconds and are converted to seconds in some command outputs. To convert to seconds, multiply the TxWait or RxWait value by 2.5 and divide by 1,000,000.

This example displays the status and statistics of priority-flow-control on all interfaces:

```
switch# show interface priority-flow-control
RxPause: No. of pause frames received
TxPause: No. of pause frames transmitted
TxWait: Time in 2.5uSec a link is not transmitting data[received pause]
RxWait: Time in 2.5uSec a link is not receiving data[transmitted pause]
================================================================================
Interface         Admin Oper (VL bmap) VL  RxPause  TxPause RxWait-     TxWait-
                                                             2.5us(sec)  2.5us(sec)
================================================================================

Po1               Auto  NA      (8)     3    0        0        0(0)      0(0)
Po350             Auto  NA      (8)     3    0        0        0(0)      0(0)
Po351             Auto  NA      (8)     3    0        0        0(0)      0(0)
Po552             Auto  NA      (8)     3    111506   0        0(0)      5014944(12)
Po700             Auto  NA      (8)     3    0        0        0(0)      0(0)
Eth2/17           Auto  Off
Eth2/18           Auto  Off
Eth2/19           Auto  Off
Eth2/20           Auto  Off
Eth2/25           Auto  On      (8)     3    0        0        0(0)      0(0)
Eth2/26           Auto  On      (8)     3    0        0        0(0)      0(0)
```

This example displays the detailed configuration and statistics of a specified virtual Fibre Channel interface:

```
switch# show interface vfc 9/11 counters detailed
vfc9/11
    3108091433 fcoe in packets
    6564116595616 fcoe in octets
    30676987 fcoe out packets
    2553913687 fcoe out octets
    0 2.5us TxWait due to pause frames (VL3)
    134795 2.5us RxWait due to pause frames (VL3)
    0 Tx frames with pause opcode (VL3)
    0 Rx frames with pause opcode (VL3)
    Percentage pause in TxWait per VL3  for last 1s/1m/1h/72h: 0%/0%/0%/0%
    Percentage pause in RxWait per VL3  for last 1s/1m/1h/72h: 0%/0%/0%/0%
```

This example displays the TxWait history information for Ethernet 2/47:

```
switch# show interface e2/47 txwait-history
```

```
TxWait history for port Eth2/47:
================================
     3333333333433333333333333333333333333333333333333333333333333333
     5555655555055546375555555554556555555555465554635665555555555
     5570286961955529745559779559555555580995925559585259577986955
1000
 900
 800
 700
 600
 500
 400 ############## # ########## ############ #### # ############
300 ############################################################
200 ############################################################
100 ############################################################
    0....5....1....1....2....2....3....3....4....4....5....5....6
              0    5    0    5    0    5    0    5    0    5    0

          TxWait per second (last 60 seconds)
                 # = TxWait (ms)

   22222222222
   1111111111600000000000000000000000000000000000000000000000000000
   ..............................................................
   3344454444320000000000000000000000000000000000000000000000000000
60
54
48
42
36
30
24 ############
18 ############
12 ############
 6 ############
  0....5....1....1....2....2....3....3....4....4....5....5....6
            0    5    0    5    0    5    0    5    0    5    0

    TxWait per minute (last 60 minutes)
                 # = TxWait (secs)


     2
     3
     1000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
3600
3240
2880
2520
2160
1800
1440
1080
 720
 360 #
    0....5....1....1....2....2....3....3....4....4....5....5....6....6....7.7
              0    5    0    5    0    5    0    5    0    5    0    5    0 2

    TxWait per hour (last 72 hours)
                 # = TxWait (secs)
```

This example displays the RxWait history information for Ethernet 1/47:

```
switch# show interface e1/47 rxwait-history

RxWait history for port Eth1/47:
  ==============================

      7887777787777778777777788778778777787778777777777777877777778
      9009999909999990999999900990990999909990999999999999199999990
  1000
   900
   800
   700
   600
   500
   400
   300
   200
   100 ############################################################
      0....5....1....1....2....2....3....3....4....4....5....5....6
                0    5    0    5    0    5    0    5    0    5    0

            RxWait per second (last 60 seconds)
                   # = RxWait (ms)

               1
     4444444445570000000000000000000000000000000000000000000000000
     ............................................................
     7777777775870000000000000000000000000000000000000000000000000
  60
  54
  48
  42
  36
  30
  24
  18            #
  12            #
   6 ############
    0....5....1....1....2....2....3....3....4....4....5....5....6
              0    5    0    5    0    5    0    5    0    5    0

            RxWait per minute (last 60 minutes)
                   # = RxWait (secs)


      2                  1                            1
      7                2 5                            9
      0000000000000000060600020000000000000000000000090000000000000000000001
  3600
  3240
  2880
  2520
  2160
  1800
  1440
  1080
   720
   360 #                                               #
      0....5....1....1....2....2....3....3....4....4....5....5....6....6....7.7
                0    5    0    5    0    5    0    5    0    5    0    5    0 2

      RxWait per hour (last 72 hours)
                   # = RxWait (secs)
```

This example displays the onboard failure log (OBFL) for TxWait caused by receiving PFC pause frames:

```
module# show logging onboard txwait
--------------------------------
 Module: 2 txwait count
--------------------------------
----------------------------
 Show Clock
----------------------------
2017-09-22 06:22:17
Notes:
     - Sampling period is 20 seconds
     - Only txwait delta >= 100 ms are logged
-------------------------------------------------------------------------------------
|   Interface     | Delta TxWait Time     | Congestion | Timestamp                |
|                 | 2.5us ticks | seconds |            |                          |
-------------------------------------------------------------------------------------
|   Eth2/1(VL3)   |  2508936    |    6    |      31%   | Fri Sep 22 05:29:21 2017 |
|   Eth2/1(VL3)   |  3355580    |    8    |      41%   | Mon Sep 11 17:55:52 2017 |
|   Eth2/1(VL3)   |  8000000    |   20    |     100%   | Mon Sep 11 17:55:31 2017 |
|   Eth2/1(VL3)   |  8000000    |   20    |     100%   | Mon Sep 11 17:55:11 2017 |
|   Eth2/1(VL3)   |  8000000    |   20    |     100%   | Mon Sep 11 17:54:50 2017 |
```

This example displays the onboard failure log (OBFL) for RxWait caused by transmitting PFC pause frames:

```
module# show logging onboard rxwait
--------------------------------
 Module: 14 rxwait count
--------------------------------
----------------------------
 Show Clock
----------------------------
2017-09-22 11:53:53
Notes:
     - Sampling period is 20 seconds
     - Only rxwait delta >= 100 ms are logged
-------------------------------------------------------------------------------------
|   Interface       | Delta RxWait Time     | Congestion | Timestamp                |
|                   | 2.5us ticks | seconds |            |                          |
-------------------------------------------------------------------------------------
|   Eth14/21(VL3)   |  2860225    |    7    |      35%   | Thu Sep 21 23:59:46 2017 |
|   Eth14/30(VL3)   |    42989    |    0    |       0%   | Thu Sep 14 14:53:57 2017 |
|   Eth14/29(VL3)   |    45477    |    0    |       0%   | Thu Sep 14 14:47:56 2017 |
|   Eth14/30(VL3)   |    61216    |    0    |       0%   | Thu Sep 14 14:47:56 2017 |
|   Eth14/29(VL3)   |    43241    |    0    |       0%   | Thu Sep 14 14:47:36 2017 |
|   Eth14/30(VL3)   |    43845    |    0    |       0%   | Thu Sep 14 14:47:36 2017 |
|   Eth14/29(VL3)   |    79512    |    0    |       0%   | Thu Sep 14 14:47:16 2017 |
|   Eth14/30(VL3)   |    62529    |    0    |       0%   | Thu Sep 14 14:47:16 2017 |
|   Eth14/29(VL3)   |    50699    |    0    |       0%   | Thu Sep 14 14:45:56 2017 |
|   Eth14/30(VL3)   |    47839    |    0    |       0%   | Thu Sep 14 14:45:56 2017 |
```

This example displays the error statistics onboard failure log (OBFL) for a switch:

```
switch# show logging onboard error-stats
----------------------------
 Show Clock
----------------------------
2017-09-22 15:35:31
```

```
-------------------------------------------------------------------------------------------
 STATISTICS INFORMATION FOR DEVICE ID 166 DEVICE Clipper MAC
-------------------------------------------------------------------------------------------
 Port |                                          |               |    Time Stamp   |In
Range |            Error Stat Counter Name       |    Count      |MM/DD/YY HH:MM:SS|st
      |                                          |               |                 |Id
-------------------------------------------------------------------------------------------
11    |GD rx pause transitions of XOFF-XON VL3   |2147           |09/22/17 00:11:24|02
11    |GD uSecs VL3 is in internal pause rx state |7205308        |09/22/17 00:11:24|02
11    |GD rx frames with pause opcode for VL3    |6439           |09/22/17 00:11:24|02
11    |PL SW pause event (vl3)                   |113            |09/22/17 00:11:24|02
```

Note For 16-Gbps modules, 32-Gbps modules, and Cisco MDS 9700, 9148S, 9250i, and 9396S switches, if **no-credit-drop** timeout is configured, the maximum value of **tx-slowport-oper-delay** as shown in slow-port monitor events is limited by the **no-credit-drop timeout**. So, the maximum value for **tx-slowport-oper-delay** can reach the level of the **no-credit-drop** timeout even if the actual slow-port delay from the device is higher because the frames are forcefully dropped by the hardware when **tx-slowport-oper-delay** reaches the level of the **no-credit-drop** timeout.

Verifying Congestion Isolation

This example show how to verify system flow-control mode:

```
switch# show system fc flow-control
System flow control is ER_RDY
```

This example shows how to verify the Congestion Isolation status:

```
switch# show congestion-isolation status
Flow Control Mode     : ER_RDY
Congestion Isolation  : Enabled
Sampling Interval     : 1
Timeout               : 0
ESS Cap Details
-------------------------------------------
VSAN: 0x1(1)
Enabled domain-list: 0x4(4 - local)
Disabled domain-list: None
Unsupported domain-list: 0x61(97)
VSAN: 0x2(2)
Enabled domain-list: 0x4(4 - local)
Disabled domain-list: None
Unsupported domain-list: 0xb8(184)
VSAN: 0x3(3)
Enabled domain-list: 0x4(4 - local)
Disabled domain-list: None
Unsupported domain-list: None
VSAN: 0x4(4)
Enabled domain-list: 0x4(4 - local) 0xbb(187)
Disabled domain-list: None
Unsupported domain-list: None
```

This example shows how to verify the list of devices that were detected as slow on a local switch:

```
switch# show congestion-isolation pmon-list vsan 4
PMON detected list for vsan 4    : PWWN(FCID)
================================================
10:00:00:00:c9:f9:16:8d(0xbe0000)
```

This example shows how to verify the global list of devices that were detected as slow in a fabric when the Congestion Isolation feature was enabled. The global list should be the same on all switches in the fabric where the Congestion Isolation feature is enabled.

```
switch# show congestion-isolation global-list vsan 4
Global list for vsan 4 PWWN(FCID)
=======================================
10:00:00:00:c9:f9:16:8d(0xbe0000)
```

This example shows the list of devices that were detected as slow on remote switches (not locally detected slow devices):

```
switch# show congestion-isolation remote-list vsan 4
Remote list for vsan 4    : PWWN(FCID)
================================================
10:00:00:00:c9:f9:16:8d(0xbe0000)
```

This example shows a single device that is marked as slow (feature slow-dev) either via the port monitor or the **congestion isolation include** command:

```
switch# show congestion-isolation include-list vsan 4
Include list for vsan 4   : PWWN(FCID) (online/offline)
=========================================================
10:00:00:00:c9:f9:16:8d(0xbe0000) - (Online)

switch# show fcns database vsan 4
VSAN 4:
--------------------------------------------------------------------------
FCID        TYPE  PWWN                    (VENDOR)        FC4-TYPE:FEATURE
--------------------------------------------------------------------------
0x040000    N     10:00:40:55:39:0c:80:85 (Cisco)         ipfc
0x040020    N     21:00:00:24:ff:4f:70:47 (Qlogic)        scsi-fcp:target
0xbe0000    N     10:00:00:00:c9:f9:16:8d (Emulex)        scsi-fcp:init slow-dev <<<slow
device
[testing]Total number of entries = 3
```

This example shows the list of devices that were manually configured using Congestion Isolation exclude list command on a local switch:

```
switch# show congestion-isolation exclude-list vsan 4
Exclude list for vsan 4   : PWWN(FCID) (online/offline)
=========================================================
10:00:00:00:c9:f9:16:8d(0xbe0000) - (Online)
```

Verifying Congestion Isolation Recovery

This example shows how to check the configured *isolate-duration*, *recovery-interval*, and number of rising threshold occurrences:

```
switch# show port-monitor

Port Monitor : enabled
DIRL :
      Recovery Interval   : 60 seconds
FPIN :
      Recovery Interval   : 900 seconds
Cong-isolate-recover :
      Recovery Interval   : 900 seconds
      Isolation Duration   : 24 hours
      Number of Isolation occurrences   : 3
--------------------------------------------------------------------------------

Policy Name : default
Admin status : Not Active
Oper status  : Not Active
Logical type : All Ports
--------------------------------------------------------------------------------------------------------------------------------------------
|        Counter        | Threshold  | Interval |     Warning      |     Thresholds     |         Rising/Falling actions     | Congestion-signal
 |
|                       |    Type    |  (Secs)
|-----------------------|------------|----------------------------------------------|------|
|                       |            |          | Threshold | Alerts |  Rising  | Falling  | Event |     Alerts       | PortGuard| Warning |
Alarm   |
--------------------------------------------------------------------------------------------------------------------------------------------
| Link Loss             | Delta      | 60       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| Sync Loss             | Delta      | 60       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| Signal Loss           | Delta      | 60       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| Invalid Words         | Delta      | 60       | none      | n/a    | 1        | 0        | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| Invalid CRC's         | Delta      | 60       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| State Change          | Delta      | 60       | none      | n/a    | 5        | 0        | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| TX Discards           | Delta      | 60       | none      | n/a    | 200      | 10       | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| LR RX                 | Delta      | 60       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| LR TX                 | Delta      | 60       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| Timeout Discards      | Delta      | 60       | none      | n/a    | 200      | 10       | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| Credit Loss Reco      | Delta      | 60       | none      | n/a    | 1        | 0        | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| TX Credit Not Available | Delta    | 60       | none      | n/a    | 10%      | 0%       | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| RX Datarate           | Delta      | 10       | none      | n/a    | 80%      | 70%      | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| TX Datarate           | Delta      | 10       | none      | n/a    | 80%      | 70%      | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| TX-Slowport-Oper-Delay | Absolute  | 60       | none      | n/a    | 50ms     | 0ms      | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| TXWait                | Delta      | 60       | none      | n/a    | 30%      | 10%      | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| RX Datarate Burst     | Delta      | 10       | none      | n/a    | 5@90%    | 1@90%    | 4     | syslog,rmon,obfl | none     | n/a     |
n/a     |
| TX Datarate Burst     | Delta      | 10       | none      | n/a    | 5@90%    | 1@90%    | 4     | syslog,rmon,obfl | none     | n/a     |
n/a     |
| Input Errors          | Delta      | 60       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | none     | n/a     |
n/a     |
--------------------------------------------------------------------------------------------------------------------------------------------

Policy Name : slowdrain
Admin status : Not Active
Oper status  : Not Active
Logical type : All Edge Ports
--------------------------------------------------------------------------------------------------------------------------------------------
|        Counter        | Threshold  | Interval |     Warning      |     Thresholds     |   Rising/Falling actions          | Congestion-signal
 |
|                     |    Type    |  (Secs)  |------------------|--------------------|-----------------------------------|-----|
|                       |            |          | Threshold | Alerts |  Rising  | Falling  | Event |     Alerts       |PortGuard| Warning |
Alarm   |
--------------------------------------------------------------------------------------------------------------------------------------------
| Credit Loss Reco      | Delta      | 1        | none      | n/a    | 1        | 0        | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| TX Credit Not Available | Delta    | 1        | none      | n/a    | 10%      | 0%       | 4     | syslog,rmon      | none     | n/a     |
n/a     |
| TX Datarate           | Delta      | 10       | none      | n/a    | 80%      | 70%      | 4     | syslog,obfl      | none     | n/a     |
n/a     |
--------------------------------------------------------------------------------------------------------------------------------------------

Policy Name : fabricmon_edge_policy
Admin status : Not Active
Oper status  : Not Active
Logical type : All Edge Ports
--------------------------------------------------------------------------------------------------------------------------------------------
|        Counter        | Threshold  | Interval |     Warning      |     Thresholds     |     Rising/Falling actions        | Congestion-signal
 |
|                     |    Type    |  (Secs)  |------------------|--------------------|-----------------------------------|-----|
|                       |            |          | Threshold | Alerts |  Rising  | Falling  | Event |     Alerts       |PortGuard| Warning |
```

```
Alarm   |
--------------------------------------------------------------------------------------------------------------------------------
| Link Loss               | Delta      | 30       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | FPIN    | n/a     | n/a
    |
| Sync Loss               | Delta      | 30       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | FPIN    | n/a     | n/a
    |
| Signal Loss             | Delta      | 30       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | FPIN    | n/a     | n/a
    |
| Invalid Words           | Delta      | 30       | none      | n/a    | 1        | 0        | 4     | syslog,rmon      | FPIN    | n/a     | n/a
    |
| Invalid CRC's           | Delta      | 30       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | FPIN    | n/a     | n/a
    |
| State Change            | Delta      | 60       | none      | n/a    | 5        | 0        | 4     | syslog,rmon      | none    | n/a     | n/a
    |
| TX Discards             | Delta      | 60       | none      | n/a    | 200      | 10       | 4     | syslog,rmon      | none    | n/a     | n/a
    |
| LR RX                   | Delta      | 60       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | none    | n/a     | n/a
    |
| LR TX                   | Delta      | 60       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | none    | n/a     | n/a
    |
| Timeout Discards        | Delta      | 60       | none      | n/a    | 200      | 10       | 4     | syslog,rmon      | none    | n/a     | n/a
    |
| Credit Loss Reco        | Delta      | 1        | none      | n/a    | 1        | 0        | 4     | syslog,rmon      | none    | n/a     | n/a
    |
| TX Credit Not Available | Delta      | 1        | none      | n/a    | 10%      | 0%       | 4     | syslog,rmon      | none    | n/a     | n/a
    |
| RX Datarate             | Delta      | 10       | none      | n/a    | 80%      | 70%      | 4     | syslog,rmon,obfl | none    | n/a     | n/a
    |
| TX Datarate             | Delta      | 10       | none      | n/a    | 80%      | 70%      | 4     | syslog,rmon,obfl | none    | n/a     | n/a
    |
| TX-Slowport-Oper-Delay  | Absolute   | 1        | none      | n/a    | 50ms     | 0ms      | 4     | syslog,rmon      | none    | n/a     | n/a
    |
| TXWait                  | Delta      | 1        | none      | n/a    | 30%      | 10%      | 4     | syslog,rmon      | FPIN    | 40%     | 60%
    |
| RX Datarate Burst       | Delta      | 10       | none      | n/a    | 5@90%    | 1@90%    | 4     | syslog,rmon,obfl | none    | n/a     | n/a
    |
| TX Datarate Burst       | Delta      | 10       | none      | n/a    | 5@90%    | 1@90%    | 4     | syslog,rmon,obfl | none    | n/a     | n/a
    |
| Input Errors            | Delta      | 60       | none      | n/a    | 5        | 1        | 4     | syslog,rmon      | none    | n/a     | n/a
    |
--------------------------------------------------------------------------------------------------------------------------------

 On falling threshold portguard actions FPIN, DIRL, Cong-Isolate-Recover will initiate auto recovery of ports.
```

Verifying FPIN

This example shows the number of devices registered for FPIN in each VSAN:

```
switch# show fpm fpin
C: Congestion Notification Descriptor
P: Peer Congestion Notification Descriptor
L: Link Integrity Notification Descriptor
D: Delivery Notification Descriptor
U: Priority Update Notification Descriptor
A: Alarm Signal
W: Warning Signal

VSAN: 1
--------------------------------------------------------------------------------------------
FCID                     |            RDF             | FPIN sent  | Last FPIN sent timestamp
PWWN                     | Registered | Negotiated    |   count    |
                         |          Timestamp         |            |
--------------------------------------------------------------------------------------------
0xdc06e0                 | L          | L             | L:       0 | L: --
10:00:00:10:9b:95:41:22  | Tue Feb  2 03:38:13 2021   |            |


VSAN: 50
--------------------------------------------------------------------------------------------
FCID                     |            RDF             | FPIN sent  | Last FPIN sent timestamp
PWWN                     | Registered | Negotiated    |   count    |
                         |          Timestamp         |            |
--------------------------------------------------------------------------------------------
0x7d0000                 | CPLD       | CPL           | L:       0 | L: --
21:00:f4:e9:d4:54:ac:f8  | Mon Feb  1 15:32:26 2021   | C:       0 | C: --
                         |            |               | P:       0 | P: --
```

```
0x7d0020                | CPLD       | CPL          | L:       0 | L: --
21:00:f4:e9:d4:54:ac:f9 | Mon Feb  1 15:32:27 2021 | C:       0 | C: --
                        |            |              | P:       0 | P: --
```

This example shows a summary of RDF and EDC registrations:

```
switch# show fpm registration summary
C: Congestion Notification Descriptor
P: Peer Congestion Notification Descriptor
L: Link Integrity Notification Descriptor
D: Delivery Notification Descriptor
U: Priority Update Notification Descriptor
A: Alarm Signal
W: Warning Signal

VSAN: 1
--------------------------------------------------------------------------
FCID     | PWWN                    | FPIN          | Congestion Signal
         |                         | Registrations | Registrations
--------------------------------------------------------------------------
0xdc06e0 | 10:00:00:10:9b:95:41:22 | L             | --

VSAN: 50
--------------------------------------------------------------------------
FCID     | PWWN                    | FPIN          | Congestion Signal
         |                         | Registrations | Registrations
--------------------------------------------------------------------------
0x7d0000 | 21:00:f4:e9:d4:54:ac:f8 | CPLD          | AW
0x7d0020 | 21:00:f4:e9:d4:54:ac:f9 | CPLD          | AW
```

This example shows EDC registration in detail:

```
switch# show fpm registration congestion-signal
A: Alarm
W: Warning
ms: milliseconds

VSAN: 1
-----------------------------------------------------------------------------------------------
No registered devices found

VSAN: 50
-----------------------------------------------------------------------------------------------
FCID     | PWWN                    |     Device Tx    |     Device Rx    |  Negotiated Tx
         |                         | Capa- | Interval | Capa- | Interval | Capa- | Interval

         |                         | bility|  (ms)    | bility|  (ms)    | bility|  (ms)
-----------------------------------------------------------------------------------------------
0x7d0020 | 21:00:f4:e9:d4:54:ac:f9 | AW    |       10 | AW    |       10 | AW    |     1000
0x7d0000 | 21:00:f4:e9:d4:54:ac:f8 | AW    |       10 | AW    |       10 | AW    |     1000
```

This example shows the list of devices that were detected as congested devices by port monitor:

```
switch# show fpm congested-device database local
VSAN: 1
----------------------------------
No congested devices found

VSAN: 50
```

```
-------------------------------------------------------------------------------------------------
PWWN                    | FCID     | Event type   | Detect type | Detect Time
-------------------------------------------------------------------------------------------------
21:00:f4:e9:d4:54:ac:f8 | 0x7d0000 | credit-stall | local-pmon  | Thu Jan 28 05:08:31 2021
```

This example shows a list of remote devices that are congested:

```
switch# show fpm congested-device database remote
VSAN: 1
----------------------------------
No congested devices found

VSAN: 50
----------------------------------
No congested devices found

VSAN: 70
----------------------------------
No congested devices found

VSAN: 80
----------------------------------
No congested devices found

VSAN: 1001
-------------------------------------------------------------------------------------------------
PWWN                    | FCID     | Event type   | Detect type | Detect Time
-------------------------------------------------------------------------------------------------
21:00:34:80:0d:6c:a7:63 | 0xec0000 | credit-stall | remote      | Thu Jan 28 05:12:00 2021
```

This example shows the list of devices that were manually included as congested devices:

```
switch# show fpm congested-device database static
VSAN: 1
----------------------------------
No congested devices found

VSAN: 50
-----------------------------------------------------
PWWN                    | FCID     | Event type
-----------------------------------------------------
21:00:f4:e9:d4:54:ac:f8 | 0x7d0000 | credit-stall
```

This example shows the list of congested devices that are excluded:

```
switch# show fpm congested-device database exclude
VSAN: 1
----------------------------------
No congested devices found

VSAN: 50
-------------------------------------
PWWN                    | FCID
-------------------------------------
21:00:f4:e9:d4:54:ac:f8 | 0x7d0000
```

Verifying DIRL

This example shows the configured DIRL reduction and recovery percentages:

```
switch# show fpm ingress-rate-limit status
dirl reduction rate:50%
dirl recovery rate:25%
--------------------------------------------------------------------------------
Interface  Current rate   Rate-limit-type  Previous action  Last update time
           limit(%)
--------------------------------------------------------------------------------
fc4/12     10.6435          dynamic          recovered        Wed Jan 27 20:23:34 2021

fc7/5      12.9567          dynamic          recovered        Wed Jan 27 20:23:34 2021
```

This example shows the configured DIRL reduction and recovery percentages for the port fc4/12:

```
switch# show fpm ingress-rate-limit status interface fc4/12
dirl reduction rate:50%
dirl recovery rate:25%
--------------------------------------------------------------------------------
Interface  Current rate  Rate-limit-type  Previous action  Last update time
           limit(%)
--------------------------------------------------------------------------------
fc4/12     10.6435          dynamic          recovered        Wed Jan 27 20:23:34 2021
```

This example shows the list of interfaces that are excluded from DIRL rate reduction:

```
switch# show fpm dirl exclude
All target device connected interface are excluded from DIRL
--------------
Interface
--------------
fc4/19
fc4/21
fc7/13
```