Troubleshooting FCoE Issues

Fibre Channel over Ethernet (FCoE) provides a method of transporting Fibre Channel traffic over a physical Ethernet connection. FCoE requires that the underlying Ethernet be full duplex and provides lossless behavior for Fibre Channel traffic.

This chapter describes how to identify and resolve problems that can occur with FCoE in the Cisco Nexus 5000 Series switch.

This chapter includes the following sections:

- Data Center Bridging
- FIP
- CNA
- PFC
- Registers and Counters

Data Center Bridging

VFC (FCoE) interface not online

This section includes the following topics:

- General troubleshooting
- Nexus 5548 Troubleshooting

General troubleshooting

Possible Cause

An FCoE-attached server has no connectivity to FC, or FCoE-attached storage, and the **show interface** command for the virtual Fibre Channel interface mapped to this server's port reveals that the VFC interface is down.

Solution

- Verify the configuration using the **show running-config** command.

  Example:

Note The default setting for VFC is shutdown, however, in the following example was changed by the setup script.

```
switch# show running-config
feature fcoe
vlan 1
vlan 100
fcoe
vsan database
vsan 100
interface vfc4
bind interface Ethernet1/4
no shutdown
vsan database
vsan 100 interface vfc4
interface fc2/1
no shutdown
interface Ethernet1/4
switchport mode trunk
switchport trunk allowed vlan 100
spanning-tree port type edge trunk
```

- Check to ensure that the LLDP Transmit and Receive are enabled on the interface.
  Use the **show lldp interface ethernet 1/4** command.

  Example:

```
switch# show lldp interface ethernet 1/4

Interface Information:
Enable (tx/rx/dcbx): Y/Y/Y Port Mac address: 00:0d:ec:d5:a3:8b
Peer's LLDP TLVs:
Type Length Value
---- ------ -----
001 007 0400c0dd 145486
002 007 0300c0dd 145486
003 002 0078
128 061 001b2102 020a0000 00000002 00000001 04110000 c0000001 00003232
00000000 00000206 060000c0 00080108 100000c0 00890600 1b210889
14001b21 08
000 000
```

  If LLDP is disabled, the VFC will not come online.
  You can enable LLDP transmit and receive with the **interface ethernet 1/4** command:

```
switch(config)# interface ethernet 1/4
switch(config-if)# lldp ?
receive Enable LLDP reception on interface
transmit Enable LLDP transmission on interface
```

- Check that the peer supports LLDP.
  Check if remote peers exist. Check if values exist for a peer's LLDP TLVs.
  Use the **show lldp interface ethernet 1/4** command.

  Example:

```
switch# show lldp interface ethernet 1/4
Interface Information:
Enable (tx/rx/dcbx): Y/Y/Y Port Mac address: 00:0d:ec:d5:a3:8b
Peer's LLDP TLVs:
Type Length Value
---- ------ -----
```

```
001 007 0400c0dd 145486
002 007 0300c0dd 145486
003 002 0078
128 061 001b2102 020a0000 00000002 00000001 04110000 c0000001 00003232
00000000 00000206 060000c0 00080108 100000c0 00890600 1b210889
14001b21 08
```

- Check the peer (CNA) to see if it supports DCBX.
  Use the **show system internal dcbx info interface ethernet 1/4** command.
  (For releases earlier than 4.2(1)N1, use the "sh platform software dcbx internal info interface ethernet x/y" command.)

Note In the example, DCBX is enabled and the peer supports CEE.

Example:

```
switch# show system internal dcbx info interface ethernet 1/4
Interface info for if_index: 0x1a003000(Eth1/4)
tx_enabled: TRUE
rx_enabled: TRUE
dcbx_enabled: TRUE
DCX Protocol: CEE
Port MAC address: 00:0d:ec:d5:a3:8b
DCX Control FSM Variables: seq_no: 0x1, ack_no: 0x2,my_ack_no: 0x1, peer_seq_no: 0x2
oper_version: 0x0, max_version: 0x0 fast_retries 0x0
Lock Status: UNLOCKED
PORT STATE: UP
```

- In the output from the **show system internal dcbx info interface ethernet 1/4** command, check the peers LLDP values.
  Make sure that the mandatory LLDP values exist.

  Example:

```
switch# show system internal dcbx info interface ethernet 1/4

LLDP Neighbors
Remote Peers Information on interface Eth1/4
Remote peer's MSAP: length 12 Bytes:
00 c0 dd 14 54 86 00 c0 dd 14 54 86
LLDP TLV's
LLDP TLV type:Chassis ID LLDP TLV Length: 7
04 00 c0 dd 14 54 86
Chassis type: 04 Chassis ID:00 c0 dd 14 54 86
LLDP TLV type:Port ID LLDP TLV Length: 7
03 00 c0 dd 14 54 86
Port ID subtype: 03 Port ID:00 c0 dd 14 54 86
LLDP TLV type:Time to Live LLDP TLV Length: 2
00 78
TTL = 00
LLDP TLV type:Unknown 128 LLDP TLV Length: 61
00 1b 21 02 02 0a 00 00 00 00 00 02 00 00 00 01 04 11 00 00 c0 00
00 01 00 00 32 32 00 00 00 00 00 00 02 06 06 00 00 c0 00 08 01 08
10 00 00 c0 00 89 06 00 1b 21 08 89 14 00 1b 21 08
LLDP TLV type:END of LLDP PDU LLDP TLV Length: 0
```

- In the output from the **show system internal dcbx info interface ethernet 1/4** command, check the peers DCBX TLVs.
  Make sure that PFC and FCoE TLV were negotiated as willing and enabled, and that there are no errors.

  Example:

```
switch# show system internal dcbx info interface ethernet 1/4

Peer's DCX TLV:
DCBX TLV Proto(1) type: 1(Control) DCBX TLV Length: 10 DCBX TLV Value
00 00 02 00 00 00 01 00 00 00
sub_type 0, error 0, willing 0, enable 0, max_version 0, oper_version 0
DCBX TLV Proto(1) type: 2(PriGrp) DCBX TLV Length: 17 DCBX TLV Value
00 00 c0 00 00 01 00 00 32 32 00 00 00 00 00 00 02
sub_type 0, error 0, willing 1, enable 1, max_version 0, oper_version 0
DCBX TLV Proto(1) type: 3(PFC) DCBX TLV Length: 6 DCBX TLV Value
00 00 c0 00 08 01
sub_type 0, error 0, willing 1, enable 1, max_version 0, oper_version 0
DCBX TLV Proto(1) type: 4(App(Fcoe)) DCBX TLV Length: 16 DCBX TLV Value
00 00 c0 00 89 06 00 1b 21 08 89 14 00 1b 21 08
sub_type 0, error 0, willing 1, enable 1, max_version 0, oper_version 0
```

- Check the peer PFC and FCoE subtypes.

  Use the show **system internal dcbx info interface ethernet 1/4** command.
  (For releases earlier than 4.2(1)N1, use the **sh platform software dcbx internal info interface ethernet x/y** command.)

  Example:

```
switch# show system internal dcbx info interface ethernet 1/4

Feature type PFC (3)
feature type 3(PFC)sub_type 0
Feature State Variables: oper_version 0 error 0 local error 0 oper_mode 1
feature_seq_no 0 remote_feature_tlv_present 1 remote_tlv_aged_out 0
remote_tlv_not_present_notification_sent 0
Feature Register Params: max_version 0, enable 1, willing 0 advertise 1
disruptive_error 0 mts_addr_node 0x101 mts_addr_sap 0x179
Desired config cfg length: 2 data bytes:08 08
Operating config cfg length: 2 data bytes:08 08
Peer config cfg length: 0 data bytes:
Feature type App(Fcoe) (4)sub_type FCoE (0)
feature type 4(App(Fcoe))sub_type 0
Feature State Variables: oper_version 0 error 0 local error 0 oper_mode 1
feature_seq_no 0 remote_feature_tlv_present 1 remote_tlv_aged_out 0
remote_tlv_not_present_notification_sent 0
Feature Register Params: max_version 0, enable 1, willing 0 advertise 1
disruptive_error 0 mts_addr_node 0x101 mts_addr_sap 0x179
```

- Check the DCBX counters located at the very bottom of the output display from the **show system internal dcbx info interface ethernet 1/4** command. Look for any errors.

  Example:

```
Traffic Counters
DCBX pkt stats:
Total frames out: 15383
Total Entries aged: 97
Total frames in: 15039
DCBX frames in: 15033
Total frames received in error: 6
Total frames discarded: 6
Total TLVs unrecognized: 0
```

- Check for the same values for the FCoE Data Center Bridging and the Type-Length-Value on the host CNA software.
- Ensure that the VSAN trunk protocol has been enabled.

  Use the **configuration terminal** command to enter into configuration mode and use the **trunk protocol enable** command to enable the trunking protocol.

Solution Summary

- Review every feature negotiation result.

  Use the **show system internal dcbx info interface ethernet 1/4** command.
  (For releases earlier than 4.2(1)N1, use the **sh platform software dcbx internal info interface ethernet x/y** command.)

  Example:

```
switch# show system internal dcbx info interface ethernet 1/4

feature type 3 sub_type 0
feature state variables: oper_version 0 error 0 oper_mode 1 feature_seq_no 0
remote_feature_tlv_present 1
remote_tlv_not_present_notification_sent 0 remote_tlv_aged_out 0
feature register params max_version 0, enable 1, willing 0 advertise 1,
disruptive_error 0 mts_addr_node
0x101mts_addr_sap 0x1e5
Desired config cfg length: 1 data bytes:08
Operating config cfg length: 1 data bytes:08
```

- Errors
  - Indicates negotiation error.
  - Never expected to happen when connected to CNA.
  - When two Nexus 5000 switches are connected back-to-back, and if PFC is enabled on different CoS values, then a negotiation error can occur.
- Operating configuration
  - Indicates negotiation result.
  - Absence of operating configuration indicates that the peer does not support the DCBX TLV or that there is a negotiation error.
  - The remote_feature_tlv_present message indicates whether the remote peer supports this feature TLV or not.
- DCBX feature might not be working because:
  - Peer does not support the LLDP Protocol.
  - Peer does not support the DCBX Protocol.
  - Peer does not support some DCBX TLVs.
  - Unexpected DCBX negotiation result.
- An option exists to force PFC mode on an interface.
  Use the **inerfacet ethernet 1/21** command and the **priority-flow-control mode** command to force the PFC mode.

  Example:

```
switch(config)# int eth1/21
switch(config-if)# priority-flow-control mode ?
```

```
auto Advertise priority-flow-control capability
on Turn on priority-flow-control
```

Note The default setting for this command is auto. The **no** option returns the mode to auto.

Nexus 5548 Troubleshooting

Possible Cause

The type of Converged Network Adapter might not be supported.

Solution

Ensure that the type of adapter is supported. The FCoE interface only supports a Generation-2 Converged Network Adapter.

Possible Cause

The FCoE class-fcoe system class is not enabled in the QoS configuration.

Solution

For a Cisco Nexus 5548 switch, the FCoE class-fcoe system class is not enabled by default in the QoS configuration. Before enabling FCoE, you must include class-fcoe in each of the following policy types:

- Network-QoS
- Queuing
- QoS

The following is an example of a service policy that needs to be configured:

```
F340.24.10-5548-1
class-map type qos class-fcoe
class-map type queuing class-fcoe
match qos-group 1
class-map type queuing class-all-flood
match qos-group 2
class-map type queuing class-ip-multicast
match qos-group 2
class-map type network-qos class-fcoe
match qos-group 1
class-map type network-qos class-all-flood
match qos-group 2
class-map type network-qos class-ip-multicast
match qos-group 2
system qos


 service-policy type qos input fcoe-default-in-policy
 service-policy type queuing input fcoe-default-in-policy
 service-policy type queuing output fcoe-default-out-policy
 service-policy type network-qos fcoe-default-nq-policy
```

FIP

Note FIP Generation-1 CNAs are not supported on the Nexus 2232 FEX. Only FIP Generation-2 CNAs are supported on the Nexus 2232 FEX.

VFC down due to FIP failure

Host is not capable of supporting FIP-related TLVs.

Possible Cause

When the connected host does not support FIP, the first step of VLAN-discovery fails based on which VFC is brought up. Use show commands to verify that the three basic TLVs required for FIP are exchanged by DCBX over the bound interface, and that FCOE-MGR is enabled for FIP. The three TLVs are FCoE TLV, PriGrp TLV, and PFC TLV. These three TLVs should be checked for both local and peer values.

Verify the TLVs with the following commands:

- **show system internal dcbx info interface** <*bound-ethernet-interface-id*>
- **show platform software fcoe_mgr info interface vfc**<*id*>

In the output from the commands:

- Check for FIP capable is TRUE.
- Check for triggered event: [FCOE_MGR_VFC_EV_FIP_VLAN_DISCOVERY].

The state of the VFC never progresses further to solicitation.

Solution

Make sure you check for correct FIP supporting firmware and drivers on the CNA and FIP supporting adapters.

VFC down due to FIP solicitation failure

When the FIP solicitation fails, the VFC goes down.

Possible Cause

Once the first step of FIP VLAN-discovery has succeeded, the host sends FIP solicitations. The switch should respond with FIP advertisements in detail. If the response is not sent or the advertisement is not sent back to the solicitation received, the VFC does not come up. The host continues trying to solicit, but never succeeds.

The following are possible reasons for no response or advertisement:

- No active fabric-provided MAC address exists. (Possible wrong fc-map, etc.)
- Fabric is not available for FLOGI.
- MAC address descriptor may be incorrect. (This is the address the CNA uses as the DMAC when it sends responses.)

Use the **show platform software fcoe_mgr info interface vfc**<*id*> command to view the status of the FIP solicitation.

In the output from the command, check for triggered event: [FCOE_MGR_VFC_EV_FIP_VLAN_DISCOVERY];

followed by triggered event: [FCOE_MGR_VFC_EV_FIP_SOLICITATION].

If the solicitation is successful, then triggered event: [FCOE_MGR_VFC_EV_FIP_FLOGI] is displayed.

If the solicitation has failed, then triggered event: [FCOE_MGR_VFC_EV_FIP_FLOGI] is not displayed and no further progress occurs.

Solution

Need to check and ensure that the VSAN is active, the memberships are correct, and that the fabric is available. Also while in NPV mode, check that an active border/NP port is available.

VFC down because VLAN response not received by CNA

Though the switch sends out a VLAN response, the response is not received by the CNA. This indicates that the VFC is down.

Possible Cause

A bound interface native VLAN ID should be a non-FCoE VLAN. If not, and the native VLAN matches the FCoE VLAN, the VLAN response sent out will be untagged. However, the FIP adapters expect tagged frames. This means that the native VLAN on the trunk interface should be a non-FCoE VLAN.

Solution

Check the configuration on the bound Ethernet trunk interface and ensure that it is a non-FCoE native VLAN.

VFC down because no active STP port-state on the bound Ethernet interface

No active STP port-state on the bound Ethernet interface causes the VFC to be down.

Possible Cause

The bound interface should be in a STP-forwarding state for both the native VLAN and the member FCOE VLAN mapped to the active VSAN. If there are no STP active ports on the VLAN, then the switch drops all FIP packets received on the VLAN over the bound interface. This means that the FIP is not initiated to bring up the VFC.

Solution

Check the STP port state on the bound Ethernet trunk interface for both non-FCOE native VLAN and FCOE member VLAN. Fix the STP port state and move it to forwarding, if in blocked inconsistent state or error-disable state.

VFC down due to FIP keepalive misses

The VFC goes down due to FIP keepalive misses.

Possible Cause

When FIP keepalives (FKA) are missed for a period of approximately 22 seconds, this means that approximately three FKAs are not continuously received from the host. Missed FKAs can occur for many reasons, including congestion or link issues.

FKA timeout : 2.4 * FKA_adv_period.

The FKA_adv_period is exchanged and agreed upon with the host as in the FIP advertisement when responding to a solicitation.

Observe the output from the following commands to confirm FKA misses:

- **show platform software fcoe_mgr info interface vfc**<*id*>
- **show platform software fcoe_mgr event-history errors**
- **show platform software fcoe_mgr event-history lock**
- **show platform software fcoe_mgr event-history msgs**
- **show platform fwm info pif ethernet** <*bound-ethernet-interface-id*>

Solution

Sometimes when congestion is relieved, the VFC comes back up. If the symptom persists, then additional analysis is required. The possible considerations are:

- The host stopped sending the FKA.
- The switch dropped the FKA that was received.

CNA

This section includes an overview of best practices for the topology of the Converged Network Adapter (CNA), a description of troubleshooting with host-based tools, followed by a description of common problems and their solutions.

Best practice topology for CNA

Best Practice Topology for Direct Connected CNA

- A unique dedicated VLAN must be configured at every converged access switch to carry traffic for each virtual fabric (VSAN) in the SAN (for example, VLAN 1002 for VSAN 1, VLAN 1003 for VSAN 2, and so on). If MSTP is enabled, a separate MST instance must be used for FCoE VLANs
- Unified Fabric (UF) links must be configured as trunk ports. FCoE VLAN must not be configured as a native VLAN. All FCoE VLANs must be configured as members of the UF links. This allows it to be extendible for VF_Port trunking and VSAN management for the VFC interfaces.
- UF links must be configured as spanning tree edge ports.
- FCoE VLANs must not be configured as members of Ethernet links that are not designated to carry FCoE traffic. This ensures to limit the scope of the spanning-tree protocol for FCoE VLANs to UF links only.
- If the converged access switches (in the same SAN fabric or in the other) need to be connected to each over Ethernet links for the purposes of LAN alternate pathing, then such links must explicitly be configured to exclude all FCoE VLANs from membership. This ensures to limit the scope of the Spanning Tree Protocol for FCoE VLANs to UF links only.
- Separate FCoE VLANs must be used for FCoE in SAN-A and SAN-B.

Best Practice Topology for Remote Connected CNAs

- A unique dedicated VLAN must be configured at every converged access switch and every blade switch to carry traffic for each virtual fabric (VSAN) in the SAN (for example, VLAN 1002 for VSAN 1, VLAN 1002 for VSAN 2, and so on). If MSTP is enabled, a separate MST instance must be used for FCoE VLANs.
- Unified Fabric (UF) links must be configured as trunk ports. FCoE VLAN must not be configured as a native VLAN. All FCoE VLANs must be configured as members of the UF links. This allows it to be extendible for VF_Port trunking and VSAN management for the VFCs.
- UF links between the CNAs and the blade switches must be configured as spanning tree edge ports.
- A blade switch must connect to exactly one converged access switch, preferably over an Ethernet port channel to avoid disruption due to STP reconvergence on events such as provisioning of new links or blade switches.

Troubleshooting with Host tools

You can troubleshoot the CNA with following host-based tools:

- Emulex
  - Emulex provides the OneCommand GUI tool to manage Emulex CNAs. The CEE tab of this tool displays details about DCB configurations and FIP settings within the FC interface.
- Qlogic
  - Qlogic provides the SanSurfer tool. The Data Center Bridging tab of this tool displays the DCB configuration learned from the switch alone with TLV exchange data. The DCE Statistics tab of this tool displays the ethernet statistics.
- Microsoft Windows
  - Microsoft Windows provides tools to view the configuration and registers for many CNA vendor products.

CNA not recognized by Host OS

Although the CNA is installed on the host, the Converged Network Adapter (CNA) is not recognized.

Possible Cause

The host operating system may not have the appropriate drivers to support the installed Converged Network Adapter model.

Solution

Step 1 1) Obtain the following information:

a. Operating system of the host.

b. Specific model of installed CNA.

Step 2 Reference the appropriate vendor support page for the CNA model and host OS.

Step 3 Determine if an existing driver is already installed on the host OS.

Step 4 Ensure that the latest driver is installed from the CNA vendor support page or the host OS support page.

PFC

This section includes an overview of how to view standard pause frames, followed by a description of common problems and their solutions.

Standard pause frames

For ports with standard, non-CNA type host connections, the Nexus 5000 supports standard pause frames. These are enabled with the interface setting, as shown in the following example:

Example:

```
switch(config)# interface ethernet 1/16
switch(config-if)# flowcontrol ?
     receive  Receive pause frames
     send     Send pause frames
switch(config-if)# flowcontrol receive on
switch(config-if)# flowcontrol send on
```

To view standard pause frames, use the **show interface flowcontrol** command.

Example:

```
switch(config-if)# show interface flowcontrol

--------------------------------------------------------------------------------
Port      Send      FlowControl    Receive   FlowControl    RxPause   TxPause
          admin         oper       admin       oper
--------------------------------------------------------------------------------
Eth1/1    off           off        off         off          0         0
Eth1/2    off           off        off         off          0         0
Eth1/3    off           off        off         off          0         0
Eth1/4    off           off        off         off          0         0
Eth1/5    off           off        off         off          0         0
Eth1/6    off           off        off         off          0         0
Eth1/7    off           off        off         off          0         0
Eth1/8    off           off        off         off          0         0
Eth1/9    off           off        off         off          0         0
Eth1/10   off           off        off         off          0         0
Eth1/11   off           off        off         off          0         0
Eth1/12   off           off        off         off          0         0
Eth1/13   off           off        off         off          0         0
Eth1/14   off           off        off         off          0         0
Eth1/15   off           off        off         off          0         0
Eth1/16   on            on         on          on           0         0
Eth1/17   off           off        off         off          0         0
```

PFC not negotiated with FCOE-capable adapters (CNA)

Priority flow control (PFC) is not negotiated with FCOE-capable adapters (CNA).

This causes packet drop to be noticed on FCoE traffic from the servers.

Possible Causes

The CNA may not support DCBX and the PFC TLV is not negotiated.

Solution

Use the following information to verify DCBX support and that the PFC TLV is negotiated:

- Check the status of the PFC. Use the **show int ethx/x priority-flow-control** command. (Connected to CNA.)

  Example:

```
switch# show interface ethernet 1/13 priority-flow-control
===============================================================
Port               Mode Oper(VL bmap)  RxPPP      TxPPP
===============================================================

Ethernet1/13       Auto Off         0          0
```

- Check for LLDP neighbor or PFC/DCBX TLV advertised by the peer. Use the **show system internal dcbx info int ethx/x** command.

  Example:

```
switch(config-if)# show system internal dcbx info interface ethernet 1/1

Interface info for if_index: 0x1a000000(Eth1/1)
tx_enabled: FALSE
rx_enabled: FALSE
dcbx_enabled: TRUE
DCX Protocol: CIN


Port MAC address:  00:0d:ec:c9:c8:08


DCX Control FSM Variables: seq_no: 0x1, ack_no: 0x0,my_ack_no: 0x0, peer_seq_no:
 0x0 oper_version: 0x0,  max_version: 0x0 fast_retries 0x0

Lock Status: UNLOCKED
PORT STATE: UP
LLDP Neighbors
No DCX tlvs from the remote peer
```

- If the peer does not support DCBX, configure the priority-flow-control mode setting to on to enable PFC.

Switch Interface connected to CNA receives constant pause frames (PFC)

Constant pause frames (PFC) are received when the switch interface is connected to a CNA.

Possible Cause

If the Nexus 5000 switch is connected to a CNA, then the CNA might be sending Xon PFC frames to the switch. This increments pause counters when using the **show interface ethx/x** command.

To verify this situation, perform the following:

- For a few iterations, check using the **show interface ethx/x** command and make sure the pause frame count is incrementing using the **show interface ethx/x |grep - i pause** command.
- For a few iterations, **check using the show interface ethx/x** command and ensure that the PFC frame count is incrementing **using the show interface ethx/x priority-flow-control** command.
- For a few iterations, use the **show queuing interface ethx/x** command to check the pause status.

  Example:

```
Per-priority-pause status                 : Rx (Inactive), Tx (Inactive)
```

If the Rx (Inactive) and pause counter increment over time (as shown with the **show interface ethx/x priority-flow-control** command), then this indicates that the issue is due to Xon frames received from the CNA.

Possible Cause

If the Nexus 5000 switch is connected to a CNA along with slow servers that are not able to handle the traffic from the switch port, then the server sends Xoff pause frames to the switch to slow it down. This increments the pause counters when using the **show interface ethx/x** command.

To verify this situation, perform the following:

- For a few iterations, check using the **show interface ethx/x |grep - i pause** command and ensure that the pause frame count is incrementing.
- For a few iterations, check using the **show interface ethx/x priority-flow-control** command and ensure that the PFC frame count is incrementing.
- For a few iterations, use the **show queuing interface ethx/x** command and check the pause status.

  Example:

```
Per-priority-pause status                 : Rx (Active), Tx (Inactive)
```

If the Rx (Active) and pause counter increment (as shown with the **show interface ethx/x priority-flow-control** command), this indicates that the issue is due to Xoff frames received from the server.

Solution

Xoff pause frames from the server pause the Nexus 5000 interface and reduces the throughput from the switch to the CNA. On the server, investigate the OS/PCI slot to ensure that they are high-speed servers. Replace the servers that can run 10gb throughput.

Check if switch is sending pause frames or getting paused

FCoE throughput on servers is very low due to pause frames from the switch. It is then necessary to check if the switch is sending pause frames or if it is getting paused.

Possible Cause

If the egress FC port is congested, the switch sends PFC frames to the servers. The PFC frames are sent to reduce its FCoE rate and avoid a drop. If the server is slow or congested, the server sends PFC frames to the switch interface.

To verify this situation, perform the following:

- For a few iterations, check using the **show interface ethx/x |grep - i pause** command and ensure that the pause frame count (Rx/TX) is incrementing.
- For a few iterations, check using the **show interface ethx/x priority-flow-control** command and ensure that the PFC frame count (RX/TX) is incrementing.
- For a few iterations, check using the **show queuing interface ethx/x** command to check the pause status.

Note PFC frames are a MAC-level type of packet and cannot be viewed using the SPAN feature. Analyzer in-line is required to actually see the PFC frames on the wire.

Example:

```
Per-priority-pause status               : Rx (Active), Tx (Inactive)
```

If the Rx (Active) and pause RX counter increment (as shown with the **show interface ethx/x priority-flow-control** command), then this indicates that this issue is due to Xoff frames received from the server.

If the Tx(Active) and pause TX counter increment (as shown with the **show interface ethx/x priority-flow-control** command), this indicates that this issue is due to Xoff frames transmitted by the switch.

Solution

Identify the source of the congestion and try to resolve it by increasing the FC bandwidth or change it to a more powerful server. If congestion is expected, then Pause is expected for FCoE traffic.

Switch ports err-disabled due to pause rate-limit

Switch ports go into error-disable state due to pause rate limit

Possible Cause

If the switch interface receives excessive Xoff pause frames from the server, ports become error-disabled due to the high rate of pause frames received. Usually the port goes into an err-disable state due to pause frames, only if the drain rate is less than 5Mbps on a 10Gb port. This means that the server is very slow and is sending a large number of pause frames to the switch ports.

To verify this situation, use the **show interface ethernet 1/14 brief** command.

Example:

```
switch# show int eth1/14 brief

--------------------------------------------------------------------------------
Ethernet      VLAN   Type Mode   Status  Reason                   Speed     Port
Interface                                                                  Ch #
--------------------------------------------------------------------------------
Eth1/14   110    eth  trunk  down    pauseRateLimitErrDisable    100(D) 110
```

- Check if the RX pause count is a large value. Use the **show interface ethx/x** command to display the pause counters.
- Check for pause error-disable logs using the **show hardware internal gatos event-history errors |grep -i err** command.

Solution

Pause error-disable recovery can be enabled to get the ports out of this state, if the port is error-disabled due to transient condition as follows:

If the port is error-disabled due to transient condition listed below, then pause error-disable recovery can be enabled to move the ports out of this state.

- Error-disable recovery causes the pause rate limit.
- The error-disable recovery interval is 30.

If there is a consistent port error-disable condition due to the pause rate limit, determine if the issue is that the server is too slow. Replace the slow server.

How to enable link pause (flow control) on switch that connects DCBX capable devices

Link pause is not enabled on the switch ports that are connected to servers. It is necessary to enable link pause (flow control) on a Nexus 5000 switch that connects DCBX-capable devices.

Possible Cause

If the peer supports PFC TLV with DCBX, then configuring the flowcontrol send on and the flowcontrol receive on does not enable link pause. You have to disable PFC TLV sent by DCBX on the interface.

To verify this situation, perform one of the following:

- Check if the operating state is off using the **show interface ethx/y flowcontrol** command.
- Check if the operating state is on using the **show interface ethx/y priority-flow-control** command.

Solution

Use the following commands under the **interface ethx/y** command to enable link pause instead of PFC with DCBX capable devices:

- **no priority-flow-control mode on**
- **flowcontrol receive on**
- **flowcontrol send on**

How to clear PFC counters

How to clear priority flow counters.

Possible Cause

Currently there are no CLI commands to clear PFC frames (Bug ID is CSCtg08068).

Solution

Although there are no CLI commands to clear the PFC counters, a workaround does exist. You can the clear interface counters and then enter the **show interface ethx/x flowcontrol** command to see the PFC frame count.

Note The PFC frame count is incremented using the **show int ethx/x flowcontrol** command. This is a known bug.

Registers and Counters

Interface level errors

To view any interface level errors, use the **show interface counters errors** command.

Example:

```
switch# show interface counters errors

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

```
Port      Align-Err      FCS-Err   Xmit-Err  Rev-Err   Undersize     OutDiscards
-------------------------------------------------------------------------------
Eth1/1          0            0          0        0           0               0
Eth1/2          0            0          0        0           0               0
Eth1/3          0            0          0        0           0               0
Eth1/4          0            0          0        0           0               0
Eth1/5          0            0          0        0           0               0
Eth1/6          0            0          0        0           0               0
Eth1/7          0            0          0        0           0               0
Eth1/8          0            0          0        0           0               0
Eth1/9          0            0          0        0           0               0
Eth1/10         0            0          0        0           0               0
Eth1/11         0            0          0        0           0               0
Eth1/12         0            0          0        0           0               0
Eth1/13         0            0          0        0           0               0
Eth1/14         0            0          0        0           0               0
Eth1/15         0            0          0        0           0               0
Eth1/16         0            0          0        0           0               0
Eth1/17         0            0          0        0           0               0
Eth1/18         0            0          0        0           0               0
Eth1/19         0            0          0        0           0               0
Eth1/20         0            0          0        0           0               0
Eth2/1          0            0          0        0           0               0
Eth2/2          0            0          0        0           0               0
Eth2/3          0            0          0        0           0               0
Eth2/4          0            0          0        0           0               0
Po300           0            0          0        0           0               0
mgmt0          --           --         --       --          --              --
```

Packet byte counts

To view packet byte counts, use the **show interface counters detailed** command.

Example:

```
show interface ethernet  1/11 counters detailed

Ethernet 1/11
  Rx Packets:                                      430908
  Rx Unicast Packets:                              129965
  Rx Multicast Packets:                            300932
  Rx Broadcast Packets:                                11
  Rx Jumbo Packets:                                     3
  Rx Bytes:                                      41893521
  Rx Packets from 0 to 64 bytes:                       47
  Rx Packets from 65 to 127 bytes:                353478
  Rx Packets from 128 to 255 bytes:                60265
  Rx Packets from 256 to 511 bytes:                17095
  Rx Packets from 512 to 1023 bytes:                  16
  Rx Packets from 1024 to 1518 bytes:                  4
  Rx Trunk Packets:                                387901
  Tx Packets:                                      172983
  Tx Unicast Packets:                              129959
  Tx Multicast Packets:                             43024
  Tx Jumbo Packets:                                     3
  Tx Bytes:                                      18220330
  Tx Packets from 0 to 64 bytes:                        7
  Tx Packets from 65 to 127 bytes:                 112452
  Tx Packets from 128 to 255 bytes:                 60461
  Tx Packets from 256 to 511 bytes:                    40
  Tx Packets from 512 to 1023 bytes:                   19
```

```
Tx Packets from 1024 to 1518 bytes:                    1
Tx Trunk Packets:                                 130019
```

Verification of SNMP readouts

To view the verification of SNMP readouts, use the **sh interface ethernet 1/11 counters snmp** command.

Example:

```
switch# show interface ethernet 1/11 counters snmp

--------------------------------------------------------------------------------
Port               InOctets                    InUcastPkts
--------------------------------------------------------------------------------
Eth1/11            41908130                         130009

--------------------------------------------------------------------------------
Port               InMcastPkts                 InBcastPkts
--------------------------------------------------------------------------------
Eth1/11              301038                           11

--------------------------------------------------------------------------------
Port               OutOctets                   OutUcastPkts
--------------------------------------------------------------------------------
Eth1/11            18226503                         130003

--------------------------------------------------------------------------------
Port               OutMcastPkts                OutBcastPkts
--------------------------------------------------------------------------------
Eth1/11              43039                            0
```

Traffic rates

To view traffic rates, use the **show interface ethernet 1/11 counters brief** command.

Example:

```
switch# show interface ethernet 1/11 counters brief

-------------------------------------------------------------------------------------
Interface          Input Rate (avg)    Output Rate (avg)
                   ----------------    -----------------
                   Rate      Total     Rate      Total          Rate averaging
                   MB/s      Frames    MB/s      Frames         interval (seconds)
-------------------------------------------------------------------------------------
Ethernet 1/11      0         0         0         0              30
                   0         0         0         0              300
```