End-to-End FCoE
Design Guide

October 2014
1. Introduction to FCoE

Simply defined, I/O consolidation is the ability to carry different types of traffic with different traffic characteristics and handling requirements over the same physical media. The most difficult challenge of I/O consolidation is to satisfy the requirements of different traffic classes within a single network. Because Fibre Channel is the dominant storage protocol in the data center, any viable I/O consolidation solution for storage must allow for transparent integration of the Fibre Channel model. Fibre Channel over Ethernet (FCoE) meets this requirement in part by encapsulating each Fibre Channel frame inside an Ethernet frame. The goal of FCoE is to provide I/O consolidation over Ethernet, allowing Fibre Channel and Ethernet networks to share a single, integrated infrastructure, thereby reducing network complexities in the data center. FCoE consolidates both SANs and Ethernet traffic onto one converged network adapter (CNA), eliminating the need for using separate host bus adapters (HBAs) and network interface cards (NICs).

2. FCoE Technology Overview

- **CNA**: CNA is similar to an HBA or a NIC, but instead of handling either Fibre Channel or IP, the CNA can handle both simultaneously. The CNA presents separate networking and storage system interfaces to the operating system.

- **Fibre Channel Forwarder (FCF)**: The purpose of the FCF is to service login requests and provide the Fibre Channel services typically associated with a Fibre Channel switch. FCFs may also optionally provide the means to:
  - De-encapsulate Fibre Channel frames that are coming from the CNA and going to the SAN
  - Encapsulate Fibre Channel frames that are coming from the SAN and going to the CNA

- **Data Center Bridging (DCB)**: DCB is a collection of standards that extend classical Ethernet protocols for use in the data center.

- **DCB Exchange Protocol (DCBX)**: DCBX is a discovery and capability exchange protocol used for conveying capabilities and configuration of the features listed previously between neighbors to ensure consistent configuration across the network. This protocol is expected to use functions provided by IEEE 802.1AB (Link Layer Discovery Protocol [LLDP]).

- **Priority-Based Flow Control (PFC)**: IEEE 802.1Qbb provides a link-level flow-control mechanism that can be controlled independently for each class of service (CoS), as defined by IEEE 802.1p. The goal of this mechanism is to ensure zero loss under congestion in DCB networks.

- **Enhanced Transmission Selection (ETS)**: This feature, defined in the IEEE 802.1Qaz standard, enables a percentage of available bandwidth on a link to be divided among specified priority groups.

- **FCoE Initialization Protocol (FIP)**: FIP is used to perform device discovery, initialization, and link maintenance. FIP performs the following protocols:
  - FIP Discovery - when an FCoE device is connected to the fabric, it sends out a discovery solicitation message. A Fibre Channel Forwarder (FCF) or a switch responds to the message with a solicited advertisement that provides an FCF MAC address to use for subsequent logins.
  - FCoE Virtual Link instantiation: FIP defines the encapsulation of fabric login (FLOGI), fabric discovery (FDISC), logout (LOGO), and exchange link parameters (ELP) frames along with the corresponding reply frames. The FCoE devices use these messages to perform a fabric login.
  - FCoE Virtual Link maintenance: FIP periodically sends maintenance messages between the switch and the CNA to ensure the connection is still valid. This message is referred to as the frame keepalive (FKA).
• FIP Snooping Bridge: A FIP Snooping Bridge is an Ethernet bridge that supports:
  ◦ Priority Flow Control (PFC; IEEE 802.1Qbb)
  ◦ Enhanced Transmission Selection (ETS; IEEE 802.1Qaz)
  ◦ DCBX (IEEE 802.1Qaz)

For more information and resources for FCoE, please visit: https://cisco.com/go/fcoe.

3. Hardware and Software Requirements

Hardware requirements vary based on the platform used for FCoE connectivity. The platform-specific requirements to build the multihop topologies discussed in this document are outlined in Table 1. Note that a multihop FCoE topology also requires CNAs for server connectivity.

Table 1. Platform-Specific Requirements

<table>
<thead>
<tr>
<th>Platform</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cisco Nexus® 7700 Switch</td>
<td>F2e and F3’ cards</td>
</tr>
<tr>
<td>Cisco Nexus 7000 Series Switches</td>
<td>F1, F2, F2e, and F3’ cards</td>
</tr>
<tr>
<td>Cisco Nexus 6000 Series Switches</td>
<td>Up to 384 10-Gbps ports or 96 40-Gbps ports</td>
</tr>
<tr>
<td>Cisco Nexus 5000 Series Switches</td>
<td>Up to 128 10-Gbps ports</td>
</tr>
<tr>
<td>Cisco MDS 9700 Series Multilayer Directors</td>
<td>48 port 10-Gbps FCoE module</td>
</tr>
<tr>
<td>Cisco MDS 9250i Multiservice Fabric Switch</td>
<td>8 ports of 10-Gbps FCoE</td>
</tr>
</tbody>
</table>

* F3 cards have hardware capability of supporting FCoE. Software enablement is in the roadmap.

Table 2 lists software requirements.

Table 2. Platform-Specific Software Requirements

<table>
<thead>
<tr>
<th>Product</th>
<th>License</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cisco Nexus 7000 Series Switches</td>
<td>Each F Series line card that runs FCoE requires an FCoE license. For a complete explanation of the Cisco NX-OS licensing scheme and how to obtain and apply licenses, refer to the Cisco NX-OS Licensing Guide.</td>
</tr>
<tr>
<td>Cisco MDS 9700 Series Multilayer Directors</td>
<td>FCoE does not require a license. For a complete explanation of the Cisco NX-OS licensing scheme and how to obtain and apply licenses, refer to the Cisco NX-OS Licensing Guide.</td>
</tr>
<tr>
<td>Cisco Nexus 6000 Series Switches</td>
<td>FCoE capability is included in the storage protocol services license (FC_FEATURES_PKG).</td>
</tr>
<tr>
<td>Cisco Nexus 5000 Series Switches</td>
<td>FCoE capability is included in the storage protocol services license (FC_FEATURES_PKG).</td>
</tr>
</tbody>
</table>

Guidelines and limitations of FCoE follow:

- You cannot enable FCoE on default VLAN.
- The quality-of-service (QoS) policy must be the same on all Cisco FCoE switches in the network.
- Beginning with Cisco NX-OS Release 6.1, FCoE is supported on F2 and F2e Series modules.
  - FCoE supports only F2e (Small Form-Factor Pluggable Plus [SFP+]) modules.
  - FCoE does not support F2e (copper) modules.

FCoE in a dedicated storage virtual device context (VDC) has the following guidelines:

- Enable the FCoE feature set in only one VDC.
- Create VLANs in the FCoE-allocated VLAN range.
- Do not enable any features other than storage-related features in the dedicated FCoE VDC.
- Allocate resources for the dedicated FCoE VDC from an F Series module, such as the 32-port 1- and 10-Gigabit Ethernet I/O module (PID N7K-F132XP-15).
- Rollback is not supported in a storage VDC.
- FCoE support on F2 and F2e Series modules (SF248XP-25) requires a Supervisor 2 module.
- F2 and F2e Series modules cannot exist in the same VDC with any other module type. This stipulation applies to both LAN and storage VDCs.

4. Data Center Fabric Design Requirements

4.1 High Availability

The best data center design should be able to sustain failures in the network and work without affecting the production traffic.

High availability is a crucial requirement in any data center design. It is especially important in a converged network, an environment in which loss-intolerant storage traffic shares the network infrastructure with loss-tolerant Ethernet traffic.

High availability is implemented at many different levels of the network:
- Chassis level
- Link level
- Process level

Cisco Nexus 7000 Series and Cisco MDS 9700 Series switches have hardware and software features that help network designers meet the high-availability requirements of their data centers and secure the uptime and integrity of the storage traffic in a consolidated environment.

Traditionally, SAN designers have used redundant network links and network equipment to create high availability in their networks. These designs provision and manage parallel, disparate SANs. These parallel, independent SANs are often referred to as SAN A and SAN B.

Data center designers need to consider high-availability options when designing a consolidated network.


4.2 Oversubscription

Oversubscription is the practice of having multiple devices share a common resource. In networks, bandwidth is commonly oversubscribed by using a set of end devices that have a greater aggregate bandwidth than the links and end devices through which they are connecting. This setup is often described in terms of an oversubscription ratio, which is the ratio between the bandwidth available to one set of ports, devices, and links and another set of ports, devices, and links.

Oversubscription ratio is of two types:
- Host bandwidth to storage bandwidth
- Edge Inter-Switch Link (ISL) bandwidth to core ISL bandwidth
The amount of oversubscription on a link depends on the I/O capabilities of the hosts and targets and the applications using them. Designs with more highly oversubscribed links may be satisfactory for applications with minimal I/O requirements, and conversely links with lower oversubscription ratios may be required for applications that have more I/O requirements.

Oversubscribing the network allows higher port-count deployments without compromising performance, because the performance requirements of the application may be far less than those of the negotiated port speed of the switch or host. Network designers should keep the requirements of their applications in mind when designing end-to-end FCoE networks.

4.3 Scalability

Data centers are growing rapidly, and the need for the new devices in a SAN is also growing. Scalability defines the ability to expand the networks on demand without being able to change the entire design. It is important to consider the scalability options during the network design so that the costs can be reduced during additional device connections.

Scalability options can be explored by checking the:

- Port density on switches
- Capability of workload balance
- Ability to adapt to changing speeds and architecture without being able to change the whole design
- Ability to accommodate the increase in the number of FLOGI and zones

4.4 SAN A and SAN B Separation

SAN A and SAN B separation in traditional Fibre Channel SANs makes sure that the design is highly available even when one of the fabrics goes down. In end-to-end FCoE designs, SAN A and SAN B separation can be achieved physically or logically by separating the FCoE VLANs that are carried across the LAN.

The FCoE VLANs synonymous to VLANs in traditional LANs can be used as a mechanism to separate the SAN fabrics from each other. Many data center architectures allow this fabric separation by not using the same VLAN across the switches carrying FCoE traffic.

4.5 Switch Considerations

Table 3 lists some of the parameters designers should consider.

<table>
<thead>
<tr>
<th>Parameters to Consider for Design</th>
<th>Cisco Nexus 7000</th>
<th>Cisco Nexus 6000</th>
<th>Cisco Nexus 5000</th>
<th>Cisco MDS 9700 and MDS 9500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resiliency</td>
<td>Highest (In-Service Software Upgrade [ISSU], management redundancy, switch fabric redundancy, and power supplies)</td>
<td>High (Power-supplies redundancy)</td>
<td>High (Power-supplies redundancy)</td>
<td>Highest (ISSU, management redundancy, switch fabric redundancy, and power-supplies redundancy)</td>
</tr>
<tr>
<td>Protocol flexibility</td>
<td>FCoE, iSCSI and NFS</td>
<td>Fibre Channel, FCoE, iSCSI, and NFS</td>
<td>Fibre Channel, FCoE, iSCSI, and NFS</td>
<td>Fibre Channel, FCoE, iSCSI, Fibre Channel over TCP/IP (FCIP), and storage services</td>
</tr>
<tr>
<td>High availability</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Port speed</td>
<td>10 and 40 Gb</td>
<td>10 and 40 Gb</td>
<td>10 and 40 Gb</td>
<td>10 Gb</td>
</tr>
</tbody>
</table>
5. Next-Generation Data Center Architecture

The data center design depends on all of the previously mentioned factors along with the application requirements. Following are the various network topologies synonymous with the traditional Fibre Channel SAN designs but running FCoE end to end.

All of the Cisco devices such as the Cisco Nexus 7000, Nexus 5000, Nexus 6000, and Cisco MDS Multilayer Directors support the topologies listed as follows. Figure 1 shows dedicated FCoE links from access to the core and to the edge.

### 5.1 Core-Edge

This design has the hosts connected to an edge device, which in turn is connected to the core switches. The core switches have the storage devices connected to them. The high availability can be achieved by using director-class switches such as the Cisco Nexus 7000 and Cisco MDS Multilayer Directors at the edge and core level. Deploying Cisco Nexus 6000 at the edge level would also give higher scalability options and provide flexibility to use the 40-Gb links in the future to connect to the core switches.

**Figure 1.** Dedicated FCoE Links from Access to Core and Edge

<table>
<thead>
<tr>
<th>Parameters to Consider for Design</th>
<th>Cisco Nexus 7000</th>
<th>Cisco Nexus 6000</th>
<th>Cisco Nexus 5000</th>
<th>Cisco MDS 9700 and MDS 9500</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Use cases</strong></td>
<td>LAN-SAN convergence and consolidation</td>
<td>LAN-SAN convergence and consolidation</td>
<td>LAN-SAN convergence with Fibre Channel bridging capabilities</td>
<td>LAN-SAN convergence and Business Continuity/Data Replication (BC/DR)</td>
</tr>
<tr>
<td><strong>Places in network</strong></td>
<td>FCoE access and core</td>
<td>FCoE access and core</td>
<td>FCoE access</td>
<td>Fibre Channel, FCoE access, and core</td>
</tr>
</tbody>
</table>
In the figure, the edge switch considerations include the Cisco Nexus 5000, Nexus 6000, and Nexus 7000. Core switch considerations include the Cisco Nexus 7000 and Cisco MDS 9000 Series.

The links from the edge and core can carry the LAN and SAN converged traffic. The host links connected to the edge switches would be converged and SAN A and SAN B separation can be achieved using different VLANs at the edge level.

This design suits the enterprise customers the best with the edge switches providing the scalability option and the core being the high-availability situation. An oversubscription ratio of 7:1 can be achieved by using Cisco Nexus 7000 Series Switches at the edge and core.

5.2 Collapsed Core

The collapsed-core design usually has a single core switch with both the storage and hosts connected to it. This single-hop design provides much less latency because both the hosts and storage are connected to the same switch. The scalability is much less compared to the other designs. The oversubscription ratio here depends only on the host-to-storage connectivity because all of them are connected to the single switch. The Cisco Nexus 7000 provides the high availability and scalability in terms of line cards, and it best suits the design (Figure 2).

Figure 2. Collapsed Core

Core switches considered include the Cisco Nexus 7000 and Nexus 6000 Series Switches.

This design best suits the small-scale data centers and the applications with minimal latency requirements. Scalability can be achieved by increasing the number of cores or connecting the hosts to edge switches.

5.3 Edge-Core-Edge

The edge-core-edge design has multiple hops from the host to the target and is a highly scalable solution. The hosts are connected to the storage through a set of edge and core switches on both ends. The solution provides flexibility in terms of the latency design and oversubscription considerations.
Cisco Nexus 7000 Switches at the core and storage edge provides higher bandwidth for ISLs and storage devices. The high availability is achieved both at the chassis level and the link level using Cisco Nexus 7000 Switches (Figure 3).

**Figure 3.** Edge-Core-Edge

In the figure, edge switch considerations include the Cisco Nexus 5000 and Nexus 6000 Switches. Core switch considerations include the Cisco Nexus 7000, and storage edge switch considerations include the Cisco Nexus 7000 and the Cisco MDS Multilayer Directors. The design best suits the enterprise data centers that provide options for scalability and high availability. An oversubscription ratio of 7:1 can be achieved using the Cisco Nexus 7000 Switches at the core and storage edge.
6. Design Consideration for Multihop FCoE with Edge-Core-Edge Design

Figure 4 shows the design considerations for the Multihop FCoE with the edge-core-edge network topology.

Figure 4. Multihop FCoE with Edge-Core-Edge Network Topology

6.1 Access Layer and SAN Edge Fabric

The access layer in a traditional LAN has the hosts connected to the access switches. The CNAs can do both Ethernet and FCoE. The converged traffic passes to the access layer switches, which should be capable of processing the Ethernet and LAN traffic.

It is important to consider the scalability option here for both the Ethernet and storage traffic. All the hosts in a rack can be connected to top-of-rack switches, and multiple rows can be connected to an end-of-row switch.

Cisco Nexus 2000 Series Switches are used as the top-of-rack switches, which can be used as fabric extenders to connect to the end-of-row switches. These devices help manage the fabrics without much hassle and reduce the cabling management. Cisco Nexus 2000 devices can be connected to either Cisco Nexus 7000 or Nexus 5000 parent switches.
The virtual port channel (vPC) feature on the parent switches provides a possibility of dual-homed FEX, which can be connected to two different end-of-row switches. The hosts connected to the FEX can be configured in a port channel for the LAN configuration in an enhanced vPC scenario, but for FCoE the traffic would pass through a single FEX device based on FCoE VLAN separation. The SAN A and SAN B separation is achieved using different FCoE VLANs on the parent switches.

The Cisco Nexus 6000 and Nexus 5000 Switches are the best designs for the access layer switches because they provide a highly scalable solution. The Cisco Nexus 6000 Switches offer 40-Gb ports that can be connected to the fabric core switches (Figure 5).

**Figure 5.** Access Layer and SAN Edge Fabric

![Access Layer and SAN Edge Fabric](image)

6.2 Aggregation Layer and SAN Core Fabric

This segment of the fabric is where the LAN and SAN or FCoE fabric gets separated. The design should incorporate switches that provide high availability and highly scalable fabrics, considering the oversubscription ratio on the number of ISLs to the edge fabrics.

The Cisco Nexus 7000 and Nexus 7700 Series Switches provide CPU, and process level high-availability options along with the scalability options with multiple modules running FCoE.
The Cisco Nexus 7000 Switches will provide 40- and 100-Gb connectivity options to the edge fabrics in the future; this connectivity will increase the throughput and reduce the latency. The 40-Gb links can be split into four 10-Gb links when connecting to the edge switches (Figure 6).

**Figure 6.** Aggregation Layer and SAN Fabric Core

The data spine in Figure 6 is a fabric path or a Cisco Dynamic Fabric Automation (DFA) link going to the LAN core.

### 6.3 Storage Edge Fabric

The storage edge fabrics have the storage arrays connected to them. Hence the high availability and bandwidth are very important factors that should be considered when designing the solution.
Cisco Nexus 7000 Series Switches and Cisco MDS Multilayer Directors provide highly available connections to the core switches. Cisco Nexus 7000 Switches also provide 40- and 100-Gb connections to the core switches, thus providing high-bandwidth solutions (Figure 7).

**Figure 7.** Storage Edge Fabric

---

6.4 Best Practices

Following are a few best practices that should be considered during the fabric design using Cisco Nexus switches:

- Configure a unique dedicated VLAN at every converged access switch to carry traffic for each Cisco Virtual SAN (VSAN) in the SAN (for example, VLAN 1002 for VSAN 1, VLAN 1003 for VSAN 2, and so on).
- Configure the unified fabric links as trunk ports. Do not configure the FCoE VLAN as a native VLAN.
- Configure all FCoE VLANs as members of the unified-fabric links to allow extensions for VF_Port trunking and VSAN management for the virtual Fibre Channel interfaces.
- Configure the unified-fabric links as spanning-tree edge ports.
- Do not configure the FCoE VLANs as members of Ethernet links that are not designated to carry FCoE traffic because you want to ensure that the scope of the shielded twisted pair (STP) for the FCoE VLANs is limited to unified-fabric links only.
- If the converged access switches (in the same SAN fabric or in another) need to be connected to each other over Ethernet links for a LAN alternate path, then such links must explicitly be configured to exclude all FCoE VLANs from membership. This action helps ensure that the scope of the STP for the FCoE VLANs is limited to unified-fabric links only.
- Use separate FCoE VLANs for FCoE in SAN A and SAN B.
- Make sure the QoS policies for the no-drop class are same throughout the fabric.

You can prevent data corruption due to cross-fabric talk by configuring an FC-Map, which identifies the Fibre Channel fabric for this Cisco Nexus device. When the FC-Map is configured, the switch discards the MAC addresses that are not part of the current fabric.
7. FCoE Scalability Considerations

The scalability options account for the supported number of fabric logins and the zoning database in the entire fabric. When designing the fabrics, it is important to consider the values mentioned previously to be able to accommodate the growing needs of the data centers.

- Fabric logins (FLOGIs): All Fibre Channel devices log in to the fabric and get a unique ID (Fibre Channel ID [FCID]) to be able to communicate with other devices. The limitation on the number of devices that can log in to a switch limits the scalability of the fabric irrespective of the core and edge designs.

- Zoning: Zoning in a traditional Fibre Channel fabric provides security by avoiding crosstalk between hosts and restricting access to the targets for unauthorized hosts.

The limitation on the size of the zone database for the fabric depends on the zone database that each switch has. The fabric can have only the zone database size of the smaller switch in the fabric. It is important to consider the database size and the number of zones allowed when considering fabric scalability.

Table 4 lists the verified scalability limits.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Cisco Nexus 7000 (Cisco NX-OS 7.1)</th>
<th>Cisco MDS 9700 with 48-Port 10-GBps FCoE Module (Cisco NX-OS 6.2(9))</th>
<th>Cisco Nexus 5000 and Nexus 6000 (Cisco NX-OS 7.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of fabric logins per switch</td>
<td>4,000</td>
<td>4,000</td>
<td>4,000</td>
</tr>
<tr>
<td>Number of fabric logins per line card</td>
<td>500</td>
<td>1000</td>
<td>-</td>
</tr>
<tr>
<td>Number of fabric logins per port</td>
<td>256</td>
<td>256</td>
<td>512 per port group</td>
</tr>
<tr>
<td>Number of FCoE hops</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Number of zone members per fabric</td>
<td>30,000</td>
<td>30,000</td>
<td>16,000</td>
</tr>
<tr>
<td>Number of zones per fabric</td>
<td>16,000</td>
<td>16,000</td>
<td>8,000</td>
</tr>
<tr>
<td>Number of zone sets per switch</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
</tbody>
</table>

For more details please visit


8. Sample Configuration

Sample configuration of the edge-core-edge topology follows.

8.1 Cisco Nexus 6000 Series Access Layer Configuration with FEX and EvPC

Figure 8 shows connections between the FEX and access layer switch.

**Figure 8.** Access Layer Configuration

VSAN 100 on SAN A with FCoE VLAN 100 and VSAN 200 on SAN B with FCoE VLAN 200.

Configuration commands on the Cisco Nexus 6000 SAN A fabric follow:

```
Nexus6000-1:

conf t
feature fcoe
feature lldp
feature vpc
feature lacp
feature fex

vsan database
vsan 100
exit

vlan 100
fcoe vsan 100
exit
```

© 2014 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.
vpc domain 1
peer-keepalive destination <ip-add of N6K-2> source <ip-add of N6K-1>
exit
!
int eth1/3-4
channel-group 1 mode active
no shut
!
int po1
switchport
vpc peer-link
!
int eth 1/1
channel-group 100
no shut
!
int eth1/2
channel-group 101
no shut
!
int po100
vpc 100
switchport mode fex-fabric
fex associate 100
no shut
!
int po101
vpc 101
switchport mode fex-fabric
fex associate 101
no shut
!
Int eth1/15
switchport mode trunk
switchport trunk allowed vlan 100
no shut
exit
!
Int eth1/16
Switchport mode trunk
Switchport trunk allowed vlan 100
No shut
Exit
!
int eth 100/1/1
channel group 2000 mode on
no shut
!
int eth101/1/1
channel group 2000 mode on
no shut
!
int po2000
switchport mode trunk
switchport trunk allowed vlan 100
spanning-tree port type edge network
no shut
!
fex 100
fcoe
exit
!
int vfc100
bind interface eth100/1/1
switchport mode F
switchport trunk allowed vsan 100
no shut
!
int vfc10
bind interface eth1/15
switchport mode E
switchport trunk allowed vsan 100
no shut
!
int vfc20
Bind interface eth1/16
Switchport mode E
Switchport trunk allowed vsan 100
No shut
Exit
!
Configuration commands on the Cisco Nexus 6000 SAN B follow:

```
Nexus6000-2##

conf t
feature fcoe
feature lldp
feature vpc
feature lacp
feature fex
!
vsan database
vsan 200
exit
!
vlan 200
fcoe vsan 200
exit
!
vpc domain 1
peer-keepalive destination <ip-add of N6K-1> source <ip-add of N6K-2>
exit
!
int eth1/3-4
channel-group 1 mode active
no shut
!
int po1
switchport
vpc peer-link
!
int eth 1/1
channel-group 100
no shut
!
int eth1/2
channel-group 101
no shut
!
int po100
vpc 100
```
switchport mode fex-fabric
fex associate 100
no shut
!
int po101
vpc 101
switchport mode fex-fabric
fex associate 101
no shut
!
int eth 1/15
switchport mode trunk
switchport trunk allowed vlan 200
no shut
exit
!
Int eth1/16
Switchport mode trunk
Switchport trunk allowed vlan 200
No shut
Exit
!
int eth 100/1/1
channel group 2000 mode on
no shut
!
int eth101/1/1
channel group 2000 mode on
no shut
!
int po2000
switchport mode trunk
switchport trunk allowed vlan 200
spanning-tree port type edge network
no shut
!
fex 101
fcoe
exit
!
int vfc200
bind interface eth101/1/1
switchport mode F
switchport trunk allowed vsan 200
no shut
!
int vfc10
bind interface eth1/15
switchport mode E
switchport trunk allowed vsan 200
no shut
!
Int vfc20
Bind interface eth1/16
Switchport mode E
Switchport trunk allowed vsan 200
No shut
!

8.2 Cisco Nexus 7700 Core Configuration

Figure 9 shows connections between access layer and SAN fabric core switches.

Figure 9. SAN Fabric Core
The configuration for the Cisco Nexus 7000-1 follows:

```plaintext
!!Storage VDC creation in Admin VDC

conf t
license fcoe module x !!license all fcoe modules!!
!
system qos
service-policy type network-qos default-nq-7e-policy
exit
!
vdc storage type storage
allocate interface eth1/1-4,1/15-16
allocate fcoe-vlan-range 100
exit
!

!!On Storage VDC of N7K-1!!

conf
feature lacp
!
vlan 100
fcoe vsan 100
exit
!
vsan database
vsan 100
exit
!
int eth 1/15
switchport mode trunk
switchport trunk allowed vlan 100
no shut
!
int eth1/1-2
channel-group 30 mode active
no shut
exit
!```
int eth1/3-4
channel-group 40 mode active
no shut
exit
!
int po30
switchport mode trunk
switchport trunk allowed vlan 100
no shut
!
int po40
switchport mode trunk
switchport trunk allowed vlan 100
no shut
!
int vfc10
bind interface ethernet 1/15
switchport mode E
switchport trunk allowed vsan 100
no shut
!
int vfc30
bind interface port-channel 30
switchport mode E
switchport trunk allowed vsan 100
no shut
!
int vfc40
bind interface port-channel 40
switchport mode E
switchport trunk allowed vsan 100
no shut
!

Configuration for the Cisco Nexus 7000-2 SAN A is similar. Replace interface Ethernet 1/15 with Ethernet 1/16.

Configuration commands are shown for the SAN A fabric. SAN B fabric configuration remains the same except for the VLAN and VSAN. Replace VLAN 100 with VLAN 200 and VSAN 100 with VSAN 200 for fabric B.
8.3 Cisco Nexus 7700 Storage Edge Fabric Configuration

Figure 10 shows connections between SAN Fabric Core and Storage Edge fabric.

**Figure 10.** SAN Fabric Core and Storage Edge Fabric

Configuration for the Cisco Nexus 7000-3 follows:

```plaintext
!!Storage VDC creation in Admin VDC

conf t
license fcoe module x !!license all fcoe modules!!
!
system qos
service-policy type network-qos default-nq-7e-policy
exit
!
vdc storage type storage
allocate interface eth1/1-4,1/11-12
allocate fcoe-vlan-range 100
```
exit
!

!!On Storage VDC of N7K-3!!

conf
feature lacp
!
vlan 100
fcoe vsan 100
exit
!
vsan database
vsan 100
!
int eth 1/11
switchport mode trunk
switchport trunk allowed vlan 100
no shut
!
int eth1/1-2
channel-group 30 mode active
no shut
exit
!
int eth1/3-4
channel-group 50 mode active
no shut
exit
!
int pc30
switchport mode trunk
switchport trunk allowed vlan 100
no shut
!
int pc50
switchport mode trunk
switchport trunk allowed vlan 100
no shut
!
Configuration for the Cisco Nexus 7000-4 SAN A is similar. Replace the interface numbers and port channels according to the design shown in Figure 10.

Configuration commands are shown for the SAN A fabric. SAN B fabric configuration remains the same except for the VLAN and VSAN. Replace VLAN 100 with VLAN 200 and VSAN 100 with VSAN 200 for fabric B.

8.4 Configuration Using DCNM

Cisco Prime™ Data Center Network Manager (DCNM) provides an easy way to configure FCoE on the Cisco Nexus 7000, Nexus 5000, and Nexus 6000 Series and Cisco MDS Switches.

The Cisco Prime DCNM SAN provides management tools to monitor the performance and maintain the operational links.

9. Managing the Fabric

9.1 Single Best Status Command for FCoE Interface Status

This simple command can be run on a questionable interface that allows a quick check that FCoE is configured properly between the CNA and the switch. The output indicates the vfc interface is completely in the “up” state showing the end device has performed a FLOGI (successfully went through the FIP process). An output similar to the following normally means any storage problem is probably not network-related:

```
N6K# show interface ethernet 101/1/1 fcoe
Ethernet101/1/1 is FCoE UP
  vfc1 is Up
  FCID is 0xd10000
  FWNN is 20:00:74:26:ac:17:2a:b1
  MAC addr is 74:26:ac:17:2a:b1
```

To be certain there are no other underlying network problems, check the output of the show interface Ethernet 101/1/1 command to ensure that no discards are occurring. This command is also valid on the Cisco Nexus 7000.

9.2 Monitoring Priority Flow Control

Pause frames issued through PFC may be an indication of performance problems at the end device. To check the status of PFC at a glance on all interfaces on a switch, use the following command:

```
N6K# show interface priority-flow-control
```

<table>
<thead>
<tr>
<th>Port</th>
<th>Mode</th>
<th>Oper (VL bmap)</th>
<th>RxPPP</th>
<th>TxPPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethernet1/1</td>
<td>Auto On</td>
<td>(8)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ethernet1/2</td>
<td>Auto On</td>
<td>(8)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ethernet1/3</td>
<td>Auto Off</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ethernet101/1/1</td>
<td>Auto On</td>
<td>(8)</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

Note the interfaces in that command that have a VL bmap of 8 are FCoE-enabled interfaces. A large count of TxPPP may indicate a host that is busy or experiencing congestion with FCoE traffic and/or data traffic. Interfaces that have a large number of RxPPPs would indicate they are the recipient of the Tx pause frames. There are also detailed Rx and Tx pause counters in the output of show interface eth 101/1/1.

9.3 Verifying FEX FCoE Configuration

Verify that FEX is configured properly to support FCoE:

```
N6K# show running-config fex
```

```
feature fex

fex 101
  pinning max-links 1
```
Note in that configuration that when using vPC with dual-attached FEXs, only one of the FEXs can be FCoE-enabled from the perspective of the respective switch (in the configuration, FEX 101 is the designated FCoE switch).

Check the status of the FEX:

N6K# show fex detail
FEX: 101 Description: FEX0101 state: Online
  FEX version: 6.0(2)N2(3) [Switch version: 6.0(2)N2(3)]
  FEX Interim version: 6.0(2)N2(3)
  Switch Interim version: 6.0(2)N2(3)
  Extender Serial: SS160309DX
  Extender Model: N2K-C2232PP-10GE, Part No: 73-12533-05
  Card Id: 82, Mac Addr: c8:f9:f9:20:c1:02, Num Macs: 64
  Module Sw Gen: 12594 [Switch Sw Gen: 21]
  post level: complete
  pinning-mode: static Max-links: 1
  Fabric port for control traffic: Eth1/3
  FCoE Admin: true
  FCoE Oper: true
  FCoE FEX AA Configured: false
  Fabric interface state:
    Eth1/3 - Interface Up. State: Active
  Fex Port State Fabric Port
  Eth101/1/1 Up Eth1/3
  Eth101/1/2 Down None

In the case of second-level vPC (EvPC), ensure the physical interface of vPC is bound to the vfc and not the port channel.
9.4 Verifying the FC-MAP Values of the Fabric

The FC-MAP value identifies the Fibre Channel fabric for the Cisco Nexus device. We can prevent data corruption due to cross-fabric talk by configuring an FC-Map. The switch discards the MAC addresses that are not part of the current fabric.

The default FC-MAP value is 0e:fc:00. This value can be changed on the Cisco Nexus devices for fabric separation.

FC-MAP is verified using the `show fcoe` command:

```
N6K# show fcoe
Global FCF details
   FCF-MAC is 00:2a:6a:1d:f9:80
   FC-MAP is 0e:fc:00
   FCF Priority is 128
   FKA Advertisement period for FCF is 20 seconds
```

Change the FC-MAP using the following command:

```
N6K# conf t
   (config)# fcoe fcmap ?
   <0x0e00-0x0eff> Enter FCMAP
```

9.5 Verifying Ethernet Interface Status

After verifying physical connectivity and putting the interface in the admin up state (no shut), check the following items to determine the problem:

```
! Verify configuration of interface
interface Ethernet101/1/1
   switchport mode trunk
   switchport trunk native vlan 10
   switchport trunk allowed vlan 10,101
   spanning-tree port type edge
   no shutdown

! Verify Ethernet interface show output. Verify VLAN allowed list:
N6K# show interface ethernet 101/1/1 switchport
Name: Ethernet101/1/1
   Switchport: Enabled
   Switchport Monitor: Not enabled
   Operational Mode: trunk
   Access Mode VLAN: 1 (default)
   Trunking Native Mode VLAN: 10 (native)
   Trunking VLANs Allowed: 10,101
```

Switchport mode trunk and native VLAN configured.

Allowed list correct, including FCoE VLAN. STP port type is edge.
Voice VLAN: none
Extended Trust State: not trusted [COS = 0]
Administrative private-vlan primary host-association: none
Administrative private-vlan secondary host-association: none
Administrative private-vlan primary mapping: none
Administrative private-vlan secondary mapping: none
Administrative private-vlan trunk native VLAN: none
Administrative private-vlan trunk encapsulation: dot1q
Administrative private-vlan trunk normal VLANs: none
Administrative private-vlan trunk private VLANs: none(0 none)
Operational private-vlan: none
Unknown unicast blocked: disabled
Unknown multicast blocked: disabled

If the previous configuration is not in the expected state, start looking at the following internal event history to figure out if any failure has happened. The normal up state follows:

ETH_PORT_FSM_ST_TRUNK_UP

N6K# show system internal ethpm event-history interface ethernet 101/1/1 |
include "Curr state"

Curr state: [ETH_PORT_FSM_ST_TRUNK_UP]

Some of the states that may be reported other than the normal up state follow:

The finite state machine (FSM) might be in one of the following states:

ETH_PORT_FSM_ST_NOT_INIT
ETH_PORT_FSM_ST_DOWN
ETH_PORT_FSM_ST_INIT_EVAL
ETH_PORT_FSM_ST_SPAN_EVAL
ETH.PORT_FSM.ST_WAIT_PRE_CFG
ETH_PORT_FSM_ST_LINK_INIT
ETH_PORT_FSM_ST_WAIT_BRINGUP
ETH.PORT_FSM.ST_WAIT_LOGICAL_UP
ETH_PORT_FSM_ST_L2_UP
ETH_PORT_FSM_ST_L3_UP
ETH_PORT_FSM_ST_PROTOCOL_DOWN
ETH_PORT_FSM_ST_SPAN_DEST_UP
ETH_PORT_FSM_ST_WAIT_PROTOCOL_DOWN
ETH_PORT_FSM_ST_WAIT_PHYSICAL_DOWN
ETH_PORT_FSM_ST_WAIT_APPLY_CONFIG
ETH_PORT_FSM_ST_WAIT_LOGICAL_DOWN
ETH_PORT_FSM_ST_WAIT_LOGICAL_CHANGE_TRUNK
ETH_PORT_FSM_ST_NOT_UP
ETH_PORT_FSM_ST_BUNDLE_MEMBER_UP
ETH_PORT_FSM_ST_WAIT_BUNDLE_PRE_CFG
ETH_PORT_FSM_ST_WAIT_BUNDLE_LOGICAL_UP
ETH_PORT_FSM_ST_WAIT_BUNDLE_LOGICAL_DOWN
ETH_PORT_FSM_ST_WAIT_BUNDLE_MEMBER_DOWN
ETH_PORT_FSM_ST_ERROR_DISABLED_LEVEL_1
ETH_PORT_FSM_ST_ERROR_DISABLED_LEVEL_2
ETH_PORT_FSM_ST_AUTH_FAIL
ETH_PORT_FSM_ST_WAIT_LOGICAL_DOWN_RNF 30
ETH_PORT_FSM_ST_WAIT_PROTOCOL_DOWN_RNF 31

It is possible that a sequence error may have occurred, which would be logged to syslog in a message similar to the following on the console:

```
2014 Jun 22 15:01:37 DCE-1 VDC-1 ETHPORT-2-IF_SEQ_ERROR: Error ("sequence timeout") while communicating with component MTS_SAP_ETH_PORT for opcode MTS_OPC_ETHPM_PORT_BRINGUP (RID_PORT: Ethernet1/16)
```

```
2014 Jun 22 06:39:11 dist-B VDC-1 ETHPORT-2-IF_DOWN_ERROR_DISABLED: Interface Ethernet1/3 is down (Error disabled. Reason:sequence timeout)
```

```
2014 Jun 22 06:39:51 dist-B VDC-1 ETHPORT-2-IF_SEQ_ERROR: Error ("sequence timeout") communicating with MTS_SAP_ETH_PORT_CHANNEL_MGR for opcode MTS_OPC_ETHPM_PORT_CLEANUP
```

If the Ethernet interface is still down, please collect the `show tech-support ethpm` output and contact the Cisco Technical Assistance Centre (TAC).

### 9.6 Verifying Ethernet Port-Channel Status

First, verify physical connections on the port-channel member interfaces. Check the cable between the port-channel members; that is, the cable between the Cisco Nexus 7700 core and the Nexus 7700 storage edge switch. If the cable is not connected, the `show interface ethernetx/y` output will show as:

```
show int ethernet 1/1

Ethernet1/1 is down (Link Not Connected)
```
If Ethernet ports between the Cisco Nexus 7700 core switch and the Nexus 7700 storage switch are up but Ethernet port channels are not up, check the running configuration on the core and storage edge switches:

```
show run interface port-channel 101
```

```
interface port-channel101
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 101
```

If the running configuration looks OK, check the trunking status of the allowed VLANs by executing:

```
N7K-storage# show interface port-channel 101 switchport
Name: port-channel101
  Switchport: Enabled
  Switchport Monitor: Not enabled
  Operational Mode: trunk
  Access Mode VLAN: 1 (default)
  Trunking Native Mode VLAN: 1 (default)
  Trunking VLANs Allowed: 101
  FabricPath Topology List Allowed: 0
  Administrative private-vlan primary host-association: none
  Administrative private-vlan secondary host-association: none
  Administrative private-vlan primary mapping: none
  Administrative private-vlan secondary mapping: none
  Administrative private-vlan trunk native VLAN: none
  Administrative private-vlan trunk encapsulation: dot1q
  Administrative private-vlan trunk normal VLANs: none
  Administrative private-vlan trunk private VLANs: none
  Operational private-vlan: none
```

Next, check the VLAN membership of the port-channel interface:

```
N7K-storage# show vlan
```

### VLAN Name  | Status  | Ports
--- | ------ | ---
1   | default | active
101 | VLAN0101 | active | Po101, Eth8/1, Eth8/2

Verify no sequence timeout errors in syslog:

```
2014 Jun 22 15:01:37 DCE-1 %$ VDC-1 %$ %ETHPORT-2-IF_SEQ_ERROR: Error ("sequence timeout")
while communicating with component MTS_SAP_ETH_PORT_CHANNEL_MGR for opcode
MTS_OPC_ETHPM_PORT_BRINGUP (RID_PORT: Ethernet1/1)
```
9.7 Verifying DCBX and LLDP for vfc Port Channels

If DCBX negotiation fails between switches for port channels, most likely there is a QoS PFC configuration incompatibility between the switches. There may be a syslog event:

```
2014 Jun 18 10:17:18 N7K-storage %IPQOSMGR-2-QOSMGR_DCBXP_PFC_CMP_FAIL_MSG: Ethernet8/2 - qos config 'Priority-flow-control' not compatible with the peer
```

Check the DCBX status on the interfaces that make up the port channel:

```
N7K-storage# show system internal dcbx info interface ethernet 8/1

Interface info for if_index: 0x1a380000(Eth8/1)
tx_enabled: TRUE
rx_enabled: TRUE
dcbx_enabled: TRUE
DCX Protocol: CEE
```

All should be set to TRUE.
Verify that the vfc port-channel interface (FCoE) comes up:

N7K-storage# show interface vfc101
vfc101 is trunking
Bound interface is port-channel101
Hardware is Ethernet
Port WWN is 20:64:00:26:98:08:fb:3f
Admin port mode is E, trunk mode is on
snmp link state traps are enabled
Port mode is TE
Port vsan is 1
Speed is 20 Gbps
Trunk vsans (admin allowed and active) (101)
Trunk vsans (up) (101)
Trunk vsans (isolated)
Trunk vsans (initializing)
256419 fcoe in packets
21936804 fcoe in octets
256418 fcoe out packets
28800216 fcoe out octets
Interface last changed at Sun Jun 11 08:04:18 2014

If the vfc and Ethernet configurations look good, verify LLDP neighbors:

N7K-storage# show lldp neighbors
Capability codes:

(R) Router, (B) Bridge, (T) Telephone, (C) DOCSIS Cable Device
(W) WLAN Access Point, (P) Repeater, (S) Station, (O) Other

<table>
<thead>
<tr>
<th>Device ID</th>
<th>Local Intf</th>
<th>Hold-time</th>
<th>Capability</th>
<th>Port ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>N6K</td>
<td>Eth8/1</td>
<td>120</td>
<td>BR</td>
<td>Eth1/1</td>
</tr>
<tr>
<td>N6K</td>
<td>Eth8/2</td>
<td>120</td>
<td>BR</td>
<td>Eth1/2</td>
</tr>
</tbody>
</table>

Total entries displayed: 2
Verify that the LLDP DCBX exchange information for the local and peer interfaces reflects the same:

```
N7K-storage# show lldp dcbx interface ethernet 8/1
```

Local DCBXP Control information:
```
Operation version: 00  Max version: 00  Seq no: 20  Ack no: 34
Type/
Subtype    Version  En/Will/Adv  Config
003/000  000  Y/N/Y  0808
004/000  000  Y/N/Y  8906001b21 08
002/000  000  Y/N/Y  1123200019 19191900 00000004
```

Peer's DCBXP Control information:
```
Operation version: 00  Max version: 00  Seq no: 34  Ack no: 20
Type/  Max/Oper
Subtype    Version  En/Will/Err  Config
003/000  000/000  Y/N/N  0808
004/000  000/000  Y/N/N  8906001b21 08
002/000  000/000  Y/N/Y  0001000032 32000000 00000002
```

In that configuration, class subtype 3 refers to FCoE APP, En refers to Enable, and Will refers to Willing. For FCoE, En should be Y after successful DCBX exchange.

Verify that `fcfwd mpmap` shows the MAC address for the respective vfcs (N7K command):

```
N7K-storage# show system internal fcfwd mpmap vfcs
```

<table>
<thead>
<tr>
<th>ID</th>
<th>if-index</th>
<th>S</th>
<th>M</th>
<th>T</th>
<th>Members</th>
<th>MAC(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>vfc101</td>
<td>0x1e000064</td>
<td>D</td>
<td>E</td>
<td>--</td>
<td>0x16000064</td>
<td>eth-pc 101 (U)+</td>
</tr>
</tbody>
</table>

In that configuration, MAC should reflect both MAC addresses; that is, the MAC addresses for the peer and local interfaces.
Verify that the port-channel vfc is in the up state with the N7K command; the N5K command is slightly different and can be seen in the following output:

N7K-storage# show system internal fcoe_mgr info interface vfc101
vfc101(0x83b6c54), if_index: 0x1e000064, VFC RID vfc101
FSM current state: FCOE_MGR_VFC_ST_PHY_UP
PSS Runtime Config:-
  Type: 3
  Bound IF: Po101
  FCF Priority: 128 (Global)
PSS Runtime Data:-
  IOD: 0x00000000, WWN: 20:64:00:26:98:08:fb:3f
  Created at: Sun Jan 11 08:03:52 2009

FC Admin State: up
Oper State: up, Reason: down
Eth IF Index: Po101
Port Vsan: 1
Port Mode: E port
Config Vsan: 101
Oper Vsan: 101
Solicits on vsan: 101
Isolated Vsan:
FIP Capable ?: TRUE
UP using DCBX ?: FALSE
PSS VN Port data:-

Run the following command on the Cisco Nexus 5600:

N6K# show platform software fcoe_mgr info interface vfc101
vfc101(0x845b264), if_index: 0x1e000064, VFC RID vfc101
FSM current state: FCOE_MGR_VFC_ST_PHY_UP
PSS Runtime Config:-
  Type: 3
  Bound IF: Po101
  FCF Priority: 128 (Global)
  Disable FKA: 0
PSS Runtime Data:-
  IOD: 0x00000000, WWN: 20:64:00:2a:6a:35:a5:3f
  Created at: Mon Jan 16 16:30:20 2012

FC Admin State: up
Oper State: up, Reason: down
To run this command on the Cisco Nexus 7000, use the following:

```
N7K# show system internal fcoe_mgr info interface vfc101
```

Note the other possible FSM states that may warrant further investigation:

- No transition
- FSM_ST_NO_CHANGE
- FSM_ST_AN
- FCOE_MGR_VE_PROTO_ST_INIT
- FCOE_MGR_VE_PROTO_ST_BRUP_WAIT
- FCOE_MGR_VE_PROTO_ST_BRDOWN_WAIT
- FCOE_MGR_VE_PROTO_ST_DOWN
- FCOE_MGR_VE_PROTO_ST_DELETE_WAIT
If vFC is up and FCF discovery has gone through, the FCoE database will show the vfc interfaces in the FCoE database:

```
N6K# show fcoe database

<table>
<thead>
<tr>
<th>INTERFACE</th>
<th>FCID</th>
<th>PORT NAME</th>
<th>MAC ADDRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>vfc1</td>
<td>0xd10000</td>
<td>20:00:74:26:ac:17:2a:b1</td>
<td>74:26:ac:17:2a:b1</td>
</tr>
</tbody>
</table>
```

Total number of flogi count from FCoE devices = 1.

**VE Ports:**

```
<table>
<thead>
<tr>
<th>INTERFACE</th>
<th>MAC ADDRESS</th>
<th>VSAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>vfc101</td>
<td>00:26:98:08:fb:00</td>
<td>101</td>
</tr>
</tbody>
</table>
```

### 9.8 Verify vfc Interface and FLOGI Status

If the Ethernet interface is up but vfc is not, execute the `show port internal event-history interface vfc1` command and check for the FSM state. The status that follows shows a normal interface status:

```
N6K# show port internal event-history interface vfc 1

>>>>FSM: <vfc1> has 200 logged transitions<<<<

snip . . .

200) FSM:<vfc1> Transition at 623770 usecs after Thu Jun 26 15:23:18 2014
    Previous state: [PI_FSM_ST_TXPORT_INIT_TRUNKING_ENABLED]
    Triggered event: [PI_FSM_EV_PACER_TIMER_EXPIRED]
    Next state: [FSM_ST_NO_CHANGE]
    Curr state: [PI_FSM_ST_TXPORT_INIT_TRUNKING_ENABLED]
```
If the state is not correct, execute the `show port internal event-history errors` command and observe any error associated with the vfc port:

```
N6K# show port internal event-history errors
snip . . .
   [102] pm_error_disable_port: parent_function nl_fsm_ac_port_init_resp_rcvd,
   Ifindex (vfc101)0x1e000064, Err disabled VLAN L2 down on Eth
   interface(0x42070010) event 0x138 reason (pre_i
   nit_from_fcce_mgr_failed), cfg_wait_str: cfg wait for none
```

Verify that the Ethernet interface belongs to the correct VLAN(s) - namely, FCoE VLAN:

```
N6K# show vlan id 101

<table>
<thead>
<tr>
<th>VLAN Name</th>
<th>Status</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>101 VLAN0101</td>
<td>active</td>
<td>Po101, Eth1/1, Eth1/2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eth101/1/1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VLAN Type</th>
<th>Vlan-mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>enet</td>
</tr>
<tr>
<td></td>
<td>CE</td>
</tr>
</tbody>
</table>
```

Confirm that the vfc belongs to the correct VSAN:

```
N6K# show vsan 101 membership
vsan 101 interfaces:
    vfc1
```
Verify queuing on the Ethernet interface:

N6K# show queuing interface ethernet 101/1/1
if_slot 33, ifidx 0x1f640000
Ethernet101/1/1 queuing information:

   Input buffer allocation:
   Qos-group: 1
         frh: 3
   drop-type: no-drop
   cos: 3
   xon    xoff     buffer-size
     --------+-----------------------
    8960    14080    24320

   Qos-group: 0
         frh: 8
   drop-type: drop
   cos: 0 1 2 4 5 6
   xon    xoff     buffer-size
     --------+-----------------------
        0    117760   126720

   Queueing:
   queue  qos-group  cos      priority  bandwidth mtu
                 --------+-----------------------+
       2       0  0 1 2 4 5 6  WRR    50   1600
       3       1     3  WRR    50   2240

   Queue limit: 66560 bytes

   Queue Statistics:
   queue  rx    tx
          ---------+--------
        2     200  3113
        3    19805 100065
Confirm that the vfc interface is trunking and the VSAN is up. The VSAN must be in the up state for FCoE traffic to traverse the vfc:

N6K# show interface vfc1
vfc1 is trunking

Bound interface is Ethernet101/1/1
Hardware is Ethernet
Port WWN is 20:00:00:2a:6a:35:a5:3f
Admin port mode is F, trunk mode is on
snmp link state traps are enabled
Port mode is TF
Port vsan is 101
Trunk vsans (admin allowed and active) (101)
Trunk vsans (up) (101)
Trunk vsans (isolated) ()
Trunk vsans (initializing) ()
1 minute input rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
1 minute output rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
1154 frames input, 201320 bytes
0 discards, 0 errors
160 frames output, 17332 bytes
0 discards, 0 errors
last clearing of "show interface" counters Tue Jun 17 21:47:00 2014

Interface last changed at Thu Jun 26 15:23:10 2014
Verify that vfc has performed FLOGI:

```bash
N6K# show flogi database interface vfc1

<table>
<thead>
<tr>
<th>INTERFACE</th>
<th>VSAN</th>
<th>FCID</th>
<th>PORT NAME</th>
<th>NODE NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>vfc1</td>
<td>101</td>
<td>0xd10000</td>
<td>20:00:74:26:ac:17:2:a:b1</td>
<td>10:00:74:26:ac:17:2:a:b1</td>
</tr>
</tbody>
</table>
```

If vfc is up but the `show flogi database` command doesn’t show an entry for it, check the FSM state using the `show flogi internal event-history interface vfc` command:

```bash
N6K# show flogi internal event-history interface vfc1

>>> FSM: <[101]20:00:74:26:ac:17:2:a:b1> has 15 logged transitions<<<<

snip . . .
```


- Previous state: [FLOGI_ST_PERFORM_CONFIG]
- Triggered event: [FLOGI_EV_CONFIG_DONE_COMPLETE]
- Next state: [FLOGI_ST_FLOGI_DONE]
- Curr state: [FLOGI_ST_FLOGI_DONE]

If the FLOGI is in any of the following states, please check the flogi event-history error (previous configuration) to investigate:

- FLOGI_ST_FLOGI_RECEIVED
- FLOGI_ST_GET_FCID
- FLOGI_ST_PERFORM_CONFIG
- FLOGI_ST_FLOGI_DONE
- FLOGI_ST_CLEANUP
- FLOGI_ST_DESTROY_FSM
- FLOGI_ST_PERFORM_FCFWD_CONFIG
- FLOGI_ST_FETCH_PRCFG_INFO
- FLOGI_ST_QUERY_PORT_NUMBER
- FLOGI_ST_CHECK_SECURITY_NEGOTIATION
- FLOGI_ST_FCSP_READY
- FLOGI_ST_FCSP_HANDSHAKE_RCVD_FLOGI_CFG
- FLOGI_ST_WAIT_FCSP_DONE
- FLOGI_ST_CHECK_PORT_LOCK
- FLOGI_ST_DPVM_CHECK
Confirms that FCoE in/out packets are incrementing. You can also verify counters on the physical Ethernet interface:

```
N6K# show interface vfc1 counters

vfc1
  1154 fcoe in packets
  201320 fcoe in octets
  160 fcoe out packets
  17332 fcoe out octets
```

**9.9 Verify Interface Status with fcoe_mgr**

The `fcoe_mgr info global` command shows the detailed status of the fcoe_mgr module running in software as well as the state transitions of all the vfc interfaces (including port channels). On switches with a large number of vfc interfaces, you need to parse the output with the "begin" utility within the `show` command. The `fcoe_mgr` commands are slightly different between the Cisco Nexus 5000 and Nexus 5600 and the Cisco Nexus 7000 platforms, but the output is very similar. The first command is what is issued on the Cisco Nexus 5000 and Nexus 5600 and the second is for the Cisco Nexus 7000 with its subsequent output:

```
N6K# show platform software fcoe_mgr info global

N7K-storage# show system internal fcoe_mgr info global

FCOE-Mgr module
Low Priority Pending queue: len(0), max len(1) [Thu Jun 26 17:48:19 2014]
High Priority Pending queue: len(0), max len(2) [Thu Jun 26 17:48:19 2014]
Log Buffer:
dequeued timer msg: rid (0x6), event_id (0)

FCOE-Mgr database
-----------------------------
------- Global Config Data ------
-----------------------------
  FCMAP: 0xefc00
  FCF Priority: 128
  FKA Adv Period: 8
  VE Loopback : disabled
-------VLAN Info--------
-----------------------------
Info for VLAN 101
  fcoe_enabled: 1
  vsan_id: 101
-------FCF Info-------
-----------------------------
```

Note VLAN enabled = 1
FCF

FCF MAC Addr: 00:00:00:00:00:00:00
FCF Num Pinned by NPM: 0
FCF Num Pinned by FIP: 0
FCF Priority Offset : 0
List of Active VSANS
FCF Vsan: 101
FCF Switch WWN: 20:65:00:26:98:08:0b:01
FCF Fabric WWN: 20:65:00:26:98:08:0b:01

vfc101(0x83b6c54), if_index: 0x1e000064, VFC RID vfc101

FSM current state: FCOE_MGR_VFC_ST_PHY_UP
PSS Runtime Config:-
  Type: 3
  Bound IF: Po101
  FCF Priority: 128 (Global)
PSS Runtime Data:-

>>>> FSM: <vfc101> has 15 logged transitions<<<<

1) FSM:<vfc101> Transition at 154738 usecs after Sun Jan 11 08:03:52 2009
   Previous state: [FCOE_MGR_VFC_ST_INIT]
   Triggered event: [FCOE_MGR_VFC_EV_CREATE]
   Next state: [FCOE_MGR_VFC_ST_CREATE_WAIT]

2) FSM:<vfc101> Transition at 156412 usecs after Sun Jan 11 08:03:52 2009
   Previous state: [FCOE_MGR_VFC_ST_CREATE_WAIT]
   Triggered event: [FCOE_MGR_VFC_EV_CREATE_SUCCE]
   Next state: [FCOE_MGR_VFC_ST_CREATED]

snip ...

Curr state: [FCOE_MGR_VFC_ST_PHY_UP]

PROTOS Info:
vfc101(0x83e9ce4), if_index: 0x1e000064, VEProto RID vfc101, vsan 101
FSM current state: FCOE_MGR_VE_PROTO_ST_UP
PSS Runtime Data:-
  Eth IF Index: Po101
  Port Mode: Unknown(0)
  FKA check enabled ? : TRUE
10. Troubleshooting FCoE Fabric

10.1 VFC VSAN Is in Initializing State

If you find a vfc with its VSAN in initializing state, check the conditions as follows. The following is an example `show interface vfc1` output. It's important to note that although the vfc is showing up because of the "vfc1 is trunking," the specified VSAN is not up, so FCoE traffic will not traverse the interface. Note in the example that follows that VSAN 101 is the configured VSAN for the vfc:

```
NEK# show interface vfc1
vfc1 is trunking
  Bound interface is Ethernet101/1/1
  Hardware is Ethernet
  Port WWN is 20:00:00:2a:6a:35:a5:3f
  Admin port mode is F, trunk mode is on
  snmp link state traps are enabled
  Port mode is TF
  Port vsan is 101
  Trunk vsans (admin allowed and active) (101)
    Trunk vsans (up)        ()
    Trunk vsans (isolated)  ()
    Trunk vsans (initializing) (101)
  1 minute input rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
  1 minute output rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
    12 frames input, 1308 bytes
      0 discards, 0 errors
    13 frames output, 1400 bytes
      0 discards, 0 errors
  last clearing of "show interface" counters Tue Jun 17 21:47:00 2014

Interface last changed at Tue Jun 17 21:47:22 2014
```
You should also investigate any discards or errors in the previous output, as well as check the bound Ethernet interface for discard errors.

The switch syslog will also indicate that the trunked VSAN is down. Note that this output does not indicate the vfc interface is down but rather the vfc status. Again, this output is an indication that the VSAN is in initializing state:

```
2014 Jun 18 20:33:04 N6K %PORT-5-IF_TRUNK_DOWN: %$VSAN 101$ Interface vfcl, vsan 101 is down (waiting for flogi)
```

### 10.2 Check QoS and PFC

If the FCoE class in network-qos settings isn’t configured, the interface won’t come out of initializing state. Because this value is a system-qos value, none of the vfc interfaces will come up as they all rely on the same valid QoS setting.

First, check the QoS settings. The following example is from a Cisco Nexus 5000. The different platforms may have slightly different outputs:

```
N6K# show running-config ipqos
system qos
    service-policy type queuing input fcoe-default-in-policy
    service-policy type queuing output fcoe-default-out-policy
    service-policy type qos input fcoe-default-in-policy
    service-policy type network-qos fcoe-default-nq-policy
```

Following is the output from a Cisco Nexus 7000:

```
N7K# show running-config ipqos
system qos
    service-policy type network-qos default-nq-7e-policy
```

You can also verify PFC at the interface level (remember PFC is Ethernet, so look at the Ethernet interface):

```
N6K# show interface ethernet 101/1/1 priority-flow-control
===============================================
<table>
<thead>
<tr>
<th>Port</th>
<th>Mode</th>
<th>Oper(VL bmap)</th>
<th>RxPPP</th>
<th>TxPPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethernet101/1/1</td>
<td>Auto</td>
<td>On (8)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
```

If PFC is correct, VL bmap should be 8.
Also check detailed queuing information on the interface:

N6K# show queuing interface ethernet 101/1/1
if_slot 33, ifidx 0x1f640000
Ethernet101/1/1 queuing information:
  Input buffer allocation:
    Qos-group: 1
    frh: 3
    drop-type: no-drop
    cos: 3
    xon xoff buffer-size
    ------------------------+--------
    8960  14080  24320

Qos-group: 0

Make sure there is group 1 and no-drop.
frh: 8
drop-type: drop
cos: 0 1 2 4 5 6
xon       xoff       buffer-size
----------------------------------------
0          117760     126720

Queueing:
queue  qos-group  cos       priority  bandwidth  mtu
--------+------------+--------+---------------------+----------+-------
2       0          0 1 2 4 5 6   WRR      50        1600
3       1          3               WRR      50        2240

Queue limit: 66560 bytes

Queue Statistics:
queue  rx        tx
--------+----------+----------
2       200      1656
3       18724    12026

Port Statistics:
rx drop  rx mcast drop  rx error  tx drop  mux ovflow
----------------+-----------------+--------+---------+---------------------
0          0        0        0        0        InActive

Priority-flow-control enabled: yes
Flow-control status:
cos  qos-group  rx pause  tx pause  masked rx pause
--------+-----------+---------+---------+------------------------
0      0          xon      xon      xon
1      0          xon      xon      xon
2      0          xon      xon      xon
3      1          xon      xon      xon
4      0          xon      xon      xon
5      0          xon      xon      xon
6      0          xon      xon      xon
7      n/a        xon      xon      xon
10.3 Check DCBX (LLDP)

DCBX is enabled on LLDP-enabled interfaces. LLDP is enabled globally when the feature is enabled. It is possible to disable LLDP (transmit or receive or both on a per-interface basis). If either transmit or receive is disabled, DCBX is automatically disabled. Check the LLDP and DCBX status on the interface:

N6K# show lldp interface e101/1/1
Interface Information:
Enable (tx/rx/dcbx): Y/Y/Y       Port Mac address: c8:f9:f9:20:cl:02

Peer's LLDP TLVs:
Type Length Value
---- ------ -----
001 007 047426ac 172aac
002 007 037426ac 172ab0
003 002 0078
127 055 001b2102 020a0000 0000001 0000000e 080a0000 c0008906 001b2108
       06060000 c000ff08 04110000 c000ffff ffff0000 00000000 000008
127 005 00014201 01
000 000

Note LLDP and DCBX are enabled.
Note TLV info received from peer (good).

Further check the LLDP/DCBX status:

N6K# show lldp dcbx interface e101/1/1

Local DCBX Control information:
Operation version: 00  Max version: 00  Seq no: 1  Ack no: 1
Type/
Subtype  Version   En/Will/Adv Config
003/000 000         Y/N/Y         0808
004/000 000         Y/N/Y         8906001b21 08
002/000 000         Y/N/Y         0001000032 32000000 00000002

Peer's DCBX Control information:
Operation version: 00  Max version: 00  Seq no: 1  Ack no: 1
Type/Max/Oper
Subtype  Version   En/Will/Err Config
004/000 000/000    Y/Y/N         8906001b21 08
003/000 000/000    Y/Y/N         ff08
002/000 000/000    Y/Y/N         fffffff00 00000000 00000008

If no type 003 on switch, then there is a PFC configuration problem.
Note subtype 004 should match.
Check that DCBX packets are incrementing. This command has a lot of detailed information. To look at just the packet status, use the `show` command:

```
N6K# show system internal dcbox info interface Ethernet 1/1 | begin "DCBX pkt"
```

DCBX pkt stats:

```
Total frames out: 9041
Total Entries aged: 38
Total frames in: 9002
DCBX frames in: 8989
Total frames received in error: 0
Total frames discarded: 0
Total TLVs unrecognized: 0
```

Run this command several times and verify the frames are incrementing.

**10.4 CNA Settings Must Support FIP Request/Response**

The CNA settings of the server need to be set correctly in order to support vfc instantiation. Because there are multiple CNA vendors with various models, it is not possible to list all the permutations of settings. The key is that CNA is configured to support FIP and FIP Request/Response. The server system administrator has tools to query the CNA for its current configuration. You can verify CNA is communicating with the switch properly by looking at fcoe_mgr information and LLDP:

```
N6K# show platform software fcoe_mgr info interface vfc1
vfc1(0x8461fa4), if_index: 0x1e000000, VFC RID vfc1
    FSM current state: FCOE_MGR_VFC_ST_PHY_UP
    PSS Runtime Config:-
        Type: 3
        Bound IF: Eth1/1/1
        FCF Priority: 128 (Global)
        Disable FPA: 0
    PSS Runtime Data:-
        IOD: 0x00000000, WWN: 20:00:00:2a:6a:35:a5:3f
        Created at: Tue Jun 17 21:46:53 2014

        FC Admin State: up
        Oper State: up, Reason: down
        Eth IF Index: Eth1/1/1
        Port Vsan: 101
        Port Mode: F port
        Config Vsan: 101
        Oper Vsan: 101
        Solicits on vsan: 101
```

If there is no VSAN “Solicits on vsan”, then FIP is not configured properly on the CNA. Value Should show the FCoE VSAN you configured for the vfc.
10.5 Cisco Nexus 7000: Packet Doesn’t Reach SUP for FCoE- Data Packet Dropped

Do the following:

1. Check to see if any packets drops are happening; look at the statistics.
   a. Check relevant drop field on application-specific integrated circuit (ASIC) instance of an Ethernet interface where VFC is bound; if it is the vFC port channel, check the first operational Ethernet port.
   b. Execute the command `show hardware internal statistics module-all pktflow dropped`.

2. Verify PAUSE enable/disable status (debugging congestion problem).
   a. PAUSE should be enabled on No-Drop VL. Because we are using qos template 7e policy, vl is 3 here.
   b. attach module <module number of the Ethernet port where vFC is bound to>
   c. vdc <vdc id of the storage vdc, can be collected from show vdc>

   module-1# show hardware internal mac port 1 qos configuration

   QOS State for port 2 (Asic 0 Internal port 0)

<table>
<thead>
<tr>
<th>VL#</th>
<th>ENABLE</th>
<th>RESUME</th>
<th>REFRESH</th>
<th>REF_PERIOD</th>
<th>QUANTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>0x0</td>
<td>0x0</td>
</tr>
<tr>
<td>1</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>0x0</td>
<td>0x0</td>
</tr>
<tr>
<td>2</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>0x0</td>
<td>0x0</td>
</tr>
<tr>
<td>3</td>
<td>ON</td>
<td>ON</td>
<td>ON</td>
<td>0xffff0</td>
<td>0xffffffff</td>
</tr>
<tr>
<td>4</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>0x0</td>
<td>0x0</td>
</tr>
<tr>
<td>5</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>0x0</td>
<td>0x0</td>
</tr>
<tr>
<td>6</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>0x0</td>
<td>0x0</td>
</tr>
<tr>
<td>7</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>0x0</td>
<td>0x0</td>
</tr>
</tbody>
</table>
   a. attach module <module number of the Ethernet port where vPC is bound to>
   b. vdc <vdc id of the storage vdc, can be collected from show vdc>

   module-1# show hardware internal mac port 1 qos configuration |in mtu
   vl  hw_mtu  pm_mtu  pm_adj  qos_mtu  qos_adj  last_mtu
   0  1522   0  22  1500  22    1500
   1  1522   0  22  1500  22    1500
   2  1522   0  22  1500  22    1500
   3  2180   0  22  2158  22    2158
   4  1522   0  22  1500  22    1500
   5  1522   0  22  1500  22    1500
   6  1522   0  22  1500  22    1500
   7  1522   0  22  1500  22    1500

11. References
   ● FCoE Initiation Protocol (White Paper)
   ● Cisco Nexus 6000 Series FCoE Configuration Guide, Release 6.0
   ● Cisco Nexus 6000 FCoE Troubleshooting Guide
   ● Cisco Nexus 7000 Troubleshooting Guide: Troubleshooting FCoE
   ● Cisco Nexus 5500 to Nexus 7000 Multihop FCoE Configuration Example
   ● FCoE Configuration Guide for Cisco Nexus 7000