The Truth about FCoE: Deployment Experiences

Mike Frase
Cisco Systems
Agenda

- **Skills Requirements - FC & Ethernet**
  - The Unified I/O Network Engineer
- **Popular Features for Data Center 10 GigE FCoE Deployments**
- **Today's Deployment Topologies for FCoE**
- **Unified I/O in Compute Blade Center**

Deployment guides
Learning Curve — Education to Users
Unified Network Troubleshooting

Bringing together Server, Network, and Storage Engineers

- Ethernet
- Layer 2 & 3 protocols
- VLANS

- Cabling and Optics
- 10 GigE NIC
- HBA
- FCoE
- DCBX
- Performance Monitoring

- Fibre Channel Protocol and FC services
- Zoning
- SCSI data
- VSANs

- O/S & Driver support
Supportability with Monitor Ports

Ability to still analyze the virtual interfaces of ethernet and Fibre Channel - same as physical interfaces

For Ethernet Protocols

<table>
<thead>
<tr>
<th>Session</th>
<th>State</th>
<th>Reason</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>up</td>
<td>The session is up.</td>
<td></td>
</tr>
</tbody>
</table>

```
MN-5020-1(config-if)# sh monitor
Session  State        Reason                  Description
------- ----------- ---------------------- --------------
1         up           The session is up
```

```
MN-5020-1# sh monitor sess 1
   session 1

--------
type   : local
state  : up
source intf :
  rx    : Eth1/29
  tx    : Eth1/29
  both  : Eth1/29
source VLANs :
  rx    :
source VSANs :
  rx    :
destination ports : Eth1/1
```

For the Fibre Channel Protocols

```
MN-5020-1(config-monitor)# sh monitor sess 2
   session 2

-------------
type       : local
state      : up
source intf :
  rx        : vfc29
  tx        : vfc29
  both      : vfc29
source VLANs :
  rx        :
source VSANs :
  rx        :
destination ports : fc2/3
```
Features Deployed with most FCoE deployments

- In service software upgrades - ISSU for Data Path
- vPC – Virtual Port-Channels but still maintain the SAN A-B separation for FC
- Switches in N_Port Virtualization mode vs FC Switching – Simplifies SAN interop
- LACP to Host - Link Aggregation Control Protocol for NIC bonding
High Availability with Virtual Port-Channels - Single-Sided vPC

This is for the Data traffic

Spanning-tree Root

logical equivalent

Spanning-tree Root
High Availability with Virtual Port-Channel - Double-Sided vPC

- **Root**: The primary and secondary ports are connected through a Peer Link.
- **Po51**: Logical equivalent of the regular STP priority.
- **Po10**: Symbolizing the double-sided vPC.
Deployments with FCoE
Deployments with FCoE

- Converged Network Adapter (CNA) presents two PCI address to the Operating System (OS)
- OS loads two unique sets of drivers and manages two unique application topologies
- Server participates in both topologies separately
- Two stacks and thus two different views of the same ‘unified wire’
  - SAN Multi-Pathing provides failover between two fabrics (SAN ‘A’ and SAN ‘B’)
  - NIC Teaming provides failover within the same fabric (VLAN)
- **Generation 2 CNA**
  - Utilizes FCoE Initialization Protocol (FIP) as defined by the T.11 FC-BB-5 specification
  - Supports both direct and via FIP Bridge attachment (example: through a Cisco Nexus 4000 FIP Snooping Bridge)

What is being deployed
Single Hop Design

![Diagram of network topology with Flagship A Fabric, DCBX, FCoE Initialization Protocol (FIP), and vPC's (Virtual Port Channels) connecting Fabric A and Fabric B through Direct attach VN_Port to VF_Port.]
What is being deployed
Attaching an Initiator

- Physical link is brought up (today requires 10GE)
- DCBX negotiation – discovers DCB capable devices and negotiates lossless Ethernet capabilities/configs
- FIP Process – discovery and negotiation of FCoE devices and characteristics
  - FCoE VLAN Discovery
  - FCF Discovery on the specific FCoE VLAN
  - Fabric Login - builds the logical wire from the end node to the FCF (VN_port to VF_port)
- FCoE traffic flows from host to target; LAN traffic flows
A VLAN is dedicated for every VSAN in the fabric

FIP discovers the FCoE VLAN and signals it to the hosts

Trunking is not required on the host driver – all FCoE frames are tagged by the CNA

FCoE VLANs must not be configured on Ethernet links that are not designate for FCoE

Maintains isolated edge switches for SAN ‘A’ and ‘B’ and separate LAN switches for NIC 1 and NIC 2 (standard NIC teaming)

! VLAN 20 is dedicated for VSAN 2 FCoE traffic
(config)# vlan 20
(config-vlan)# fcoe vsan 2
In order to maintain the integrity of FC forwarding over FCoE, FCoE VLANs are treated differently than LAN VLANs:

- No flooding, MAC learning, broadcasts, etc.
- The FCoE VLAN must not be configured as a native VLAN
- FIP uses native VLAN
- Separate FCoE VLANs must be used for FCoE in SAN-A and SAN-B
- Unified Wires must be configured as trunk ports and STP edge ports

VLAN 20 is dedicated for VSAN 2 FCoE traffic

```
(config)# vlan 20
(config-vlan)# fcoe vsan 2
```
- FCoE Fabric ‘A’ will have a different VLAN topology than FCoE Fabric ‘B’ which are different from the LAN Fabric
- PVST+ allows unique topology per VLAN
- MST requires that all switches in the same Region have the same mapping of VLANs to instances
- MST does not require that all VLANs be defined in all switches
  - A separate instance must be used for FCoE VLANs
  - **Recommended:** three separate instances – native Ethernet VLANs, SAN ‘A’ VLANs and SAN ‘B’ VLANs

```
spanning-tree mst configuration
name FCoE-Fabric
revision 5
instance 5 vlan 1-19,40-3967,4048-4093
instance 10 vlan 20-29
instance 15 vlan 30-39
```
Optimal layer 2 LAN design often leverages Multi-Chassis Etherchannel (MCEC).

Nexus utilizes Virtual Port Channel (vPC) to enable MCEC either between switches or to 802.3ad attached servers.

MCEC provides network based load sharing and redundancy without introducing layer 2 loops in the topology.

MCEC results in diverging LAN and SAN high availability topologies:
- FC maintains separate SAN ‘A’ and SAN ‘B’ topologies
- LAN utilizes a single logical topology.
vPC enabled topologies with FCoE must follow specific design and forwarding rules

The Virtual Fibre Channel (vfc) interface can only be associated with a vPC which has a single [one (1)] CNA port attached to each edge switch

While the port-channel is the same on FCF-A and FCF-B, the FCoE VLANs are different

vPC configuration works with Gen-2 FIP enabled CNAs only

FCoE VLANs are ‘not’ carried on the vPC peer-link

FCoE and FIP ethertypes are ‘not’ forwarded over the vPC peer link
Server Ethernet driver connected to the FEX in NIC Teaming or with vPC (802.3ad)

FCoE runs over vPC member port with a single link from server to Remote Line Card

Each TOR Remote Line Card single homed to upstream FCF
  FEX fabric links can be connected to FCF with individual links (static pinning) or a port channel
  oversubscribed 4:1

Consistent with separate LAN Access and SAN Edge Topologies
As an example: the Cisco Nexus 4000 is a Unified Fabric capable Blade Switch

- DCB enabled
- FIP Snooping Bridge

Dual Topology requirements for FCoE multi-hop

Servers IP connection to the FIP snooping bridge is Active/Standby

- MCEC is not currently supported from blade server to a Nexus 4000

Option 1: Unified Dedicated Wires from FIP Snooping Bridge to FCF

Option 2: Single Unified Wire Port Channel from FIP Snooping Bridge to FCF
- Virtual E_Ports (VE) deployable now
- What design considerations do we have when extending FCoE beyond the Unified Edge?
  - High Availability for both LAN and SAN
  - Oversubscription for SAN and LAN
  - Ethernet Layer 2 and STP design
- Where does Unified Wire make sense over Unified Dedicated Wire?
- Unified Wire provides for sharing of a single link for both FC and Ethernet traffic
Unified I/O in Compute Blade Center
Unified Fabric Within Compute Chassis: FCoE

- Fewer Cables
- Fewer switches
- Fewer adapters
- Overall less power
- Interoperates with existing SAN’s
Easy Start - FCoE in the Compute Blade center

Fabric Consolidation

- Fewer switches
- Fewer adapters
- All I/O types available in each chassis/blade
  - 10GE & FCoE
  - LAN, SAN, IPC
  - Easier to manage
- Blades can work with any chassis
- Small network domain

Past Approach

- All fabric types have switches in each chassis
- Repackaged switches
- Complex to manage
- Blade-chassis configuration dependency
- Costly
- Small network domain
Easy Start - FCoE in the Compute Blade center
Fabric Consolidation

Switch
- Only one type needed
  - LAN, SAN, IPC
- μsec latency
- Non-blocking
- Native uplinks
  - FC
  - Ethernet
- Not a gateway

Protocol
- Standards
  - Ethernet
  - FCoE
  - Native FC in Ethernet
- Priority flow control
  - Non-dropping

Adapter
- Only one type needed
- Multi-protocol
  - FC
  - Ethernet
Ethernet Fabric:
Single Fabric
10 GE Connected
Switch or End-host Mode
vPC / VSS on receiving Ethernet Cloud

Storage Fabric:
Dual Fabrics
4G/8G FC Connected
NPV Mode
Multipathing is host-based
Thank You

Q & A
Design Guide Links


- Google FCoE Design Guides