FCoE Direct End-Node to End-Node (aka FCoE VN2VN)

John L Hufferd
Hufferd Enterprises
A new concept has recently been accepted for standardization by the FC-BB-6 Working Group within the Fibre Channel (T11) standards committee; it is called FCoE VN2VN (aka Direct End-Node to End-Node)

T11 previously standardized the FCoE specification (which defines the encapsulation of Fibre Channel frames within Ethernet Frames) and is currently extending that specification to permit FCoE connections DIRECTLY between FCoE End-Nodes

This tutorial will show the fundamentals of the extended FCoE concept that permits it to operate without FC switches or FCoE Switches (aka FCFs) and will describe how it might be exploited in a Data Center environment
Agenda

- Introduction
- FCoE & FCoE VN2VN
- Architecture
- Discovery & Link Instantiation
- Topologies
- Scenarios
- Summary
Introduction

- This presentation provides an overview of a new proposed Standard called (herein) FCoE Direct End-Node to End-Node (aka FCoE VN2VN)
  - This is a Lossless Ethernet connection directly between Adapters’ Virtual N_Ports

- One should think about FCoE VN2VN as placing the FCoE (FC) protocol on a Lossless Ethernet without the additional requirement of FC Switches or FCoE Switches known as FC Forwarders (FCFs)
  - Permits connections through only (Lossless) Ethernet Switches
  - Permits connections via a single wire Point-to-Point

- The protocol is being defined in the INCITS Fibre Channel (T11) technical committee (FC-BB-6 Ad Hoc Work Group)
FCoE & FCoE VN2VN
FCoE is an Alternative to FC

- FCoE stands for FC over Ethernet
- FCoE was defined as an alternative network structure for carrying FC protocols
- **FCoE requires** specific Ethernet capabilities to be implemented
  - Lossless switches and fabrics (e.g., supporting IEEE 802.3 PAUSE) configurations are required
  - Jumbo frame support is strongly recommended (not a standard, but widely available)

- Deployments of FCoE should utilize the advances in Ethernet currently specified in IEEE 802.1, specifically:
  - Priority-based Flow Control (PFC) \(\rightarrow\) 802.1Qbb
  - Enhanced Transmission Selection (ETS) \(\rightarrow\) 802.1Qaz
  - DCB (capability) eXchange (DCBX) Protocol \(\rightarrow\) 802.1Qaz
  - Congestion Notification (802.1Qau),

- **Possible** future \(\rightarrow\) Multi-pathing (IETF– TRILL, IEEE 802.1aq SPB, et.al.)

- These 802.1 advance capabilities are important for Converged Flows (Messaging, Clustering and Storage)
  This set of functions is called DCB -- Data Center Bridging

- FCoE Fabrics require an FCoE Lossless Ethernet Switch that understands & supports FC protocols – These Switches are called FCFs (FCoE Forwarders)
FCoE VN2VN
(Virtual N_Port to Virtual N_Port)

- FCoE VN2VN is a Lossless Ethernet connection between End-Node Adapters’ VN_Ports

- Other than Ethernet Cables, only DCB Ethernet switches may exist between the End-Nodes (VN_Ports) – Therefore, the connection maybe either a:
  - Switched Lossless Ethernet connection
  - Or
  - Point to Point Lossless Ethernet Wire connection

- FCoE VN2VN permits FCoE networks to be built without any FC Switches or FCoE Switches (aka FCFs)

- FC Data Flow (& Packets) will flow End to End as if they were flowing over a direct (point to point) FC link
  - Must operate identically on a VN2VN connection as on a direct FC link

- No Fibre Channel services or advanced features (e.g. Name services, Zoning, virtual fabrics, IFR, security, etc.) are provided in the network
  - Use of VLANs, and ACLs, can help security issues
  - Security can also be enhanced via FC-SP
  - Storage arrays that have port and LUN masking will also enhance security
Connections to a DCB Switch

- Fibre Channel is carried over lossless Ethernet as a L3 protocol

<table>
<thead>
<tr>
<th>Lossless Ethernet MAC (DCB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lossless Ethernet port</td>
</tr>
<tr>
<td>(FCoE VN_Port)</td>
</tr>
<tr>
<td>IP address 123.45.67.89</td>
</tr>
</tbody>
</table>

Network Applications
- IP
- TCP
- UDP

Customer Applications
- FCoE
- iSCSI
- Fibre Channel
- SCSI

2013 Storage Developer Conference. © Hufferd Enterprises. All Rights Reserved.
The Simple VN2VN Interconnect

An DCB – Switch(es) may connect a number of VN2VN capable VN_Ports together

Or

A single Wire with VN2VN capable End Nodes can be interconnected
Architecture
### Fibre Channel over Ethernet (FCoE) Packets

<table>
<thead>
<tr>
<th>Ethernet Header</th>
<th>FCoE Header</th>
<th>FC Header</th>
<th>SCSI Commands/Data</th>
<th>FCS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ethertype</strong></td>
<td><strong>FCoE</strong></td>
<td>Ethertype</td>
<td><strong>SCSI Commands/Data</strong></td>
<td><strong>FCS</strong></td>
</tr>
<tr>
<td>“FCoE” (8906h)</td>
<td></td>
<td>Protocol control information: Version, SOF, EOF, etc.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Ethernet Header** provides things needed for the physical network, including “Ethertype”.
- **FC Imbedded Frames**: Same as in Physical FC
- **Frame Check Sequence** (CRC)

### FCoE Initialization Protocol (FIP) Packets

<table>
<thead>
<tr>
<th>Ethernet Header</th>
<th>FIP Header</th>
<th>Descriptors</th>
<th>FCS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ethertype</strong></td>
<td><strong>FIP</strong></td>
<td><strong>Descriptors</strong></td>
<td><strong>FCS</strong></td>
</tr>
<tr>
<td>“FIP” (8914h)</td>
<td></td>
<td>Discovery, Link establishment, maintenance &amp; disconnect (Login/Logout, etc.) Parameters</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Protocol control information: Version, Op-codes, etc.</td>
<td></td>
</tr>
</tbody>
</table>

- **Discovery, Link establishment, maintenance & disconnect (Login/Logout, etc.) Parameters**
FIP Protocol for FCoE/VN2VN Networks

- **Discovery Phase**
  - FCoE (w/o VN2VN)
    - FCFs Discover each other, & form a Fabric
    - ENodes Discover FCFs & Potential VN_Port $\leftrightarrow$ VF_Port pairing
  - FCoE VN2VN
    - VN2VN capable ENodes Discover each other

- **Login Phase**
  - FCoE (w/o VN2VN)
    - ENodes chose among discovered FCFs’ Ports for Virtual Link connections
  - FCoE VN2VN
    - VN2VN capable ENodes chose among discovered VN2VN Ports for Virtual Link connections
  - Both Use: FLOGI, FLOGI ACC, LOGO, etc …

- **End-to-End path control & Data Transfer Phase**
  - PLOGI/PRLI
  - All other FC protocol frames (FC4 ULPs. etc.)
### FC’s Encapsulation in Ethernet (FCoE)

<table>
<thead>
<tr>
<th>Word</th>
<th>31-24</th>
<th>23-16</th>
<th>15-8</th>
<th>7-0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n+1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n+2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Destination MAC Address** (6 Bytes)
- **Source MAC Address** (6 Bytes)
- **ET=FCoE** (16 bits)
- **Ver (4b)**
- **Reserved** (12 bits)
- **SOF** (8 bits)
- **Encapsulated FC Frame**
- **EOF** (8 bits)
- **Ethernet FCS**

- **FC Frame = Minimum 28 Bytes (7 Words)** → **Maximum 2180 Bytes (545 Words)**
  - (including FC-CRC)

Optional IEEE 802.1q 4 Byte Tag goes here

This field varies in size

Ethernet frame Size is 64 Bytes to 2220 Bytes
### FIP Operation Format

<table>
<thead>
<tr>
<th>Word</th>
<th>31-24</th>
<th>23-16</th>
<th>15-8</th>
<th>7-0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td><strong>ET=FIP</strong> (16 bits)</td>
<td><strong>Ver</strong> (4b)</td>
<td><strong>Reserved</strong> (12 bits)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td><strong>FIP Operation Code</strong></td>
<td><strong>Reserved</strong></td>
<td><strong>FIP SubCode</strong></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td><strong>Descriptor List Length</strong></td>
<td><strong>FP</strong></td>
<td><strong>Flags</strong></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n+1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Destination MAC Address** (6 Bytes)
- **Source MAC Address** (6 Bytes)
- **ET=FIP** (16 bits)
- **Ver** (4b)
- **Reserved** (12 bits)
- **FIP Operation Code**
- **Reserved**
- **FIP SubCode**
- **Descriptor List Length**
- **FP**
- **Flags**
- **PAD to minimum length or mini-Jumbo length**

- **Ethernet frame size** is 64Bytes to 2220Bytes
- **Ethernet FCS**

- **FPMA bit**
- **REC/P2P bit**
- **Available bit**
- **Solicited bit**
- **FCF bit**
Each ENode (HBA/CNA) may have multiple Physical Ethernet Ports
Each Physical Port may have multiple Logical VN_Ports

Each instantiation's N_Port_ID & MAC Address is independent of the others
There can be duplicates (if they are in different VLANs)
Each Logical VN_Port may connect to multiple other Logical VN_Ports
Discovery and Link Instantiation (FIP -- FCoE Initiation Protocol)
1. As each VN2VN enabled FCoE End-Node starts-up it will Randomly Generate its own:
   • FC LU-ID (N_Port_ID) and Ethernet MAC Address

2. Then Each VN2VN End-Node, on behalf of its VN_Port, **Multicasts** a **PROBE** with the Generated MAC Address & LU-ID
   • And **listens** for conflict responses

3. If Address/ID Conflict message is received, the process will repeat (at step1) until no Conflict messages are received

4. If no Conflicts are received, the End-Node instantiates the VN_Port and **Multicasts** its **CLAIM** to the MAC Address & LU-ID (N_Port_ID) and announces its “Capabilities” (FC-4 Features)
4. Each VN2VN enabled FCoE End-Node receiving a CLAIM will respond with its own information & record the received CLAIMed information into a Neighbor Table – Including:
   - N_Port_Name, MAC Address & LU-ID and
   - Capabilities of the CLAIMing VN_Port (Initiator/Target, etc.)

5. Upon receiving CLAIM response messages, the CLAIMing End-Node will record the received CLAIM response information into its own Neighbor Table – Including:
   - N_Port_Name, MAC Address & LU-ID and
   - Capabilities of the responding VN_Port (Initiator/Target, etc.)
VN2VN Initial Login Flow Ladder

End-Node

Probe & Claim
(after Randomly computing ID)
(repeat until no conflicts)

Multicast to “ALL-VN2VN-ENode-MACs”

Unicast Probe & Claim responses

Note: in P2P mode, the Probe is not performed & the Claim is Multicast to “ALL-PT2PT-ENode-MACs”

FLOGI

to chosen VN_Port Peer

FLOGI ACC to accepted VN_Port

End-Node

FC Command

to chosen VN_Port Peer

FC Command responses

FCoE Initialization Protocol (FIP)

Discovery Phase

Login Phase

Normal FC Processing

FCoE Protocol
Summary of Discovery Process (for VN2VN)

- Randomly chooses an identity (FC-ID & MAC Address)
- Multicast (& Respond with) FIP messages (called “Probe, Probe Response, Claim, & Claim Response”) which:
  - Insure ID uniqueness
    (repeat the Random ID creation and the Probe & Claim as needed)
  - Permit population of the neighbor tables

This will announce the VN_Ports’ identities and the VN_Ports’ capabilities to other ENodes:
- The capabilities will be used for choosing a peer VN_Port

Note: FC-ID & MAC Address should be saved, if possible, for next Reboot
After IDs and Potential Partners (VN_Ports) are identified within the Level 2 Ethernet:
- The Initiators chose their Targets and FLOGI & FLOGI ACC FIP frames are exchanged
- Then PLOGI/PLOGI ACC & PRLI/PRLI ACC FCoE Frames are exchanged
- Thus Instantiating the VN2VN Logical Link

➤ Then FCoE frames will be exchanged directly between the VN_Ports ≤

After Link Instantiation all VN_Port’s IDs will be Periodically Beaconsed (Multicast)

(Beaconing permits detection of link loss (via time-outs) & new or incorrect LAN joins)
Topologies
A CNA to CNA FCoE path between these Switch ports is now also possible even without an FCF using Direct VN2VN mode.

It is now possible to connect End-to-End as shown below.
DCB – Switch(es) may connect
• A number of VN2VN capable VN_Ports together
• Pairs of some configured VN2VN-P2P ENodes
  (Requires physically/logically configuring the Switch)

A single Wire may connect
• VN2VN ENodes (without P2P capability/configuration)
• A pair of VN2VN-P2P ENodes
Scenarios
Scenario 1: FCoE & IP Flows

- **Classical Ethernet Network**
- **Internet**
- **DCB Network**
- **FCoE VN2VN Flows**
- **IP Flows**
Scenario 2: FCoE Right & Wrong

Classical Ethernet Network

Internet

DCB Network

FCoE VN2VN Flows

Invalid FCoE VN2VN Flows
FCoE Summary

- T11.3’s FC-BB-6 Ad-Hoc Working Group accepted VN2VN specification for inclusion in the next published standard (due 2013 - ....)
  - Vendors may produce products before that

- FCoE VN2VN is a simple, efficient mechanism for encapsulating Fibre Channel in Ethernet frames on a Lossless Ethernet type Network
  - Not a traditional Ethernet Interface or fabric
  - A New Network – A lossless Ethernet Network
  - Defined in the IEEE 802.1 standards working group (and called DCB)
  - FC protocols frames will just be inserted into these Ethernet frames

- Specification permits the installation to evolve from simple VN2VN into Full FCoE Fabrics (and even real FC Fabrics)
  - Can start small with simple Networks (even Point-to-Point)
    - Perhaps with all software Initiators
  - May grow into larger FCoE fabrics in an evolutionally way
  - Full FCoE Fabrics and VN2VN networks can co-exist as installations grow
  - When upgrading you can continue to use the same physical components

- Now FC protocols can span the small, medium and Enterprise size networks
  - Only one protocol type (FC) is needed within the Data Center even if using different wire types
  - Different size companies can merge and their Data Center will easily merge also
Thank You!
Attribution & Feedback

The SNIA Education Committee would like to thank the following individuals for their contributions to this Tutorial.

Authorship History

Original Author: John L. Hufferd / 11/2010

Updates:
John L. Hufferd / 7/2011
John L. Hufferd / 9/2011
John L. Hufferd / 3/2012
John L. Hufferd / 9/2012
John L. Hufferd / 9/2013

Additional Contributors

Claudio DeSanti
Silvano Gai
Joe Pelissier
John Hufferd

Please send any questions or comments regarding this SNIA Tutorial to tracktutorials@snia.org
Appendix

- FCoE Relation to ISO Layers