FCoE

Why it’s more than just another protocol
A year ago at Dallas SNW, NetApp and several other companies showed the first FCoE demo. This presentation will discuss FCoE, and how this simple protocol can bring both capital cost and operational cost savings. It will look at how the protocol works and will use the NetApp demo as an example of how easy it is to implement and operate an FCoE environment.
FCoE Agenda

- FCoE Hardware
- FCoE Topologies
- FCoE from the Hosts Perspective
- FCoE and Security
FCoE Hardware
FCoE SAN Components

- Required Components
  - FCoE Switch
  - FCoE End Point

- Optional Components
  - Ethernet Switch (CEE capable)
  - Fibre Channel Switch
  - Fibre Channel End Point
FCoE End Point

- First Generation CNA (discrete components)
  - still high power use
  - full size
  - early adopters / proof of concept
FCoE Converged Network Adapter

- First Generation (discrete components)
Gen 1 CNA functional operation

All MAC, FCoE and Ethernet magic is handled here – making it transparent to the existing host FC S/W.

Regular Ethernet traffic flows through to NIC

- 10Gb Ethernet port
- FCoE Engine
- NIC
- FC
- PCIe
- Existing host NIC driver
- Existing host/target FC driver
Gen 2 CNA functional operation

NEW Integrated Component – containing all 3 functions

All MAC, FCoE and Ethernet magic is handled here – making it transparent to the existing host FC S/W.

Regular Ethernet traffic flows through to NIC
FCoE Converged Network Adapter

- First Generation CNA (discrete components)
  - still high power use
  - full size
  - early adopters / proof of concept

- Second Generation (integrated solution)
  - low power
  - smaller size
  - general adoption
  - Some companies are starting here
FCoE NIC

- FCoE can be implemented as complex S/W
  - Therefore can use NIC chips using FCoE S/W drivers
    - FC protocol (including CRC)
    - FCoE protocol headers
  - [http://open-fcoe.org/](http://open-fcoe.org/)
- Not just any old NIC
  - Must be CEE capable
FCoE Topologies
FCoE Topologies

- Today’s Topologies (FC Initiator to FC target)
- FCoE initiators direct to FCoE switch
- FCoE through 1Gb E to 10Gb E cloud
- FCoE through 10Gb E cloud
- FCoE Targets (from FCoE initiators)
- FCoE Targets (from FC initiators)
- The SNW Demo Topology
Discrete Networks

Ethernet

WWW

FC Fabric

FC Initiator

FC Target
FCoE – Topologies (add FCoE Initiators directly to switch)
FCoE – Topologies (add FCoE Initiators through Ethernet cloud)

*FCoE enabled switch*

*Ethernet*

*WWW*

*Lossless Ethernet cloud (1Gb or 10Gb)*

*FC Initiator*

*FC Target*
FCoE – Topologies (add FCoE Targets in the cloud)
**FC + FCoE = SINGLE SAN**

- **FCInitiator**
- **FCTarget**
- **WWWW**
- **FCoEInitiator**
- **FCFabric**
- **FCoEenabled switch**
- **Ethernet**
- **Consolidation without management issues**
- **Stateless, just encapsulates and decapsulates**
- **Standard FC Session**
- **Single namespace, single management space**

**FC**

**FCoE**

**Target**

**Initiator**

**WWW**

**FCoE**
FCoE with FC Targets
FCoE with FCoE Targets

Classical Ethernet

Internet

Lossless Ethernet

FCoE

1.2.1 1.2.2 1.3.1

1.2.20

Lossless Ethernet

Fibre Channel

2.2.1 2.2.2 2.3.2

6.2.20 6.31.1

3.2.2 3.3.1 9.4.1 9.2.2
FCoE Host Perspective
System Implications

Systems can contain fewer HBAs (can use lower slot count systems); use less power, requires less A/C. Lower installation and management costs.
Cabling Savings in 16 Server rack

Servers Using FC
- 4+ cables per server
- 32+ FC server cables
- 32+ Ethernet server cables
- 4 FC uplink cables
- 8+ Ethernet uplink cables
- 76+ total cables

Servers Using FCoE
- 2 cables per server
- 32 FCoE + 10GbE server cables
- 4 FC uplink cables
- 4 Ethernet uplink cables
- 40 total cables
FCoE Protocol Layering

- The Protocol layering contributes to the high degree of transparency
- Applications, Filesystems, SCSI, FCP (and other FC-4 protocols) all remain the same
Fibre Channel Protocol layering

- SCSI Protocol remains the same
- Basic Fibre Channel Protocol remains the same
- Fibre Channel Name Services remain the same
- Fibre Channel Zoning remains the same
- Only the Physical layer changes
## Sample Host S/W stack

<table>
<thead>
<tr>
<th>File System</th>
<th>(UFS, ZFS, VxFS, CFS etc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume Manager</td>
<td>(SVM, VxVM, CVM etc)</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>(SCSI Disk driver)</td>
</tr>
<tr>
<td><strong>SG</strong></td>
<td>(SCSI Generic driver)</td>
</tr>
<tr>
<td><strong>ST</strong></td>
<td>(SCSI Tape driver)</td>
</tr>
<tr>
<td><strong>SCSI</strong></td>
<td>(SUN calls this SCSA)</td>
</tr>
<tr>
<td><strong>MPXIO</strong></td>
<td>(aka STMS, scsi vhci)</td>
</tr>
<tr>
<td><strong>FCP, FCSM</strong></td>
<td>(FCP driver, FC SAN Mgmt driver)</td>
</tr>
<tr>
<td><strong>FP, FCTL</strong></td>
<td>(FC Port Driver, FC Transport Layer driver)</td>
</tr>
<tr>
<td><strong>emlxs</strong></td>
<td>(SUN Native driver for Emulex HBAs)</td>
</tr>
<tr>
<td><strong>qlc</strong></td>
<td>(SUN Native driver for QLogic HBAs)</td>
</tr>
</tbody>
</table>

**FC HBAs**
### Sample Host S/W stack + FCoE

<table>
<thead>
<tr>
<th>File System</th>
<th>(UFS, ZFS, VxFS, CFS etc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume Manager</td>
<td>(SVM, VxVM, CVM etc)</td>
</tr>
<tr>
<td>SD</td>
<td>(SCSI Disk driver)</td>
</tr>
<tr>
<td>SG</td>
<td>(SCSI Generic driver)</td>
</tr>
<tr>
<td>ST</td>
<td>(SCSI Tape driver)</td>
</tr>
<tr>
<td>SCSI</td>
<td>(SUN calls this SCSA)</td>
</tr>
<tr>
<td>MPXIO</td>
<td>(aka STMS, scsi_vhci)</td>
</tr>
<tr>
<td>FCP, FCSM</td>
<td>(FCP driver, FC SAN Mgmt driver)</td>
</tr>
<tr>
<td>FP, FCTL</td>
<td>(FC Port Driver, FC Transport Layer driver)</td>
</tr>
<tr>
<td>emlx</td>
<td>(SUN Native driver for Emulex HBAs)</td>
</tr>
<tr>
<td>qlc</td>
<td>(SUN Native driver for QLogic HBAs)</td>
</tr>
</tbody>
</table>

**FCoE HBAs**

Everything else remains the SAME

**Updated Drivers**
FCoE Host Administration

Typical SAN

Storage Admin
Accessing Host & FC Switches

FC Edge

FC Initiator

FC Initiator

SAN

FC Core

FCoE SAN

Storage Admin
Accessing Host & FC Switches

FC Initiator

FC Initiator

FCoE Edge

FC Initiator

FCoE

SAN

FC Core
Host Side Summary

- Virtually unchanged device drivers
- Existing file systems and applications work
  - No change to higher layers required
- Existing networking applications work
- Converged traffic is a reality
Linux S/W FCoE

```
[root@rhelinstall ~]# fconf
Name      WWPN                  FC-ID   FC  Network
eth0      20:00:00:1b:21:19:39:7a  c80009 Online  UP
eth1      20:00:00:14:5e:17:bf:c0  0    No HBA  UP
eth2      20:00:00:14:5e:17:bf:c2  0    No HBA  DOWN
[root@rhelinstall ~]# fconf target
eth0 target ports:
FC ID      WWPN                  State Role
  c80008    50:0a:09:38:83:44:3c:84:44 Online FCP Target
[root@rhelinstall ~]# fconf lun
eth0 LUN mappings:
Device FC-ID LUN LUN ID Capacity Vendor Model
  sdb    c80008 0   -   10.00 GB NETAPP LUN
  sdc    c80008 1   -   10.00 GB NETAPP LUN
```
Showing 2 FCoE + 2 FC ports

Head1 > fcp show adapter
Slot: 0c
Description: Fibre Channel Target Adapter 0c (Dual-channel, QLogic 2432 (2462) rev. 2)
Adapter Type: Local
Status: LINK NOT CONNECTED
FC Nodename: 50:0a:09:80:87:49:2b:7d (500a098087492b7d)
FC Portname: 50:0a:09:81:87:49:2b:7d (500a098187492b7d)
Standby: No

Slot: 0d
Description: Fibre Channel Target Adapter 0d (Dual-channel, QLogic 2432 (2462) rev. 2)
Adapter Type: Local
Status: LINK NOT CONNECTED
FC Nodename: 50:0a:09:80:87:49:2b:7d (500a098087492b7d)
FC Portname: 50:0a:09:82:87:49:2b:7d (500a098287492b7d)
Standby: No

Slot: 4a
Description: Fibre Channel Target Adapter 4a (Dual-channel, QLogic 2432 (8432) rev. 3)
Adapter Type: Local
Status: LINK NOT CONNECTED
FC Nodename: 50:0a:09:80:87:49:2b:7d (500a098087492b7d)
FC Portname: 50:0a:09:85:87:49:2b:7d (500a098587492b7d)
Standby: No

Slot: 4b
Description: Fibre Channel Target Adapter 4b (Dual-channel, QLogic 2432 (8432) rev. 3)
Adapter Type: Local
Status: LINK NOT CONNECTED
FC Nodename: 50:0a:09:80:87:49:2b:7d (500a098087492b7d)
FC Portname: 50:0a:09:86:87:49:2b:7d (500a098687492b7d)
Standby: No

Head1 >
A Specific (Generation 1) FCoE Example – the proof of concept
Updated Driver Source Code

from diff --c config.c.orig config.c:

```
    VID_QLOGIC,   DID_QLOGIC_EP2432,   BT_SCSI_HA,
    NULL,        ispfc_common_init,   ispfc_common_info,   TRUE
},
+ {   /* QLogic ISP 8042 FCoE card */
+    VID_QLOGIC,   DID_QLOGIC_FCOE,   BT_SCSI_HA,
+    NULL,        ispfc_common_init,   ispfc_common_info,   TRUE
+ },
{   /* LSI 929X */
    VID_LSI,      DID_LSI_929X,       BT_SCSI_HA,
    NULL,        mptfc_init,         mptfc_info,   TRUE
```
Updated Driver Source Code

from ispfct_driver.c:

/*
* This routine is used to send the iocb to the ISP so it will verify the firmware in the FCoE chip. If this is successful,
* then the FCoE login sequence will begin and we are ready to go. Without this, the FCoE login will not even start.
*/
Static int
isp2400fct_verify_fcoe_fw_iocb( ispfct_cb_t * cb )
{
    /*
    * Start with the cb (control block for the HBA) and get a pointer to a command request structure for this HBA.
    */
    isp2400_fcoe_mgmt_iocb *fcoe = (isp2400_fcoe_mgmt_iocb *) &(cb->request[cb->request_in]);
    /*
    * Fill in the command entry with the fcoe enable function.
    */
    bzero(fcoe, sizeof(isp_command_Entry)); /* zero everything */
    fcoe->Header.entry_type = ISP_ENTRY_FCOE_FW;
    fcoe->Header.entry_cnt = 1;
    fcoe->Options = ISP2400_FCOE_START;
    isp2400fct_send_iocb( cb );
    return ( THA_SUCCESS );
}
Updated Driver Source Code

from ispfct_driver.c (ispfct_initialize_adapter() routine):

...  
if (dp->deviceld == DID_QLOGIC_FCOE) {
   /*
   * If we have a FCoE chip, then we need to
   * kick start it by sending the verify iocb.
   */
   isp2400fct_verify_fcoe_fw_iocb(cb);
}
...

The Real World

- The proof of concept demo did not use FIP
- Standard FCoE will use FIP during initialization
  - Real World driver changes will involve more than just shown
  - FIP involvement may be needed
- FIP will not impact mainline I/O or applications
Security Considerations
FCoE Security Considerations

- FCoE is no more secure than Ethernet
- FCoE is no more secure than your FC network
- FCoE has no new Security Features
- FCoE uses existing Security Features
FCoE Security Considerations

- FC-SP
  - Used for Authentication at the FC layer
  - Transparent to physical connection layer; therefore works over FCoE
  - Consistent Management across both FC and FCoE physical layers
FCoE Security Considerations

- Ethernet ACLs (Access Control List)
  - Automatic setup (via FIP snooping)
    <more later>
  - Manual setup
    - Allow FIP packets to go to FCFs:
      - DA = ALL_FCF_MACs, Type = FIP_TYPE, permit;
      - DA = {FCFs}, Type = FIP_TYPE, permit;
      - Type = FIP_TYPE, deny;
      - <note: you must know all FCF addresses>
FCoE Security Considerations

- Manual setup (continued):
  - Prevent non-FIP before FLOGI:
    - Type = FCoE_Type, deny;
    - SApre = FC_MAP, deny;
  - Prevent E-node to E-node:
    - DApre = FC-MAP, deny;
  - Enable traffic after FLOGI:
    - SA = FCF assigned MAC address, DA = FCF MAC address,
      Type = FCoE, permit;
    - <note: usually at top of ACE list>
Manual ACL Summary

- Prior to receipt of any Discovery Advertisements, the initial ACL is:
  - DA = ALL_FCF_MACs, Type = FIP_TYPE, permit;
  - Type = FIP_TYPE, deny;
  - Type = FCoE_Type, deny;
  - <followed by Any non-FCoE related ACEs>

- As the result of administrative FCoE address knowledge, the ACL is expanded to:
  - DA = ALL_FCF_MACs, Type = FIP_TYPE, permit;
  - DA = {FCFs}, Type = FIP_TYPE, permit;
  - Type = FIP_TYPE, deny;
  - Type = FCoE_Type, deny;
  - SApre = FC_MAP, deny;
  - DApre = FC-MAP, deny;
  - <followed by Any non-FCoE related ACEs>

- For each successful FLOGI (or FDISC), an ACE is added (prior to SApre = FC-MAP, deny;) of the form:
  - SA = FCF assigned MAC address, DA = FCF MAC address, Type = FCoE, permit;
Automatic ACL Summary

- Ethernet ACLs
  - Automatic setup (via FIP snooping)
    - Switch examines FIP protocol and automatically builds the appropriate ACLs for you.
    - Only available to an FCoE aware Ethernet switch (so it can examine the FIP packets).
Summary
Summary

- FCoE is easily added to existing FC environments
  - Enabling Concurrent use, or
  - Enabling Migration as desired
- Contributes to reduced host adapter count, cable count, power usage, and A/C usage
- Caveats:
  - Standard not done, targeting October complete with April 2009 approval.
  - IEEE CEE activities targeting summer 2009
  - Gen 1 H/W vs. Gen 2 H/W vs. S/W