Ethernet Fabric Requirements for FCoE in the Data Center

Gary Lee
Director of Product Marketing

glee@fulcrummicro.com
FCoE Market Overview

- FC networks are relatively high cost solutions provided by a limited number of vendors
- The data center is moving toward converged Ethernet fabrics (DCB, DCE, CEE etc)
- Several vendors are providing FCoE CNAs today
- FCoE storage array targets will be next
- Ethernet switches with FC ports
  - Needed today to connect to legacy FC fabrics
  - Requirement will disappear over time
- More important are Ethernet switches with DCB features
FCoE Fabric Requirements

- 10GbE data rates
  - Required to support existing FC bandwidth requirements
- Lossless operation
  - Storage protocols cannot tolerate retransmission latencies
- Bounded latency through the fabric
  - Cannot tolerate delays due to blocking in fabric
- Control plane support
  - FIP snooping for discovery and port associations
- Data plane forwarding
  - Forwarding tables must be supported in the switch
- FCoE endpoint requires FCF encapsulation services
  - Only required on certain datacenter fabric endpoints
Latency is Important in Data Center Networks

- **Application Performance**
  - Deterministic Forwarding Delay
  - Low Jitter

- **Output Queued Architecture**
  - Memory efficient

- **Scaleable**
  - Latency per stage adds up

- **Congestion Management**
  - High Priority, Low latency Feedback Path

### Latency Comparison (Cut-through vs. Store-n-forward)

<table>
<thead>
<tr>
<th>Packet Size (Bytes)</th>
<th>Latency (μS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1500</td>
<td>10</td>
</tr>
</tbody>
</table>

- **Cut-through**
- **Store-n-forward**
Data Center Bridging and FCoE

- **Priority Flow Control (PFC)**
  - Provides lossless operation
  - Required for FCoE fabrics

- **Enhanced Transmission Selection (ETS)**
  - Provides bounded latency
  - Required for FCoE fabrics

- **Quantized Congestion Notification (QCN)**
  - Flow control in large scale fabrics

- **Data Center Bridging Exchange Protocol (DCBx)**
  - Allows switches to exchange capabilities
PFC Provides Lossless Ethernet for FCoE

- Priority Flow Control (PFC)
  - ACL rules can be used to classify FCoE traffic
  - FCoE traffic can be assigned to a special memory partition

- At the switch ingress
  - PFC frames are generated when shared memory watermarks are crossed in a given memory partition providing lossless operation for FCoE

- At the switch egress
  - Enhanced Transmission Selection is supported (ETS)
  - QCN frames minimize multistage congestion
ETS Provides Bounded Latency for FCoE

- Priority Flow Control
  - Received PFC frames can be used for egress scheduling
    - Up to 8 CoS queues per each egress port

- Egress Scheduling Example
  - Strict priority
    - Priority groups
    - WRR within group
  - Deficit Round Robin
    - DRR groups
  - Traffic shaping
    - Shaping groups

San José, CA USA
February 2010
FCoE Forwarding support using a TCAM

- The FCoE packet can be parsed as a non-L3; EtherType = FCoE
- L2 deep inspection can be enabled to pass data to the TCAM
- Forwarding decisions can be based on FC header
- When TCAM finds a match in the FC_ID destination field, the action may produce an index into an ARP table
- ARP table is used to maintain next hop MAC address similar to IP routing
- Frame DMAC is replaced with next-hop MAC from ARP table and SMAC is replaced by switch MAC
FCoE Traffic Forwarding

- **ACL rules can be used to identify FCoE traffic**
  - Maps FCoE frames to a unique traffic class
  - FCoE traffic is mapped to storage memory partition and egress queue
  - Other traffic is mapped to other memory partitions

- **Lossless operation**
  - IEEE PFC frames can be generated per memory partition at ingress
  - Data traffic can be paused without effecting storage traffic

- **Forwarding using L3 routing (ARP) table**
  - ACL rules can check for DMAC of Fibre Channel Forwarder (FCF) and ethertype
  - VLAN and D_ID can be used to lookup new DMAC and VLAN from ARP table
  - SMAC is set to MAC of FCF

- **Egress schedulers guarantee bounded latency**
  - DDR scheduler can give FCoE minimum bandwidth guarantee
  - Egress shapers can be used to limit data traffic to downstream switches
  - Supports IEEE PFC frames to differentiate FCoE traffic

- **Baby jumbo frame support**
  - Baby FCoE jumbo frames are about 2.5KB
  - Cut-through operation improves system performance
FC Initialization Protocol (FIP) Snooping

- Control plane mechanism for endpoint discovery
  - Separate Ethertype for discovery and login
  - Works using existing FocalPoint switch mechanisms
- Redirect FIP frames to switch supervisor port using ACL rules
  - Mirror to CPU port or
  - Mirror to Ethernet port
- CPU monitors control plane and updates switch
  - Update ACL rules to recognize FCoE frames by DMAC, SMAC and ethertype
  - VLAN and FCF discovery
  - Update ARP table entries based on FCF ports
  - Virtual link instantiation
  - Virtual link maintenance
FCoE Solution Example

- **FCoE to FC Bridge**
  - Supports legacy FC
  - Only at the edge of future networks
  - Acts as FCF

**Future Networks**
- Storage arrays with FCoE ports
- Servers with CNAs

Diagram:
- 288-port 10GbE Fabric
- CEE/DCB Features
- FocalPoint Core Switch
- FocalPoint Edge Switch
- FCoE to FC Bridge
- Legacy FC Switch
- Legacy FC Storage Array
- NIC
- HBA
- Legacy Servers
- Server
- FCoE Storage Array
- CNA
Fulcrum Company Snapshot

Located in Calabasas, CA

Formed out of Caltech

Shipping low-latency 10G switching silicon

Located in Calabasas, CA

Backed by top-tier investors

San José, CA  USA
February 2010
FocalPoint Product Overview

- High performance switching and routing
  - 24 ports, 200ns latency, 360 Mpps
  - 300ns latency for L3 with L3/L4 ACLs enabled
  - 2MB shared packet memory
- IPv4 & IPv6 routing
  - 16k IP addresses
- L2-L4+ ACLs with deep inspection
  - At full line-rate on all ports
- Scalable architecture
  - Non-blocking fat tree support
  - QCN support
- Converged Enhanced Ethernet (CEE)
  - PFC for lossless operation
  - ETS for bandwidth guarantees
  - DCBx support
  - FCoE support
Conclusions

- Converged fabrics will replace FC in the data center
- FCoE requires some key DCB features
  - Priority Flow Control (PFC)
  - Enhanced Transmission Selection (ETS)
- Deep header inspection is required for FCF
- FocalPoint provides and efficient scalable solution for FCoE in the data center