



FCoE aware Ethernet Switches

T11/07-694v0, December 2007
JR Rivers



Background

Concepts discussed in this presentation are independent of addressing schemes

- Mapped addresses used as an example

Goal is robust, interoperable transit Ethernet switch solutions

May be deployed on many existing Ethernet switch platforms

Existence of FCoE will cause an evolution of Ethernet switches

Ethernet Switching beyond IEEE

Basic switch defined in IEEE 802.1 and 802.3

Today's enterprise and data center switches are Upper Layer Protocol aware

- Not an unmanaged switch from Circuit City...
- This includes all form factors – modular, fixed, blade

Additional functionality defined by other bodies

- IGMP Snooping (IP Multicast) – rfc4541
- DHCP Snooping (Relay Information Option 82) – rfc3046

Mandatory Transit Switch Requirements

Lossless Ethernet - 802.3X PAUSE

- Some might argue “Priority Pause”

Large MTU support

- ~2.5KBytes to transport 2112 byte FC payloads

Maximum bridge transit time

- Mandatory to have known, enforced value
- Desired to have “FC scale” time (0.5-1 sec)

Evolving Transit Switch Requirements

Data and Control Plane Integrity

- Protection from **accidental** or intentional Denial-of-Service or Man-in-the-middle attacks
 - Mis-cabling, bad HW, bad SW, etc.
- **Many existing Ethernet switch HW platforms meet these requirements**
 - Some may require firmware upgrades

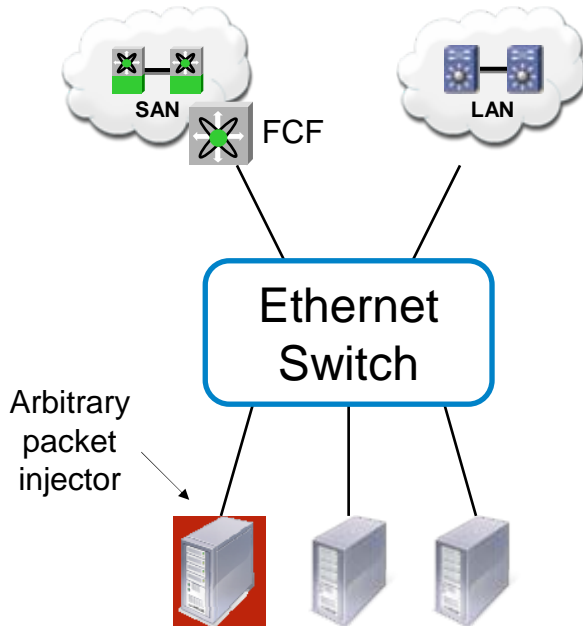
Data Plane Quality-of-Service

- Priority Pause, Congestion Notification (802.1Qau), Link Scheduling
- Evolving standards, likely to require new HW

Management Plane

- Configuration, monitoring, diagnostics, and reporting
- **Firmware solutions with no HW implications**

Focus on Data Integrity



Any scenario that can be created with a wiring or configuration problem can also be created with a “raw” Ethernet packet

Goal: Prevent looping, black-hole, denial-of-service, and spoofing attacks that lead to data corruption, loss, or theft

Many scenarios to evaluate

- Complete analysis with “arbitrary packet injector”
 - “System breaks if it sees a packet like this...”

Eventually solution must be dynamic

- Manual configuration burdens deployment and cost time/money
 - Example: Replacing a bad server blade

Solution should...

- Rely on normal switch forwarding operation
- Avoid deep packet inspection in data plane

Solution should be defined by FC-BB-5 for multi-vendor interoperability

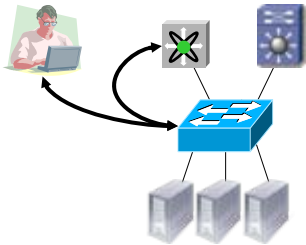
- Normative or informative as appropriate

Data Integrity Solutions



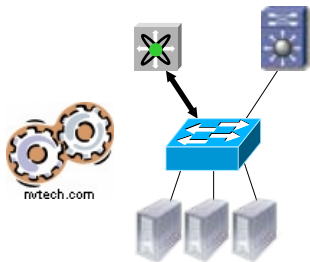
Ignore the problem

- Reasonable for proof-of-concept and small deployments
- Downside – Certification concerns and large scale deployment



Management plane

- Use SNMP, CLI, or API to achieve desired behavior
 - Separate management tool or based in FCF
- Downsides
 - Proprietary solutions/interactions
 - Various levels of scaling and support
 - Enacting trust (credential management)



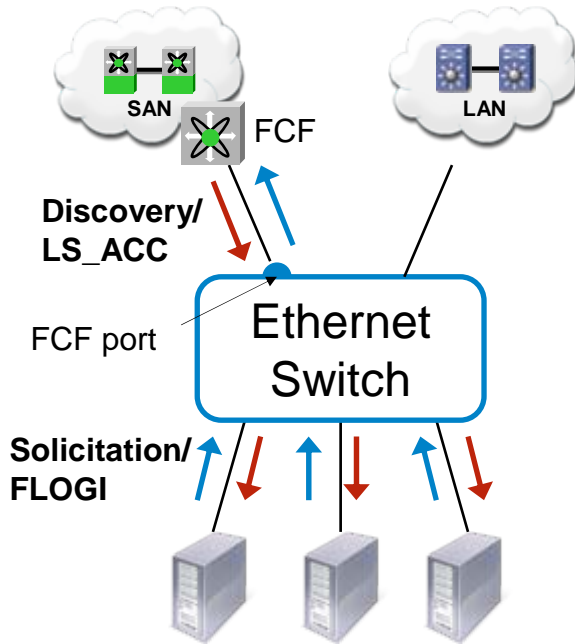
Standard based control plane

- A cog in the machine
- Discussed in this presentation

Natural evolution independent of addressing scheme

Let's learn a lesson from the Ethernet/IP marriage

Dynamic Solution to Data Integrity



Fibre Channel Forwarders are part of network infrastructure

- Configure Ethernet switch ports to FCFs
 - Could also learn from 802.1X

Switch learns Fabric and FCF parameters from Discovery

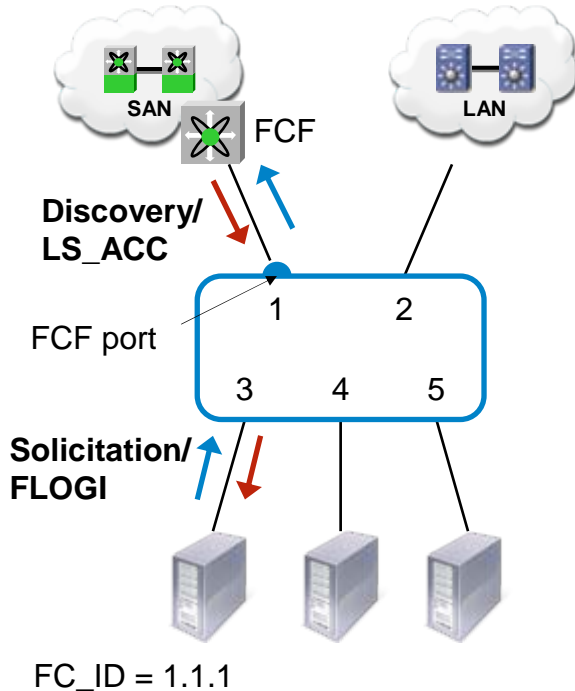
- FCF addresses and FC-MAP value (SAN A/B)

Filtering on non-FCF ports

- Start with full denial in data plane
- Snoop FLOGI exchanges to open holes for valid logins

Works with either addressing scheme!

Mapped Address Example



Ports to Fibre Channel Forwarder(s) manually configured

FC-MAP value and FCF address learned from Discovery

Port 3 FC_ID (1.1.1) learned through FLOGI snooping

Default FCoE filters applied to ports 2, 4, and 5

- Define characteristics, implementation dependent

Permissive FCoE filters applied to port 3

- Define characteristics, implementation dependent

Port 2, 4, and 5 filter example

```
deny mac_sa FC-MAP:00:00:00
    mask 00:00:00:ff:ff:ff
deny mac_sa FCF-MAC
permit mac_da FCF-MAC type FCoE
permit mac_da All-FCFs type FCoE
deny type FCoE
```

Port 3 filter example

```
permit mac_sa FC-MAP:01:01:01
    mac_da FCF-MAC
    type FCoE
deny mac_sa FC-MAP:00:00:00
    mask 00:00:00:ff:ff:ff
deny mac_sa FCF-MAC
permit mac_da FCF-MAC type FCoE
permit mac_da All-FCFs type FCoE
deny type FCoE
```

Filter Technical Underpinnings

Multi-field classifier (MFC)

- Select frames by comparison against known fields
 - Many implementations use ternary bit-wise match... 0, 1, or ?
- HW capability of modern Ethernet switch silicon
 - Supported by multiple vendors

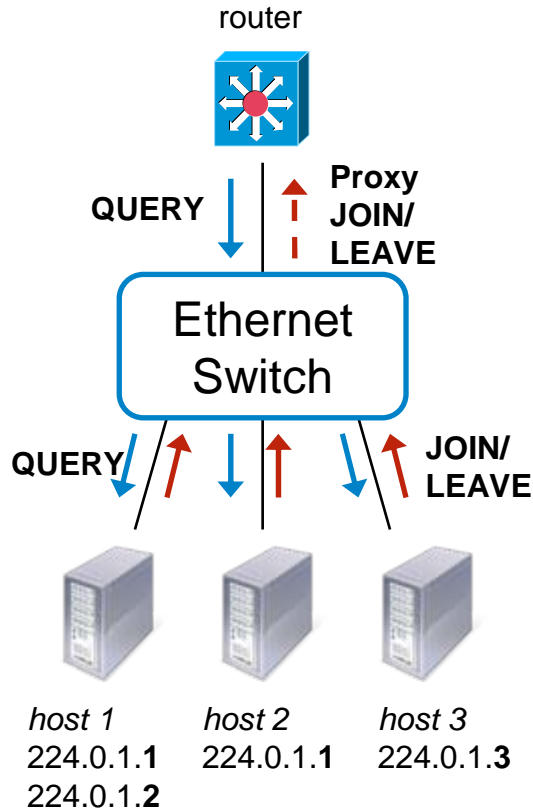
Control plane protocol proxy

- MFC to detect control frames and redirect to supervisor

Data Plane filter

- MFC for filter lists
- 802.1D forwarding table for address “pinning”

Example – IGMP Snooping (rfc4541)



Switch forwards...

224.0.1.1 → hosts 1 and 2

224.0.1.2 → host 1

224.0.1.3 → host 3

Ethernet switch filters IP multicast groups

- Otherwise all hosts see all groups

Switch knows which ports lead to routers

- Configuration
- Protocol snooping

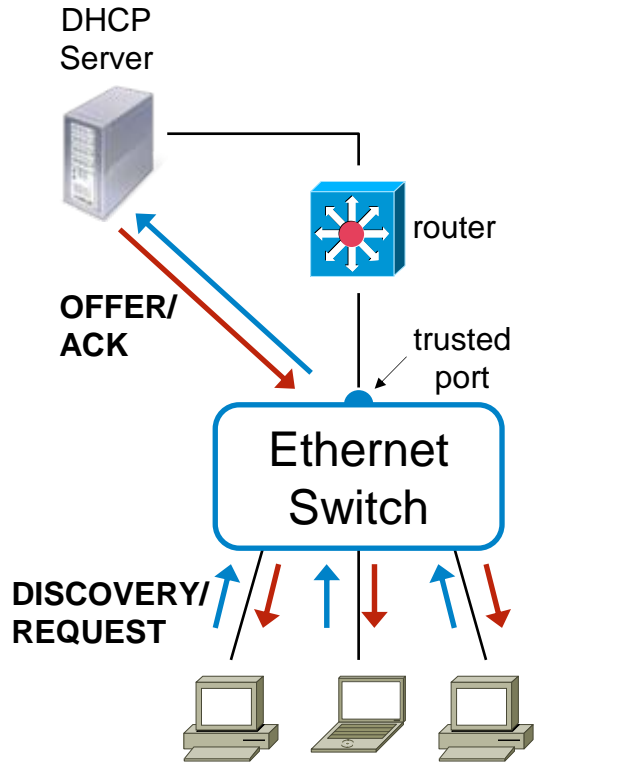
Switch forwards QUERY to all hosts

- Hosts indicate interest in a group (JOIN)

Switch captures JOIN/LEAVE from hosts

- Proxies to router(s)
 - JOIN for first member in group
 - LEAVE for last member out of group
- Forwarding table has static entries with port membership for IP multicast group

Example – DHCP Snooping (rfc3046)



Ethernet switch relays DHCP frames between DHCP client and server to...

- Forward requests to server beyond subnet
- Track allocated IP addresses
 - VLAN, MAC, switch port, office, jack
 - At switch and in DHCP server database
- Prevent DHCP server attacks
 - DOS, spoofing, etc
- Source binding checks
 - Thwart data and control plane DOS/MiM attacks

Switch is configured with list of trusted DHCP servers

Captures requests from untrusted ports

- Evaluates against existing bindings
- Filter invalid DHCP frames

Relays acks from server

- Builds new bindings in data plane

Switch captures DISCOVERY/REQUEST and relays to server

Server sends OFFER/ACK to switch who relays to client

Technical Underpinnings for IGMP/DHCP

IGMP and DHCP Snooping use similar mechanisms

Configuration to define trusted entities

- Casual entities detected through protocol snooping

Control plane protocol proxy

- Multi-field classifier to detect control frames and redirect to supervisor

Data Plane filter

- Multi-field classifier for protocol filtering
- 802.1D static entries for targeted forwarding

Same constructs and techniques may be applied to Dynamic Data Integrity solution for FCoE

Conclusions

Upper Layer Protocol awareness is a part of modern Ethernet switches

Mapped MAC addresses work in dynamic Data Integrity solution

Transit switch functions can be implemented with many existing 10GE switches

- Modular, fixed, blade form-factors
- Multiple vendors (some may require firmware changes)

Call to action... **FC-BB-5 Considerations for Transit Ethernet Switches**



Thank You

