



# **FCoE Aware Ethernet Switches**

## **FIP for a Dynamic Control Plane**

T11/08-079v0, February 2008

JR Rivers



# Background

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

*The first version of this presentation was given in Dec (07-694v0)*

# Updates from December

## Fibre Channel Initialization Protocol (FIP)

- L2 Protocol for end point discovery and fabric association
  - Separate Ethertype for Discovery and Login
- Easily distinguished from FCoE data plane

FIP Snooping in Ethernet switches provides a simple control plane for dynamic FCoE data integrity mechanisms

- Access Control lists
- Static address entries

## Presentation updated to include FIP

- Unchanged content moved to “context” section for reference

# 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

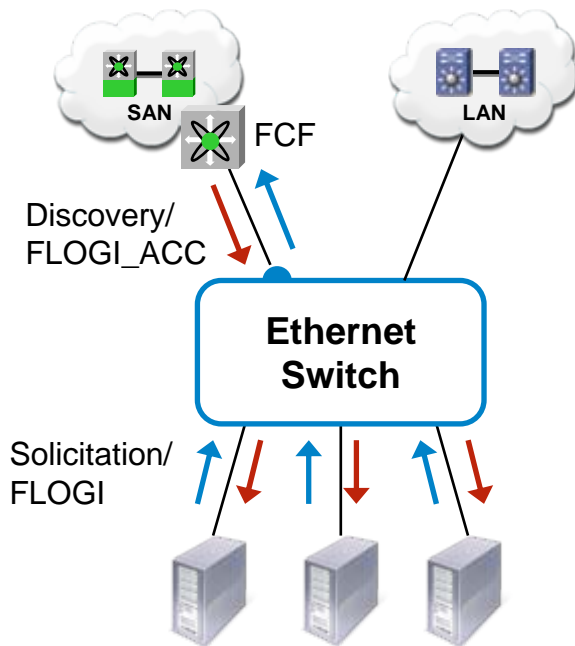
- Not an unmanaged switch from Circuit City...
- This includes all form factors – modular, fixed, blade

Additional functionality defined by other bodies

- IGMP Snooping (IP Multicast) – rfc4541
- DHCP Snooping (Relay Information Option 82) – rfc3046
- **FIP Snooping (FCoE Data Integrity)**

# FIP Snooping

## *A Dynamic Solution to Data Integrity*



Fibre Channel Forwarders are trusted part of network infrastructure

- FCFs can be identified via port, address configuration, 802.1X, etc

Switch learns Fabric and FCF parameters from FIP Discovery

Filtering on non-FCF ports

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

**Works with both FPMA and SPMA addressing schemes!**

# FIP Snooping 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 silicon vendors in shipping products

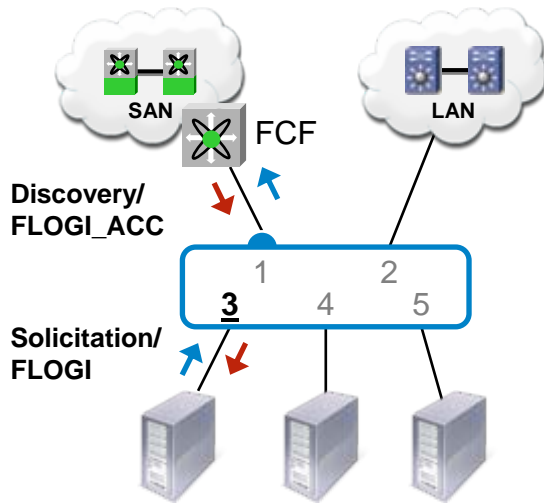
## Control plane protocol intercept

- MFC redirects FIP frames from FCF to switch supervisor
  - Installed on FCF facing ports

## Data Plane filter

- MFC for filter lists
- Static forwarding table entries for “pinning”

# FIP Snooping Example



**After FIP** →

<u>Port Filters</u>	
<b>permit</b>	port 3 mac_sa FIP_ACC_MAC mac_da FCF-MAC type FCoE
<b>redirect</b>	port 1 mac_sa FCF_MAC type FIP
<b>permit</b>	port 1 mac_sa FCF-MAC
<b>deny</b>	mac_sa FCF-MAC
<b>permit</b>	port 1 mac_sa FCF-MAC type FCoE
<b>deny</b>	type FCoE

**Default FCoE** →

## FCoE Forwarding Entries

**ALL-FCFS** : static ports 1,  
**FCF-MAC** : static port 1  
**FIP\_ACC\_MAC** : static port 3

Identify ports to Fibre Channel Forwarder

- Manually configured in this example

Default FCoE filters applied to non-FCF ports

FCF address learned from FIP snooping

- Added to address table entries

FIP\_ACC\_MAC learned through FIP snooping

- Snoop FIP replies from FCF
- From MAC Address TLV in FLOGI ACCEPT frames

Permissive FCoE filters applied to port 3

- Implementation dependent

# Issues to Explore

Transit switch state refresh and timeout

Multiple transit switch hops and link failures

Multiple FCFs and FCF failover

FIP Integrity Requirements

# Conclusions

Upper Layer Protocol awareness is a part of modern Ethernet switches

Dynamic data integrity solutions can be built on top of FIP

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**



# Context from December Presentation

# Mandatory Transit Switch Requirements

## Lossless Ethernet - 802.3X PAUSE

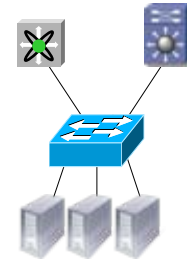
- Some might argue “Priority Pause”

## 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

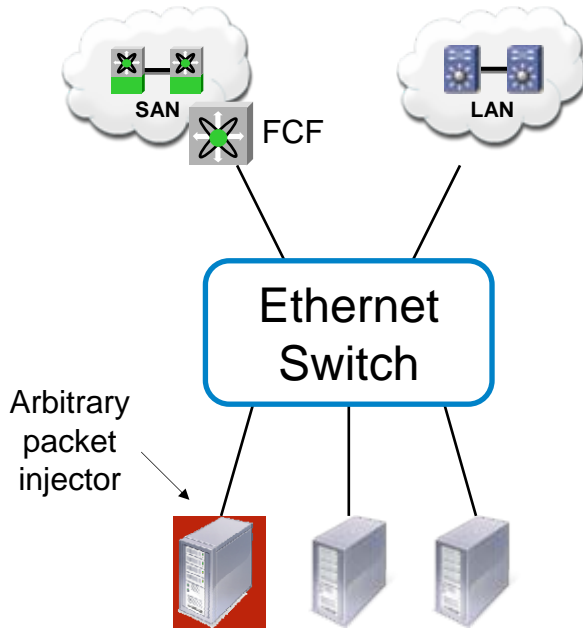
## Data Plane Quality-of-Service

- 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

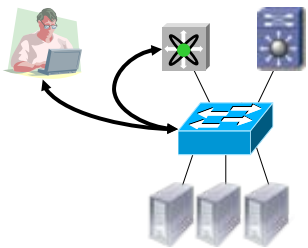
- Normative or informative as appropriate

# Data Integrity Solutions



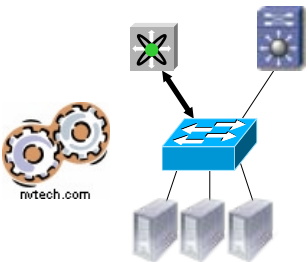
## Ignore the problem

- 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)



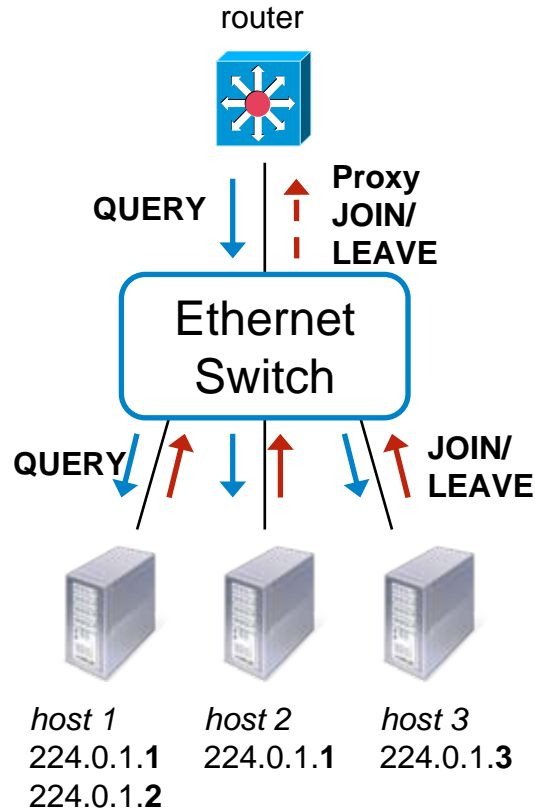
## Standard based control plane

- 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

# 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

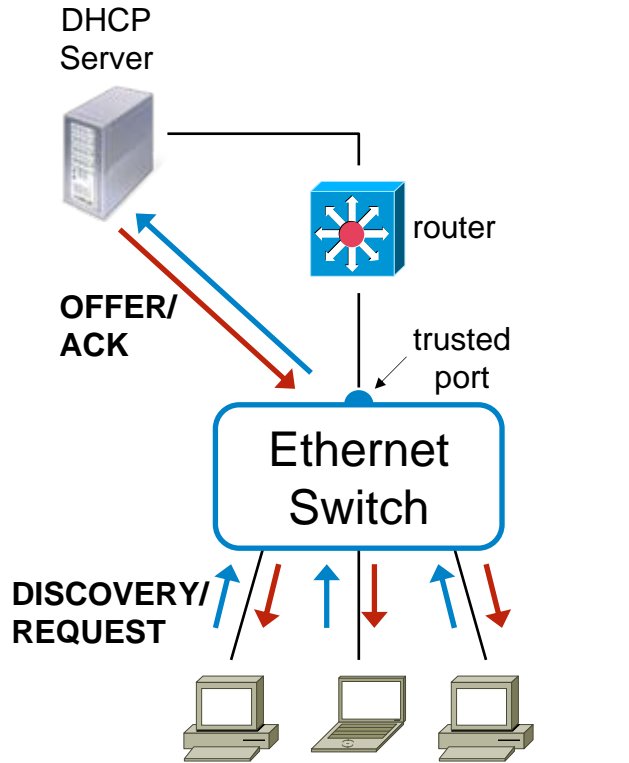
Switch forwards QUERY to all hosts

- 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)



Switch captures DISCOVERY/REQUEST and relays to server

Server sends OFFER/ACK to switch who relays to client

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

# 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**



**Thank You**

