# FCoE sync up

**Liz Stine**

**estine@cisco.com**

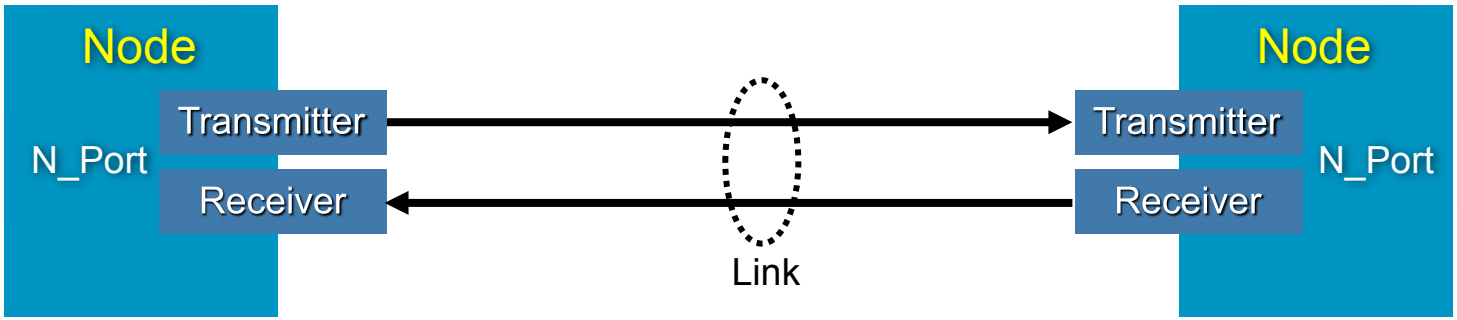
# Agenda

- Brief SAN Technology Overview
- Introduction to Fibre Channel over Ethernet
- Standards Defined
- FCoE Data/Control Plane
- Current and Future FCoE topologies
- FCoE Configuration
- Q & A

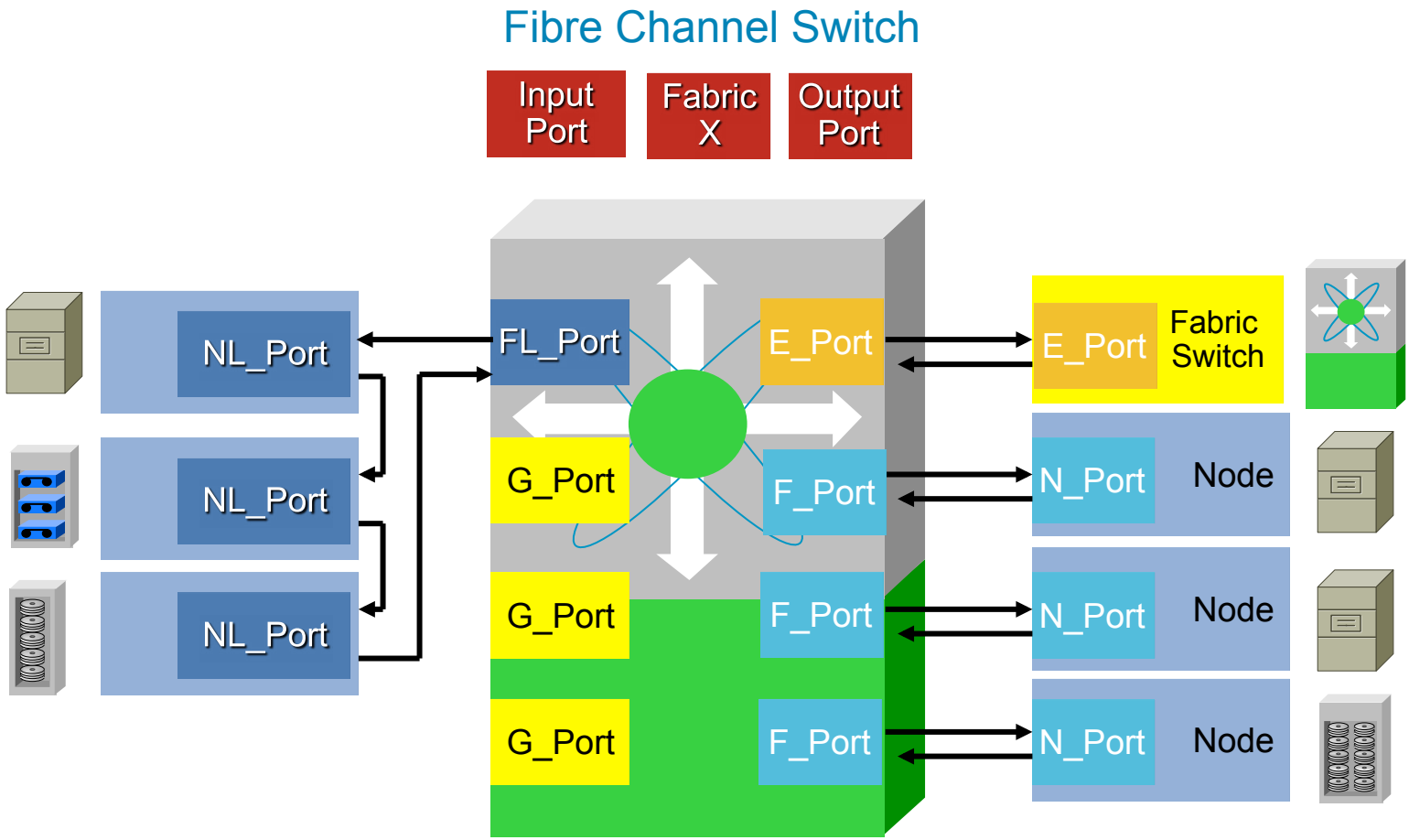
# SAN Technology Overview

# Fibre Channel Communications

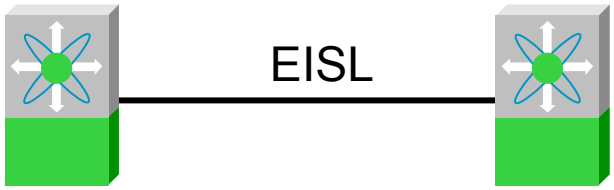
- Point to point oriented
  - Facilitated through device login
- N_Port to N_Port connection
  - Logical node connection point
- Flow Controlled
  - Buffer-to-Buffer Credits and end-to-end basis
- Acknowledged
  - For certain classes of traffic, none for others
- Multiple connections allowed per device



# Fibre Channel Port Types



# Inter-Switch Link (ISL)




- The interconnection between switches is called the ISL

  E_Port to E_Port (‘Expansion port)

- Supports all classes of service

  Class 1, 2, 3, and a special Class F (switch-to-switch)

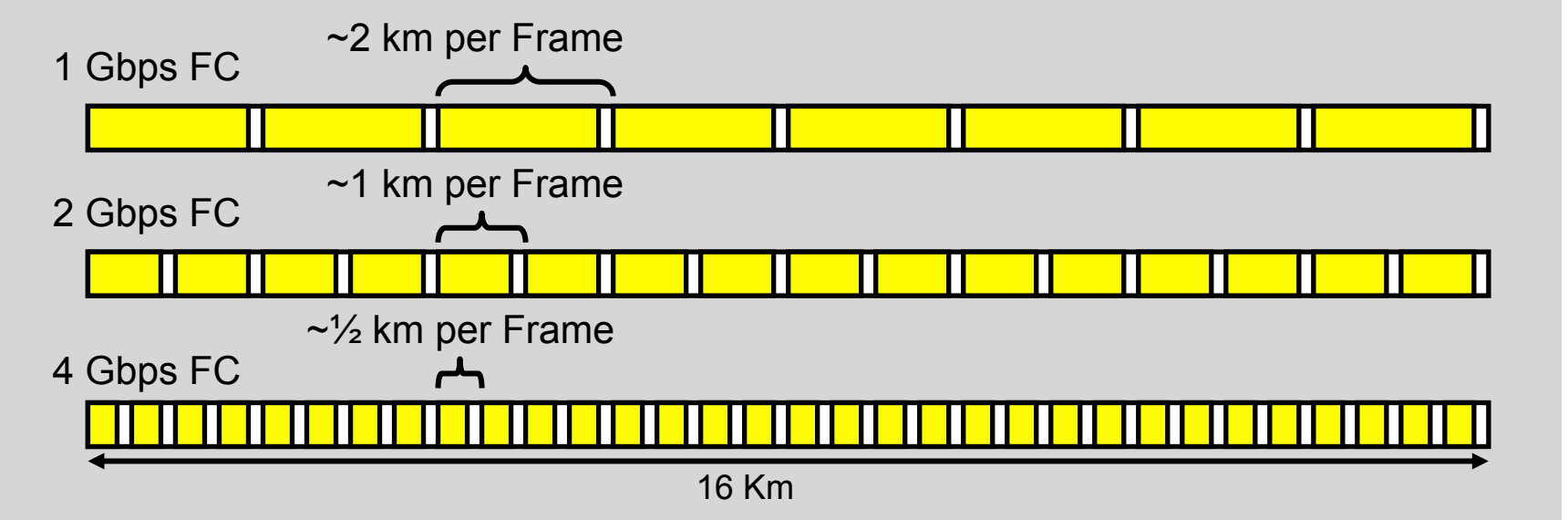
- FC-PH permits consecutive frames of a sequence to be routed over different ISL links for maximum throughput

- Cisco’s implementation is to dedicate an FC_ID pair and/or a given exchange to an ISL bundle member to guarantee in-order delivery for exchange/sequence frames

- Cisco Extended ISL (EISL, TE port)

# Buffer to Buffer Credit Flow Control

## BB_Credits and Distance



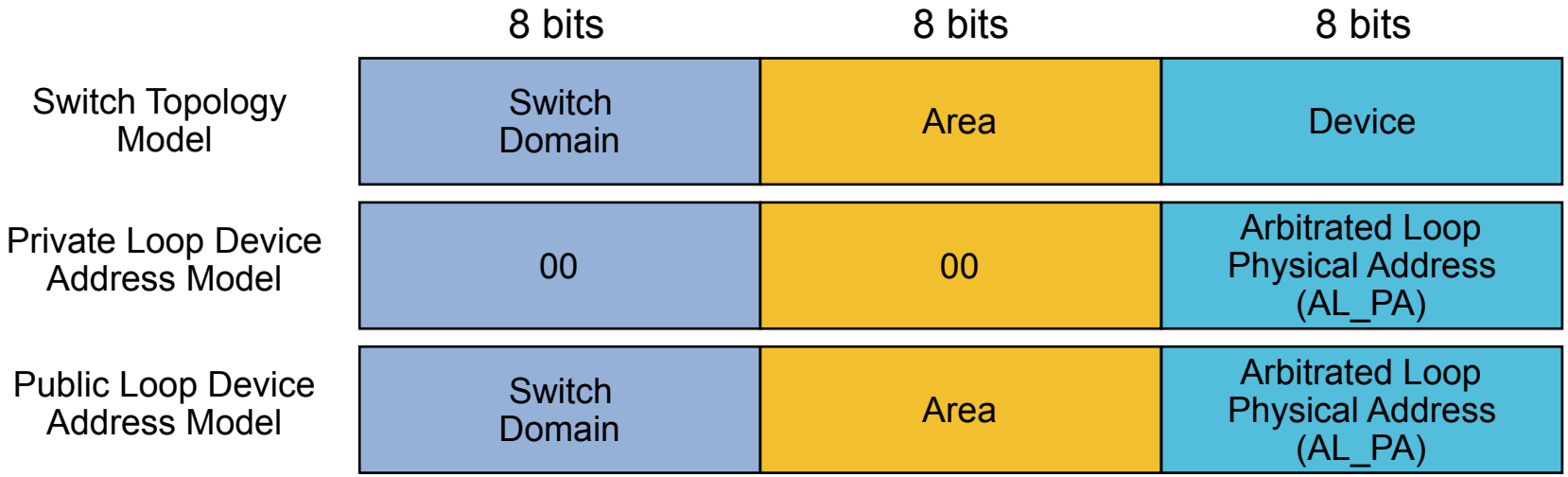

- BB_Credits are used to ensure enough FC frames in flight
- A full (2112 byte) FC frame is approx 2 km long @ 1 Gbps, 1 km long @ 2 Gbps and ½ km long at 4 Gbps
- As distance increases, the number of available BB_Credits need to increase as well
- Insufficient BB_Credits will throttle performance—no data will be transmitted until R_RDY is returned

# FSPF Protocol

- FSPF stands for fabric shortest path first
- Path selection protocol used in fiber channel
- Based on link state protocol
- Fiber channel standard defined in FC-SW2
- Conceptually based on open shortest path first (OSPF) internet routing protocol

# FC_ID Address Model

- FC_ID address models help speed up routing
- Switches assign FC_ID addresses to N_Ports
- Some addresses are reserved for fabric services
- Private loop devices only understand 8-bit address (0x0000xx)
- FL_Port can provide proxy service for public address translation
- Maximum switch domains = 239 (based on standard)



# NPV Enabled Environment

Provides physical port-level virtualization of multiple Fibre Channel end nodes to one F_Port on a Fibre Channel switch:

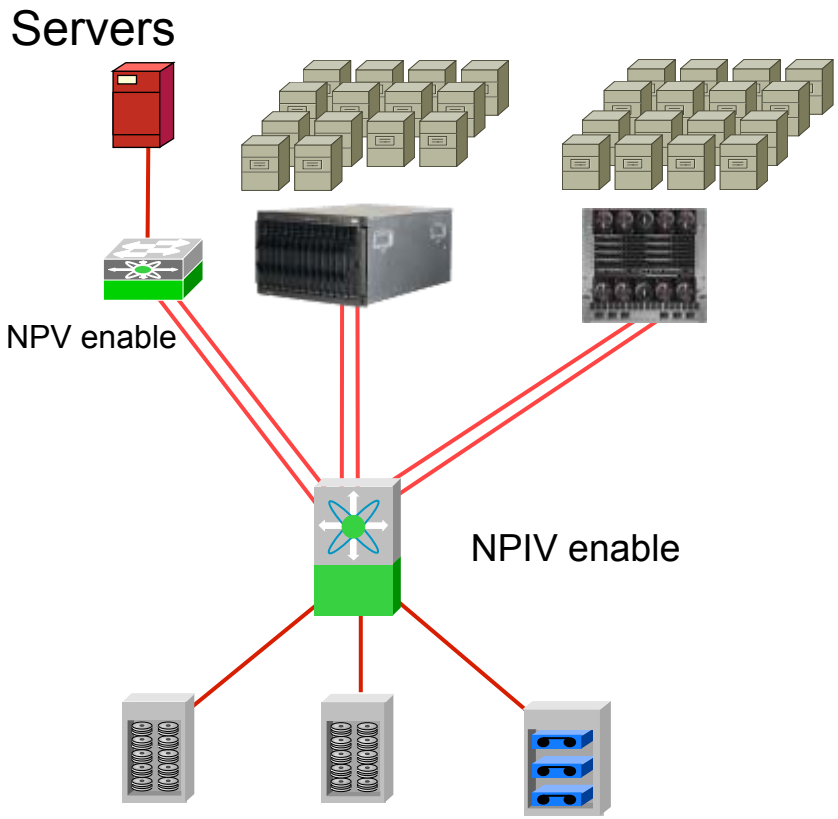
- can operate in N_Port proxy mode and in Fibre Channel switch mode.

NPV Simplifies multivendor interoperation.

Eliminates the Fibre Channel  domain on Each Switch in the fabric.

Simplifies management

Typically used in conjunction with NPIV



# Introduction to FCoE

# What is Fibre Channel over Ethernet?

- From a Fibre Channel standpoint it's

  FC connectivity over a new type of cable called… an Ethernet cloud

- From an Ethernet standpoints it's

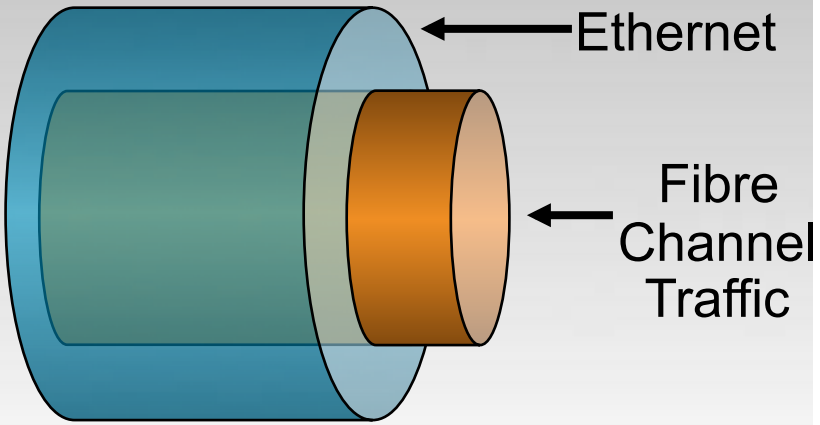
  Yet another ULP (Upper Layer Protocol) to be transported

*FCoE is an extension of Fibre Channel onto a Lossless Ethernet fabric*

# Unified Fabric Overview
## Fibre Channel over Ethernet (FCoE)

### FCoE

- Mapping of FC Frames over Ethernet
- Enables FC to Run on a Lossless Ethernet Network



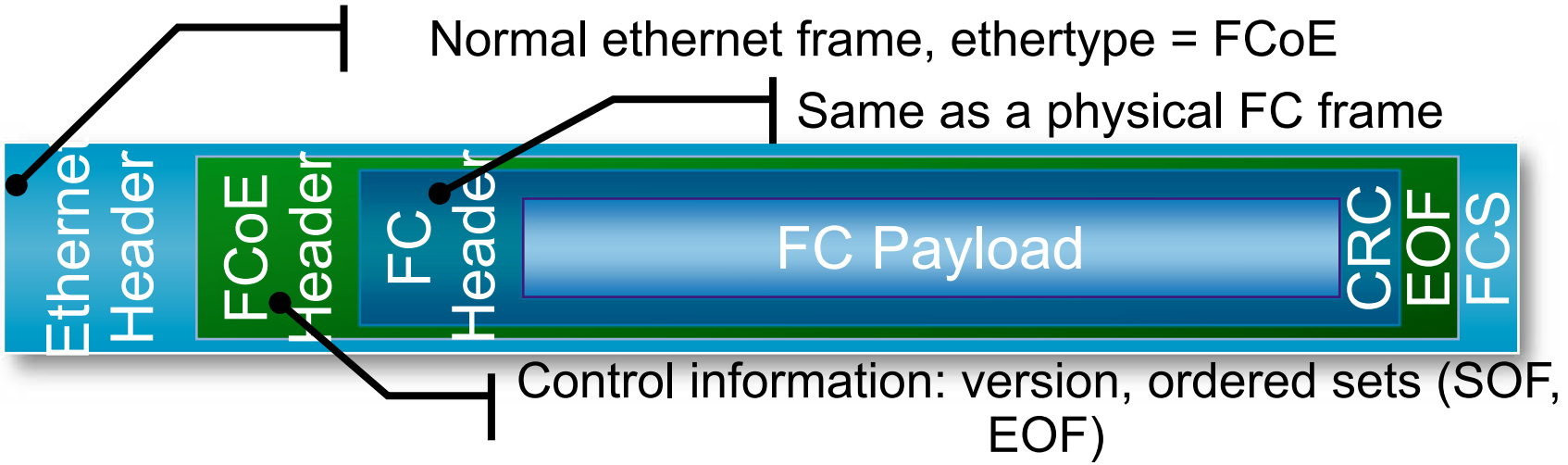

### Benefits

- Fewer Cables
  - Both block I/O & Ethernet traffic co-exist on same cable
- Fewer adapters needed
- Overall less power
- Interoperates with existing SAN's
  - Management SAN's remains constant
- No Gateway

# FCoE Enablers

- 10Gbps Ethernet
- Lossless Ethernet

  Matches the lossless behavior guaranteed in FC by B2B credits
- Ethernet jumbo frames



# FCoE Standards Defined

# What's FC-BB-5

- FC-BB-5 covers the majority of the FC features, using Ethernet
- From an Ethernet perspective, FC-BB-5 is

  Ethernet control plane referred to as FIP (Fibre Channel over Ethernet Initiation Protocol)

  discover and build virtual paths between end points

  Ethernet data plane providing FCoE forwarding

  including both FC control plane and FC data plane (FCF)

# Protocol Organization

***FCoE is really two different protocols:***

## FCoE itself …

- Is the data plane protocol
- It is used to carry most of the FC frames and all the SCSI traffic

## FIP (FCoE initiation protocol)

- It is the control plane protocol
- It is used to discover the FC entities connected to an Ethernet cloud
- It is used to login to and logout from the FC fabric

***The two protocols have:***

- Two different Ethertypes
- Two different frame formats

# What's NOT FC-BB-5

- FC-BB-5 doesn't deal with how lossless is realized in Ethernet

  no Priority Flow Control, Bandwidth Management, etc.

- FC-BB-5 doesn't deal with management functions

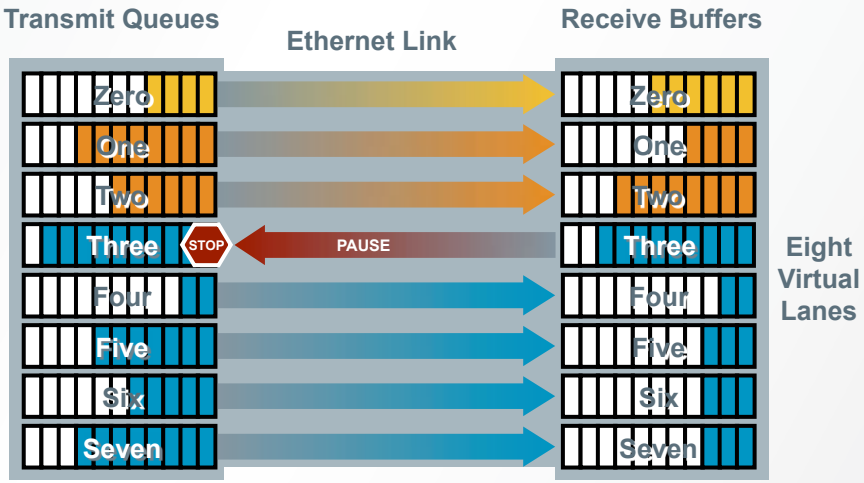
# IEEE DCB standards status

DCB technologies allow Ethernet to be lossless and to manage bandwidth allocation of SAN and LAN flows

| Feature / Standard | Standards Status |
|---|---|
| IEEE 802.1Qbb<br>Priority Flow Control (PFC)<br>Enable multiple traffic types to share a common Ethernet link without interfering with each other | PAR approved<br>Draft 1.0 published, expected WG ballot in 11/09 |
| IEEE 802.1Qaz<br>Bandwidth Management (ETS)<br>Enable consistent management of QoS at the network level by providing consistent scheduling | PAR approved<br>Draft 0.2 published, expected WG ballot in 11/09 |
| Data Center Bridging Exchange Protocol (DCBX)<br>Management protocol for enhanced Ethernet capabilities | This is part of IEEE 802.1Qaz |

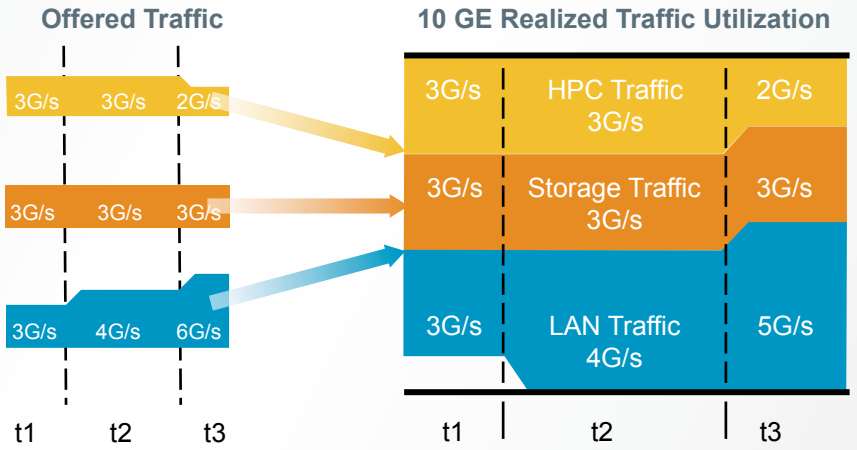
# Data Center Ethernet: PFC & Bandwidth Management

## Priority Flow Control




- *Enables lossless behavior for each class of service*
- *PAUSE sent per virtual lane when buffers limit exceeded*

## CoS based Bandwidth Management




- *Enables Intelligent sharing of bandwidth between traffic classes control of bandwidth*
- *802.1Qaz Enhanced Transmission*

# DCBX Overview

- Auto-negotiation of capability and configuration
  - Priority Flow Control capability and associated CoS values
- Allows one link peer to push config to other link peer
  - Link partners can choose supported features and willingness to accept
- Discovers FCoE Capabilities
- Responsible for Logical Link Up/Down signaling of Ethernet and FC
- DCBX negotiation failures will result in:
  - vfc not coming up
  - Per-priority-pause not enabled on CoS values with PFC configuration

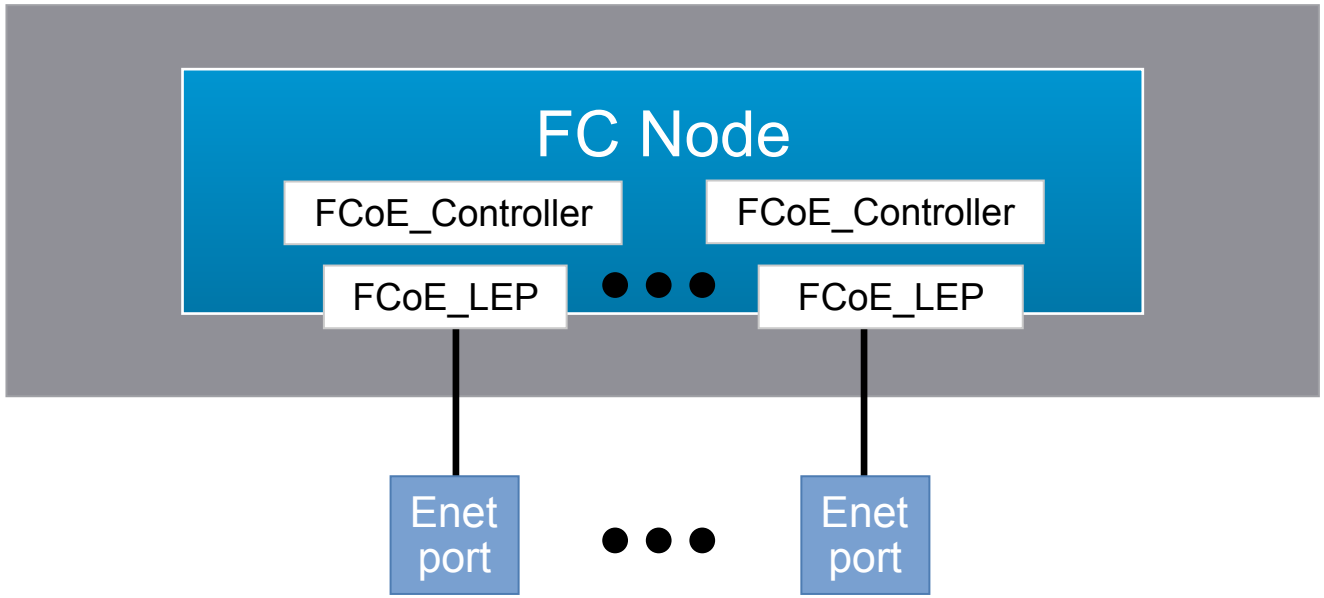
http://download.intel.com/technology/eedc/dcb_cep_spec.pdf

http://www.ieee802.org/1/files/public/docs2008/

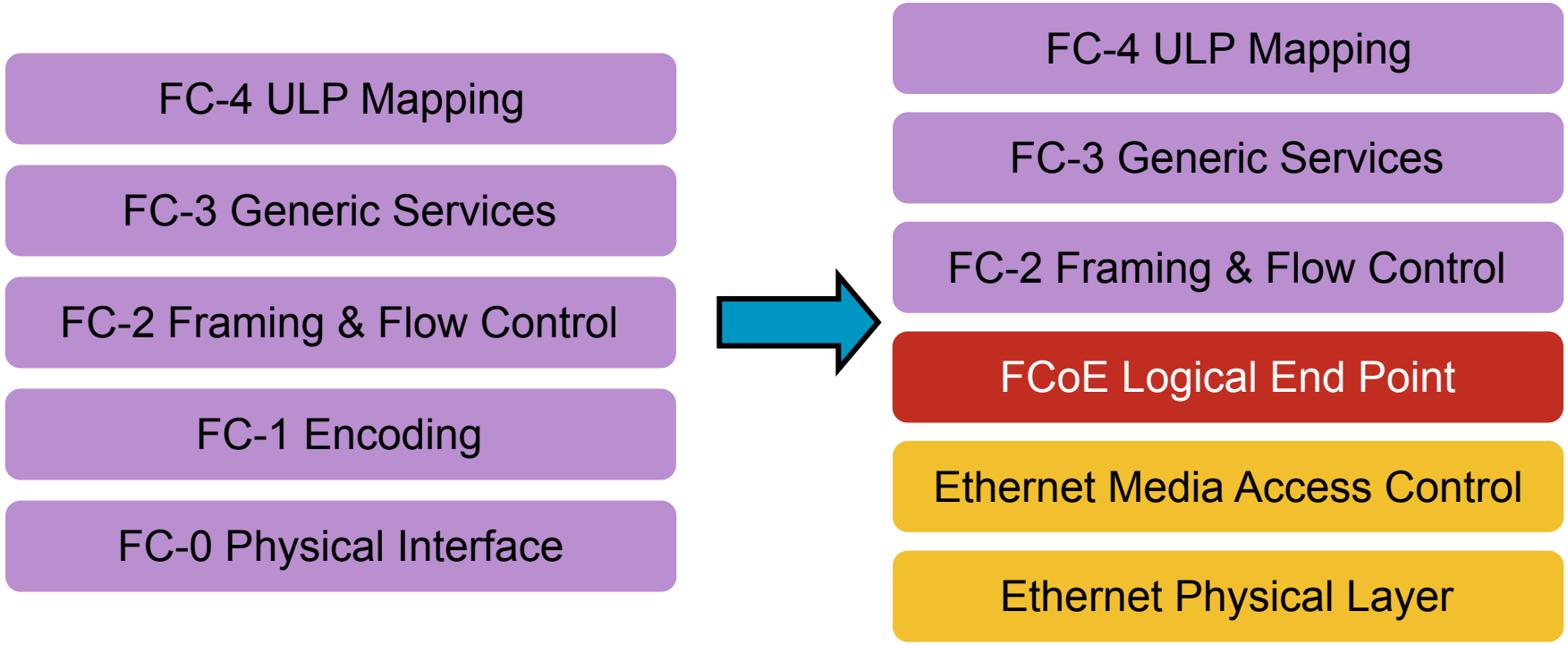
# FCoE data plane

# ENode: Simplified Model

- ENode (FCoE Node): a Fibre Channel HBA implemented within an Ethernet NIC aka CNA (Converged Network Adapter)
- FCoE LEP : The data forwarding component that handles FC frame encapsulation/decapsulation
- FCoE Controller is the functional entity that performs the FIP and instantiates VN_Port/FCoE_LEP pairs.

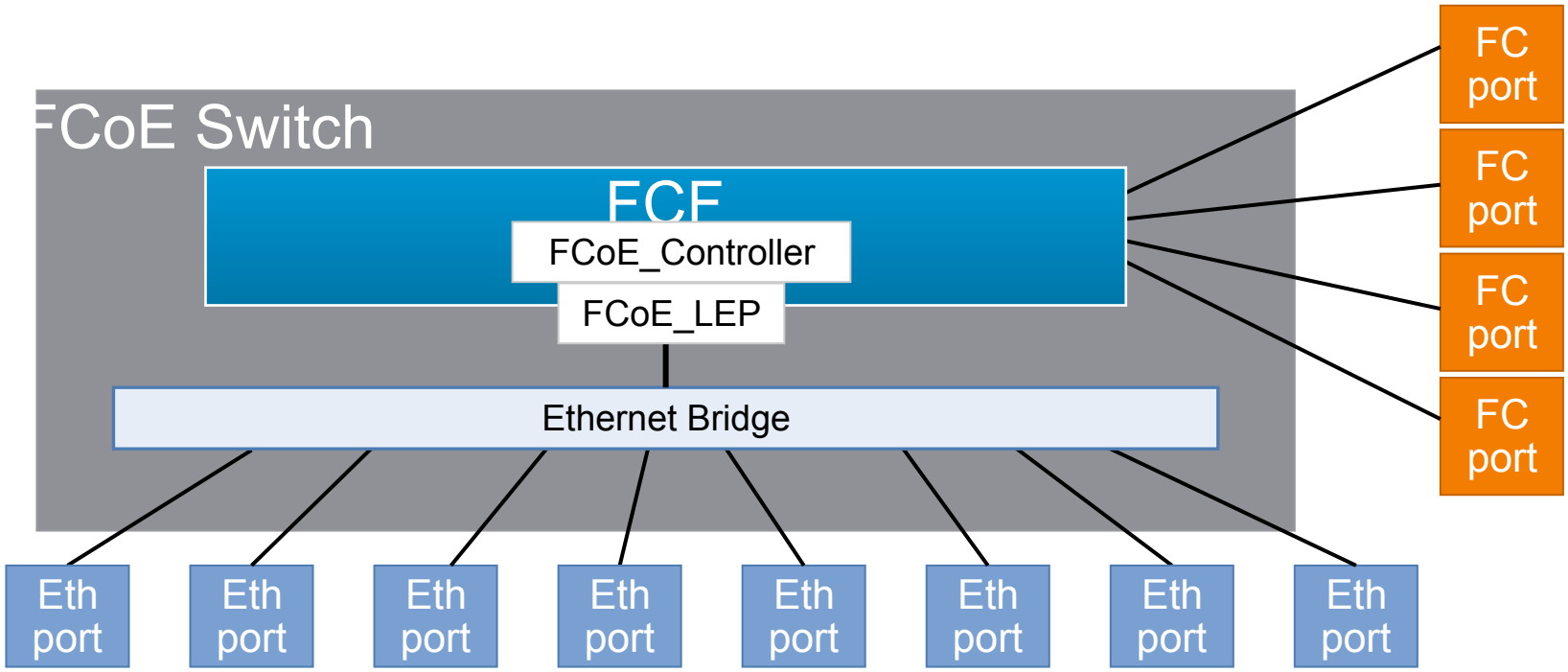


# FCoE Logical End Point



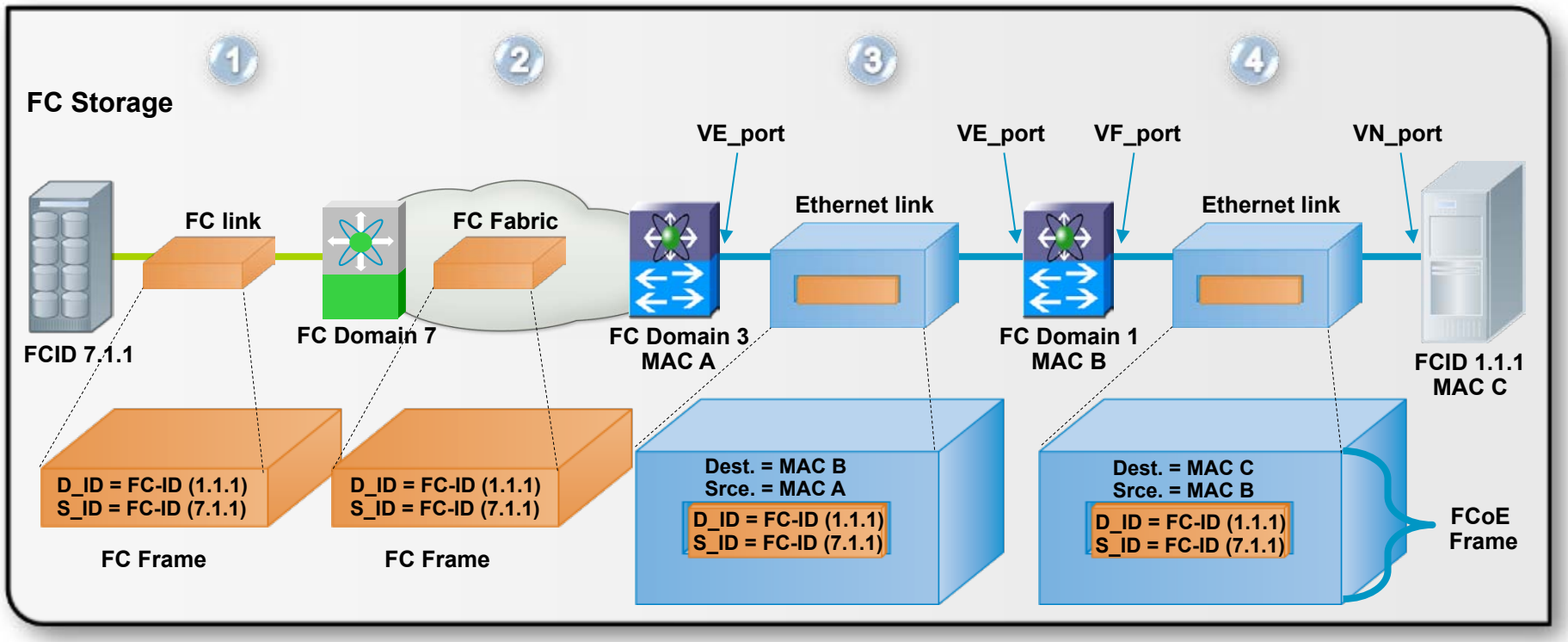
# FCoE Switch: Simplified Model

- FCF (Fibre Channel Forwarder), the forwarding entity inside an FCoE switch

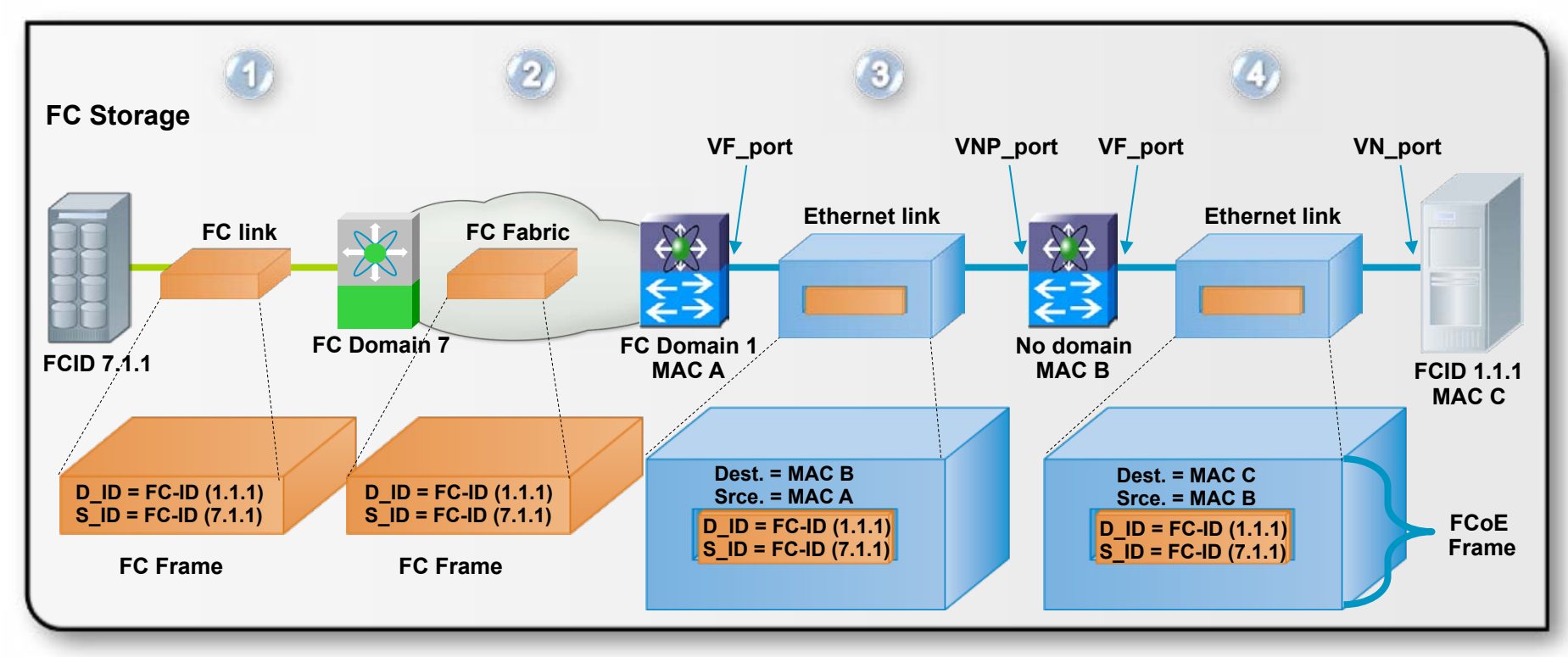


# FCoE Forwarding (VE_ports)

- FCoE frames have:
  - MAC addresses (hop-by-hop)
  - FC addresses (end-to-end)



# FCoE Forwarding (NPV)



# FCoE control plane

# What is FIP and Why FIP ?

- CIN-DCBX – Cisco, Intel, Nuova Data Center Bridging Exchange protocol, pre-standard
- CEE-DCBX – Converged Enhanced Ethernet Data Center Bridging Exchange protocol, which is standards base
- FIP does device discovery including discovering FCoE capable devices through ethernet-only pass-through devices

  Pre-FIP refers to the initial implementation of Unified Fabric where CNAs connected directly to the FCoE capable switch. Discovery was done through CIN
- Enables FCoE adapter to discover FCoE FCFs on a VLAN and establish a VN⇔VF link with one of the FCFs
- Supports explicit assignment of MAC address to the adapter (Server provided or Fabric provided)
- Building foundation for future multi-hop FCoE topologies



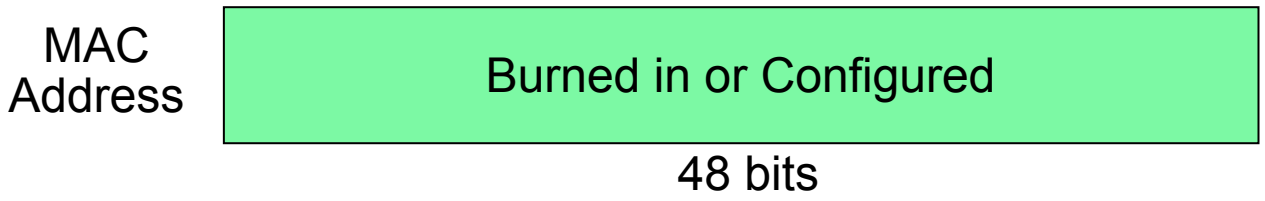
  

# Option 1: Server Provided MAC Addresses

Adapter uses burned-in or configured MAC address:

- Consistent with the Ethernet model

FCF needs a table to map between MAC addresses and FC_IDs



# Option 2: Fabric Provided MAC Addresses

MAC address assigned for each FC_ID:

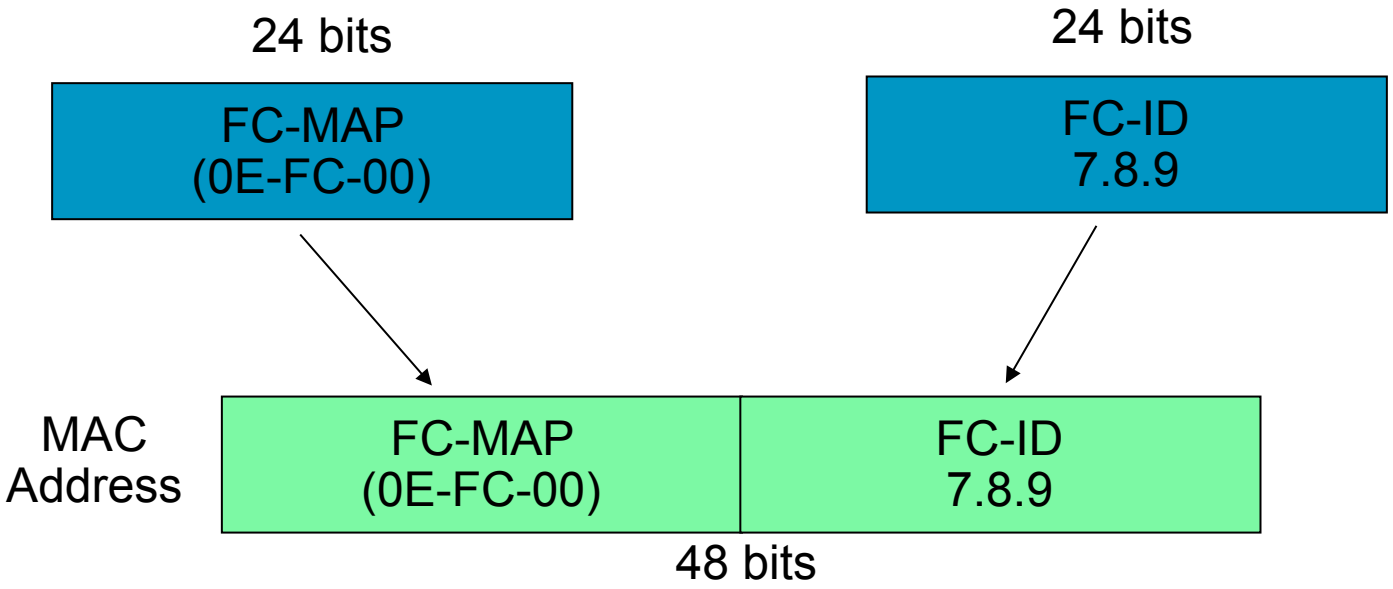
- Consistent with the Fibre Channel model

Uses OUIs with U/L = 1 (Local addressing), called FC-MAPs

Multiple FC-MAPs may be supported (one per SAN)

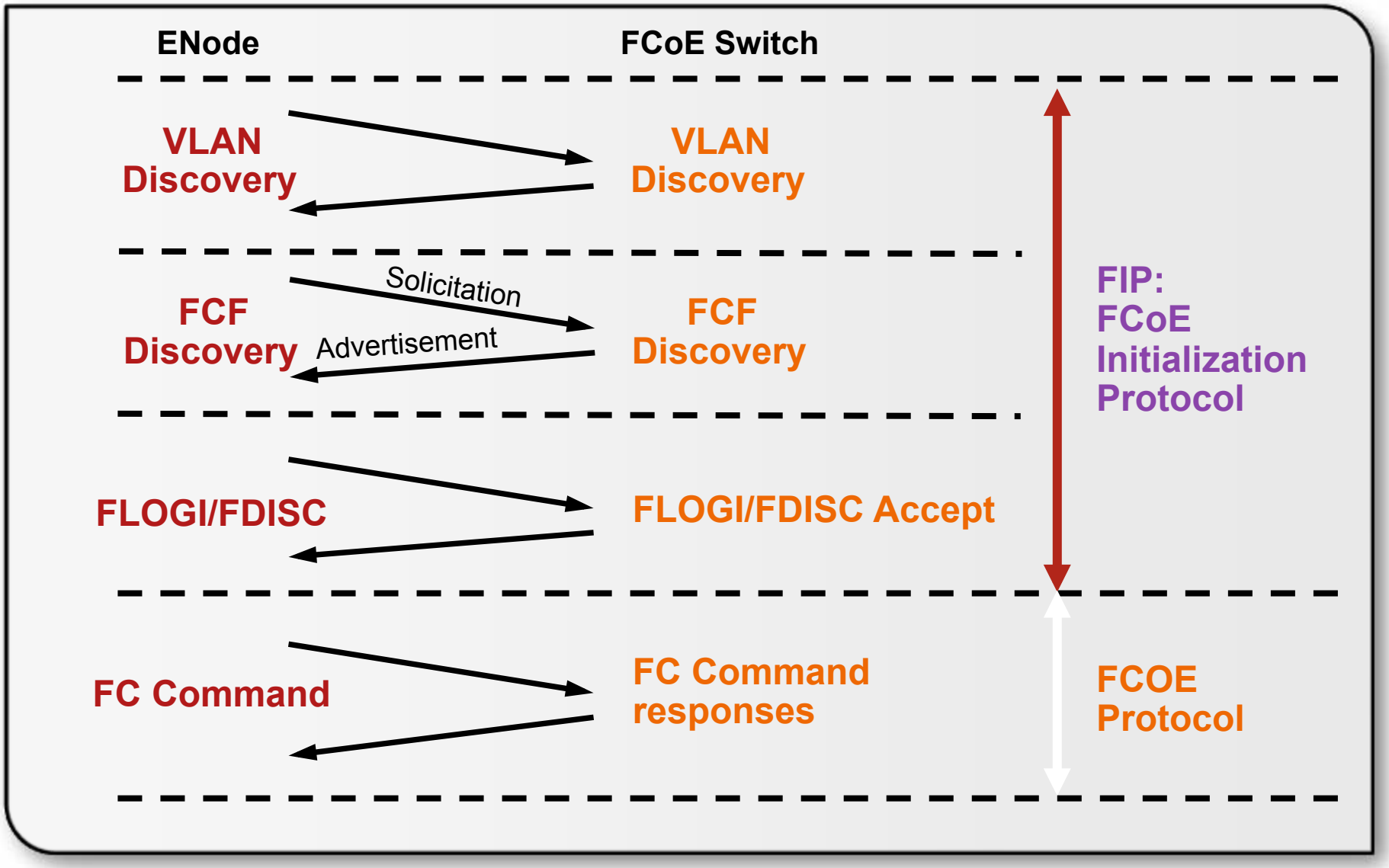
No table needed for Encapsulation

Multiple MACs may be needed for NPIV




Cisco Nexus 5000 uses FPMA

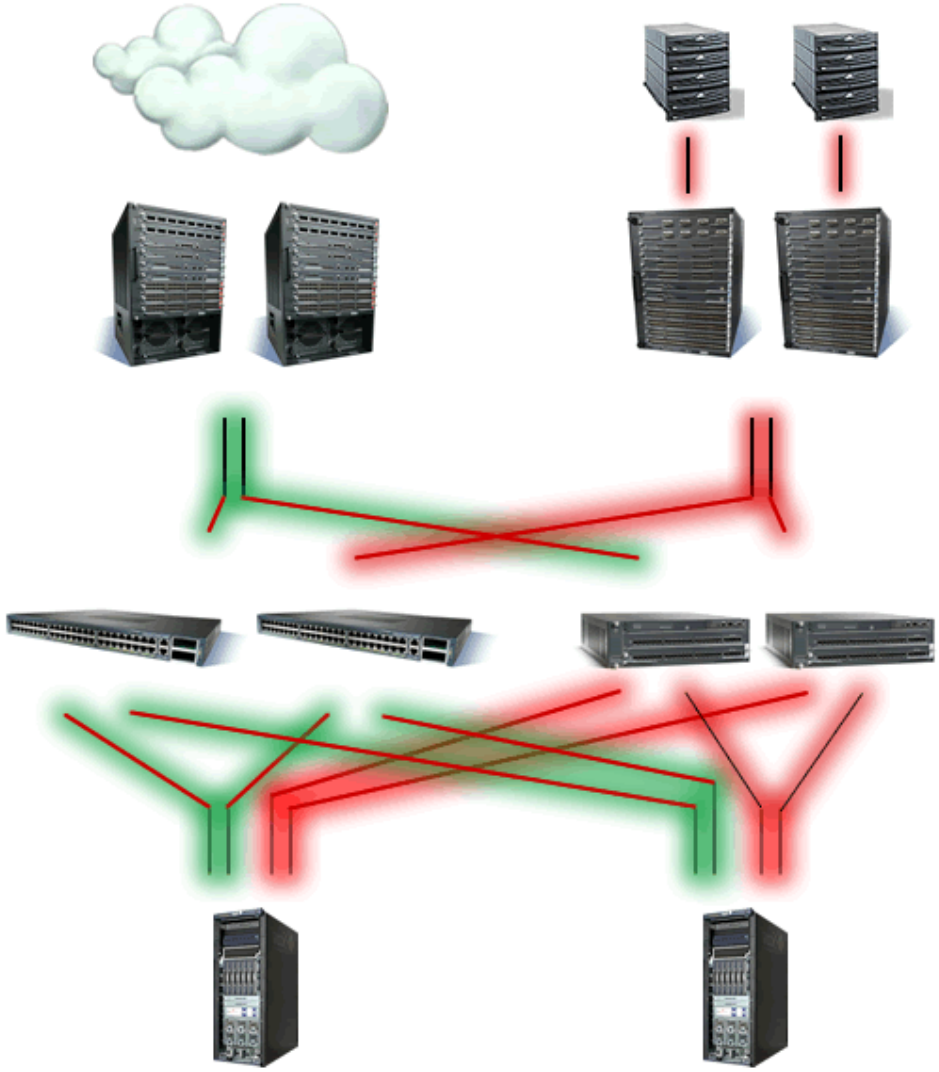
# Initial Login Flow ladder



# FCoE Network Topology

# Unified Fabric

## Where we started - Segregated Fabrics



### Seggregated LAN and SAN

**In current architectures, LAN and SAN connectivity is seggregated directly from the Servers, where NICs and HBAs connect into Ethernet switches and Fibre Channel Fabrics. This may result in excess of 8+ cables to/from each physical server**

**In Ethernet, redundancy relies upon technologies such as Spanning Tree Protocol to provide a loop-free topology...**

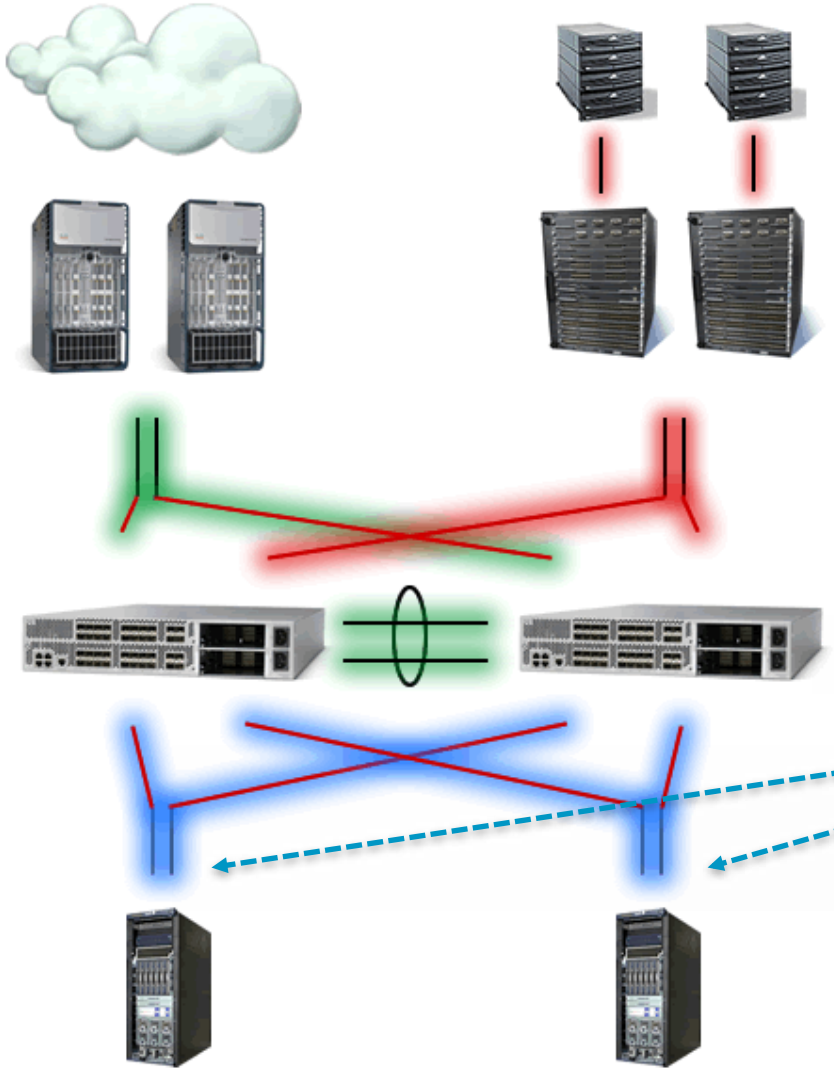
**Fibre Channel**

**Ethernet**

**FCoE**

# Unified Fabric with FCoE

## Phase 1 – Where we have been

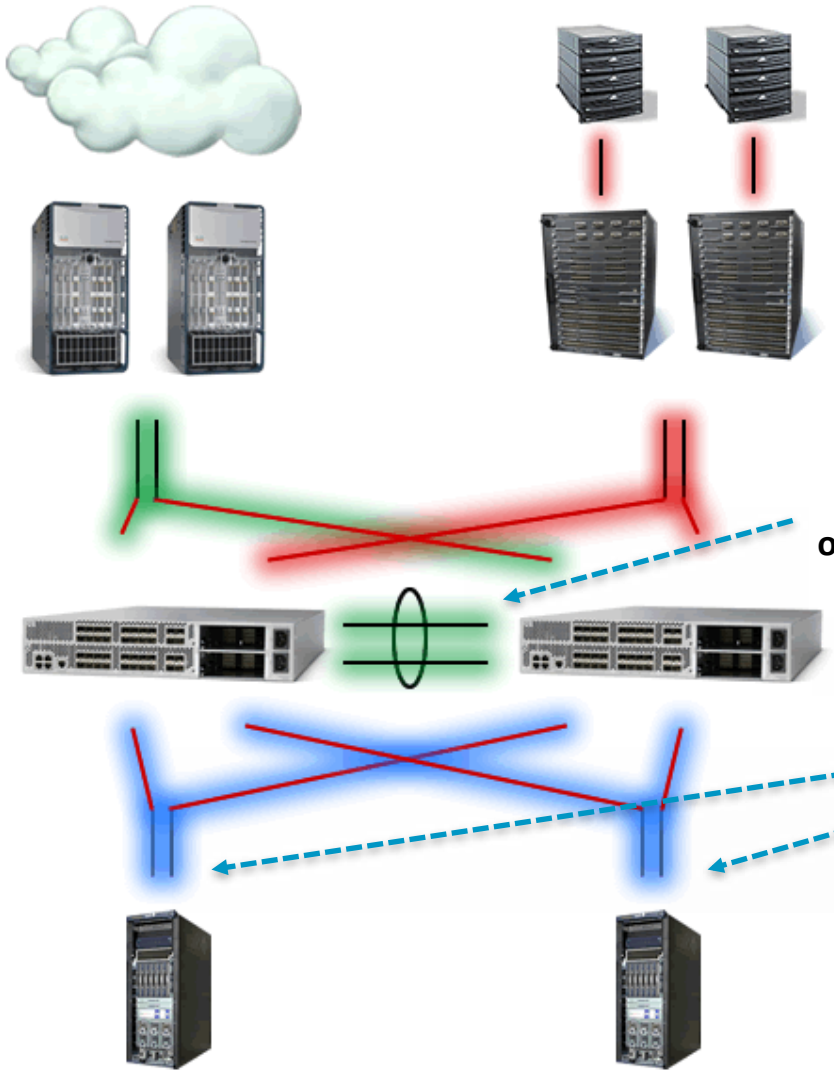



**Servers with Converged Network Adapters (CNAs) provide consolidated LAN/SAN Direct-attach connectivity to the Nexus 5000.**

**The Nexus 5000 can operate in FCF mode or NPV mode. It can then split off Ethernet traffic and Native Fibre Channel traffic**

# Unified Fabric with FCoE

## Phase 2 – Where we are today (FIP support)



Servers with Converged Network Adapters (CNAs) provide consolidated LAN/SAN connectivity to the Nexus 5000.

FIP on the Nexus 5000s now provides discovery of direct and indirect attached FCoE initiators

Supported as of NX-OS 4.1(3)N1 on Nexus 5000

VPC Peer link carrying Ethernet Traffic only maintaining SAN A and SAN B isolation

Nexus 5000s as FCF or In NPV mode

FIP on Nexus 5000 allows for direct attach And indirect attach CNAs

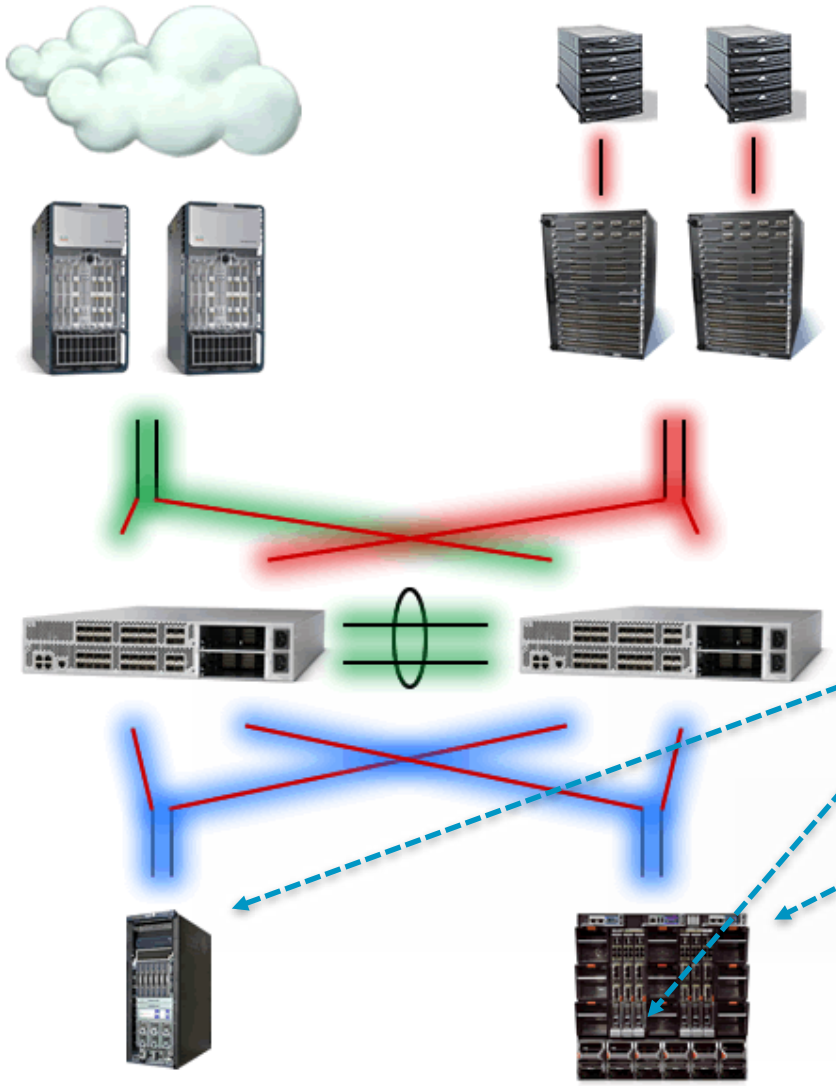
Fibre Channel

Ethernet

FCoE

# Unified Fabric with FCoE

## Phase 2 – Adding the Nexus 4000 (Q4CY10)



**Mezzanine CNAs can be installed within Blade Enclosures, connecting up to Nexus 4000 Blade Switches.**

**Nexus 4000 will provide FIP Snooping Capabilities, providing lossless servicing and Segmentation.**

**FIP on the Nexus 5000s now provides discovery of indirect attached FCoE initiators**

**FIP on Nexus 5000 allows for direct attach And indirect attach CNAs**

**Nexus 4000 Blade Switches providing FIP Snooping capabilities**

**Fibre Channel**

**Ethernet**

**FCoE**

# Unified Fabric with FCoE

## Phase 2 – Adding the Nexus 4000 (Q4CY10)

4 - 6 sfp+ 10/1G ports*

14 -16 server 10/1G ports

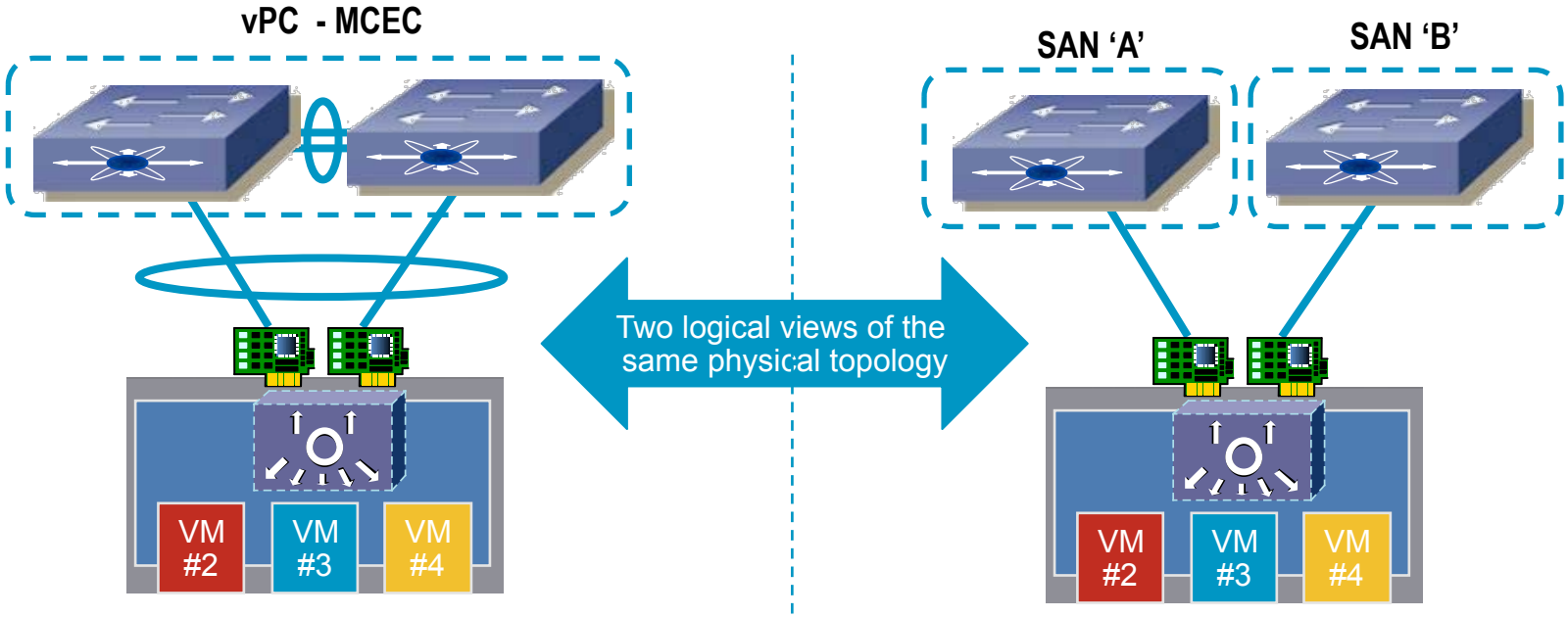
*Only supports for 1G fiber SFPs*

Data Center Bridging Enabled

- 4K IEEE 802.1Q VLAN
- 8K MAC entries
- Etherchannel
- Cut Through Switching
- Latency: ~1.2 μs
- FCOE
- FIP Snooping Bridge
- Eight Independent Virtual Lanes
- PFC and ETS
- IEEE802.3X Pause

# Unified Fabric with FCoE

## FCoE Design Considerations




- In a Unified I/O configuration (FCoE) we have two distinct topologies
  - Isolated access switches - SAN 'A' and SAN 'B'
  - Combined access switches – vPC supporting MCEC
- To ensure correct forwarding behaviour 'vfc' interface can only be associated with a vPC etherchannel (only one physical interface per switch)

# Unified Fabric with FCoE

## FCoE Design Considerations

### CNAs NX-OS FIP Support Matrix

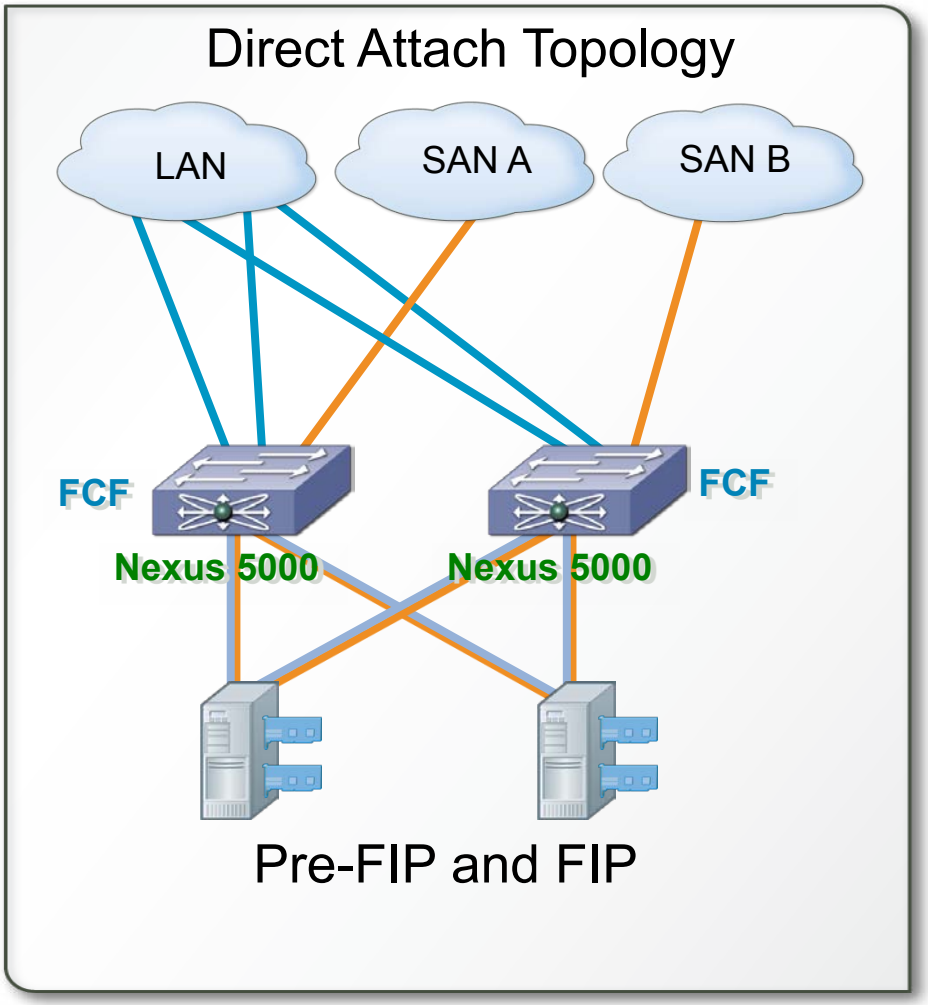
| | 4.0(x)x | 4.1(3)N1 |
|---|---|---|
| Emulex LP2100x (Gen-1) | CIN-DCBX + Pre-FIP | CIN-DCBX + Pre-FIP |
| Emulex One Connect (Gen-2) | NA | CEE-DCBX + FIP only |
| NetAPP (Gen-1) | CIN-DCBX + Pre-FIP | CIN-DCBX + Pre-FIP |
| NetAPP (Gen-2) | NA | CEE-DCBX + FIP only |
| Qlogic QLE804x (Gen-1) | CIN-DCBX + Pre-FIP | CIN-DCBX + Pre-FIP |
| Qlogic QLE814x/QLE815x (Gen-2) | NA | CEE-DCBX + FIP only |

Note: Nexus 5000 Cronulla code is **backwards compatible** with all Gen-1 CNAs and controllers.

# Unified Fabric with FCoE

## FCoE Design Considerations

- Direct Attached FCoE
- SAN 'A' and SAN 'B' with Standard NIC teaming for Ethernet
- Pre-FIP (CIN-DCBX) is supported
  - CIN-DCBX = Cisco, Intel, Nuova – Data Center Bridging Exchange
- FIP (CEE-DCBX) is supported
  - CEE-DCBX = Converged Enhanced Ethernet – Data Center Bridging Exchange

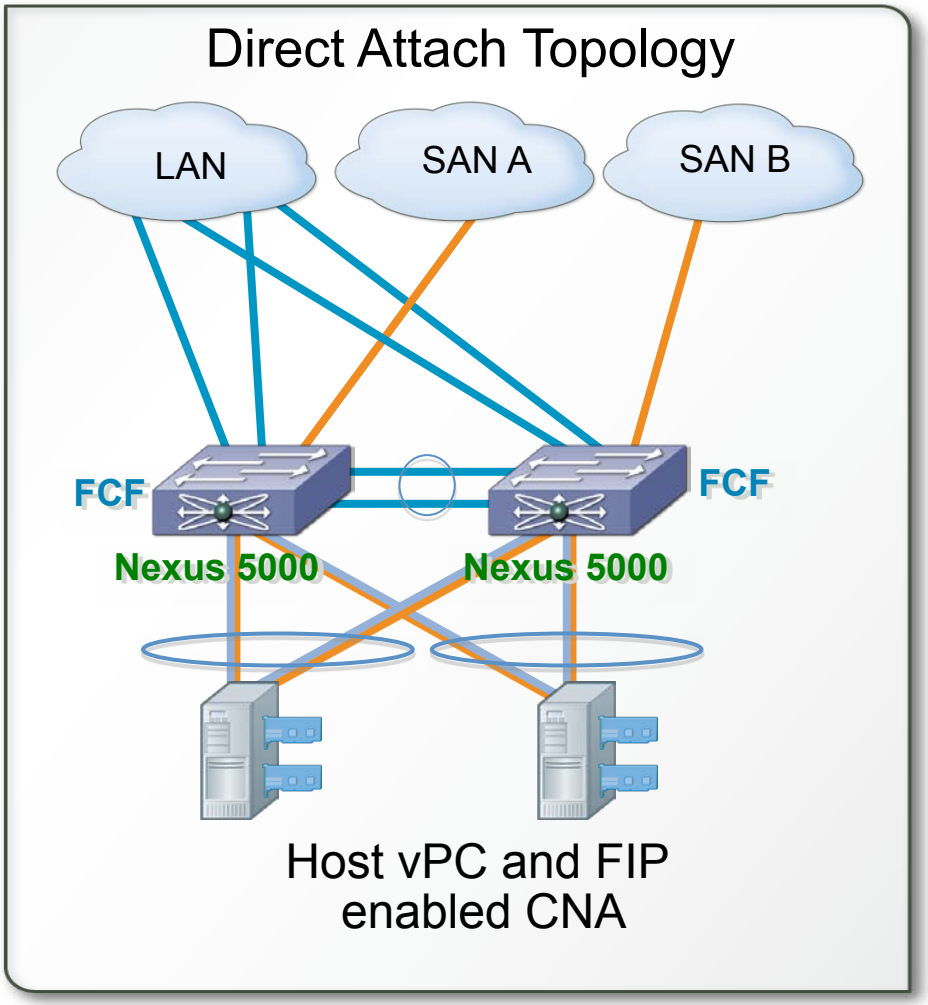



4G FC
10G Ethernet
10G Unified I/O

# Unified Fabric with FCoE

## FCoE Design Considerations

- Direct Attach using MCEC - vPC
- SAN 'A' and SAN 'B' with LACP based NIC teaming for Ethernet
- FIP (CEE-DCBX) is required even with direct attach when using vPC
- Will work with Gen-2 CNAs *only*

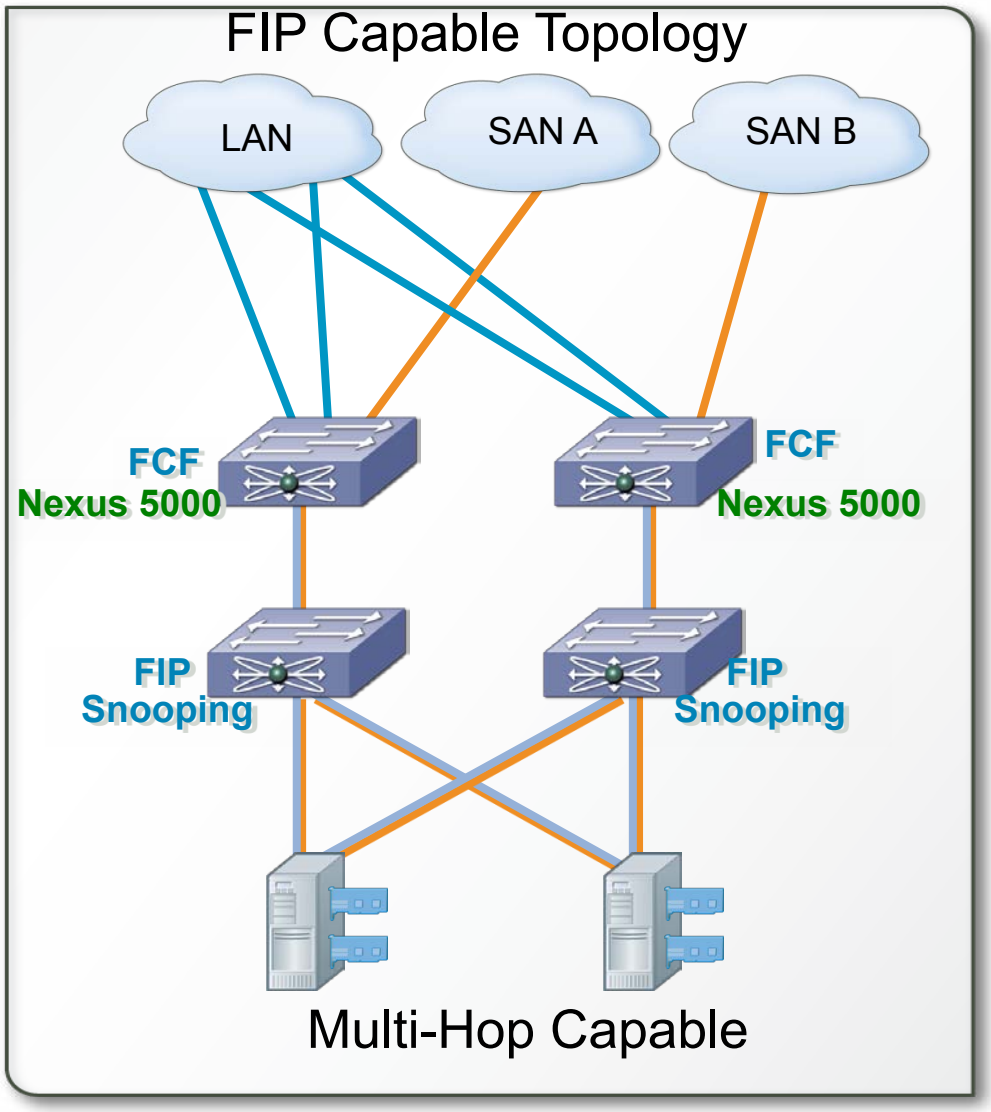



4G FC
10G Ethernet
10G Unified I/O

# Data Center Access Architecture

## FCoE Design Considerations

- Multi-hop attached initiators
- SAN ‘A’ and SAN ‘B’ with Standard NIC teaming for Ethernet
- Work with Gen-2 FIP Capable CNAs ONLY
- FIP (CEE-DCBX) is required
- First hop switch MUST HAVE “FIP Snooping” capabilities for multi-hop support for FCoE

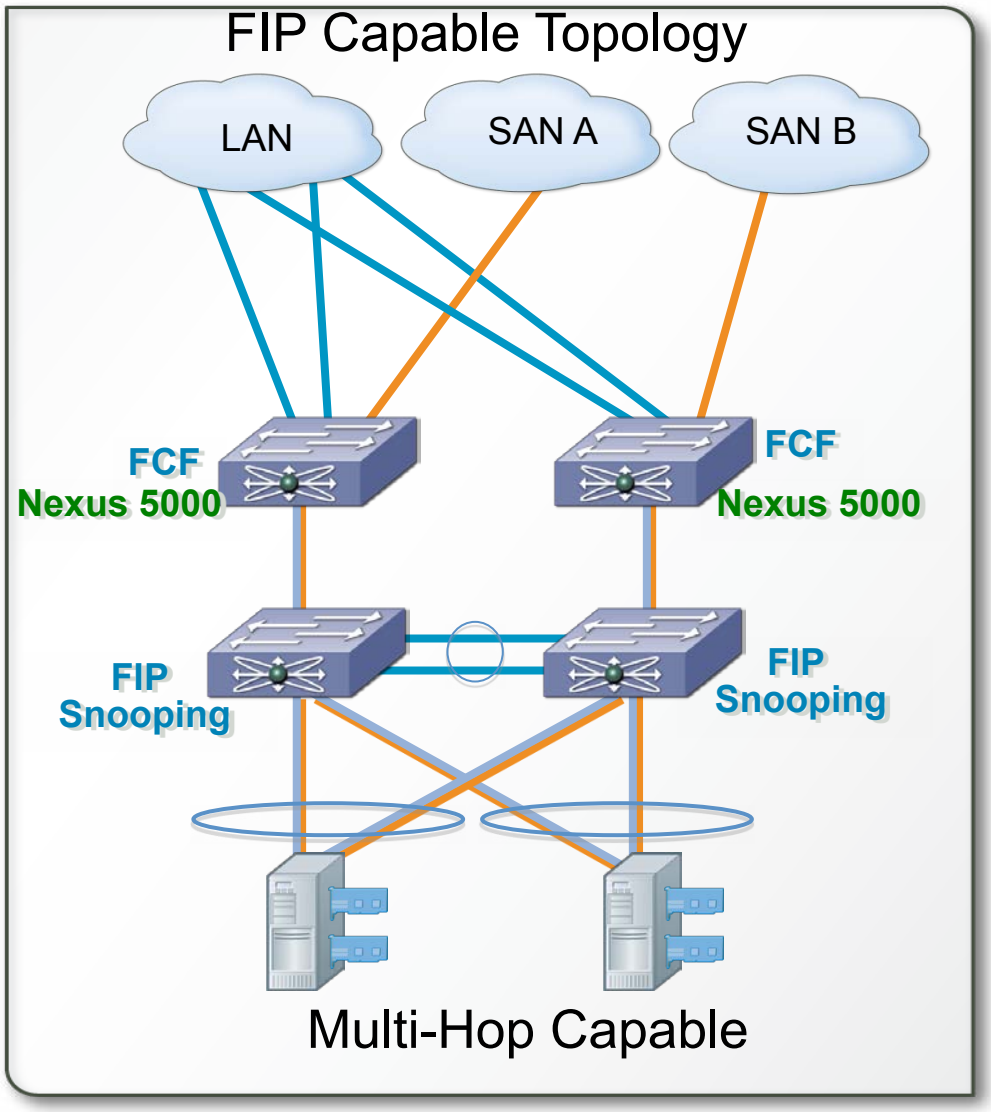



4G FC
10G Ethernet
10G Unified I/O

# Data Center Access Architecture

## FCoE Design Considerations

- Multi-hop initiators with vPC connectivity
- Single host CNA port connect to separate Nexus 5000
- FIP (CEE-DCBX) is required
- Work with Gen-2 CNAs *only*
- First hop switch MUST HAVE "FIP Snooping" capabilities for multi-hop support for FCoE

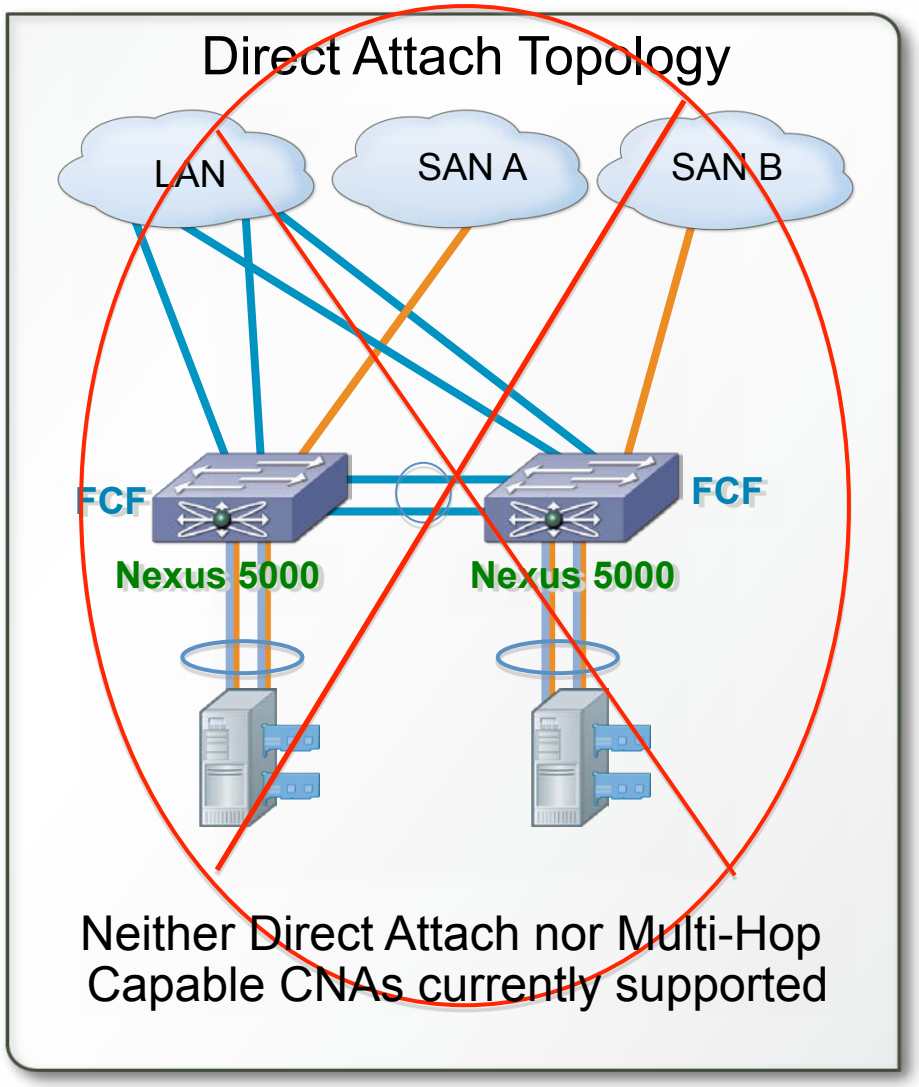



4G FC
10G Ethernet
10G Unified I/O

# Data Center Access Architecture

## FCoE Design Considerations

- Single initiator dual homed via a Port-Channel to a single Nexus 5000
- No ability to isolate SAN ‘A’ and SAN ‘B’
- This is an *unsupported* configuration
- Not consistent with Fibre Channel High Availability design requirements

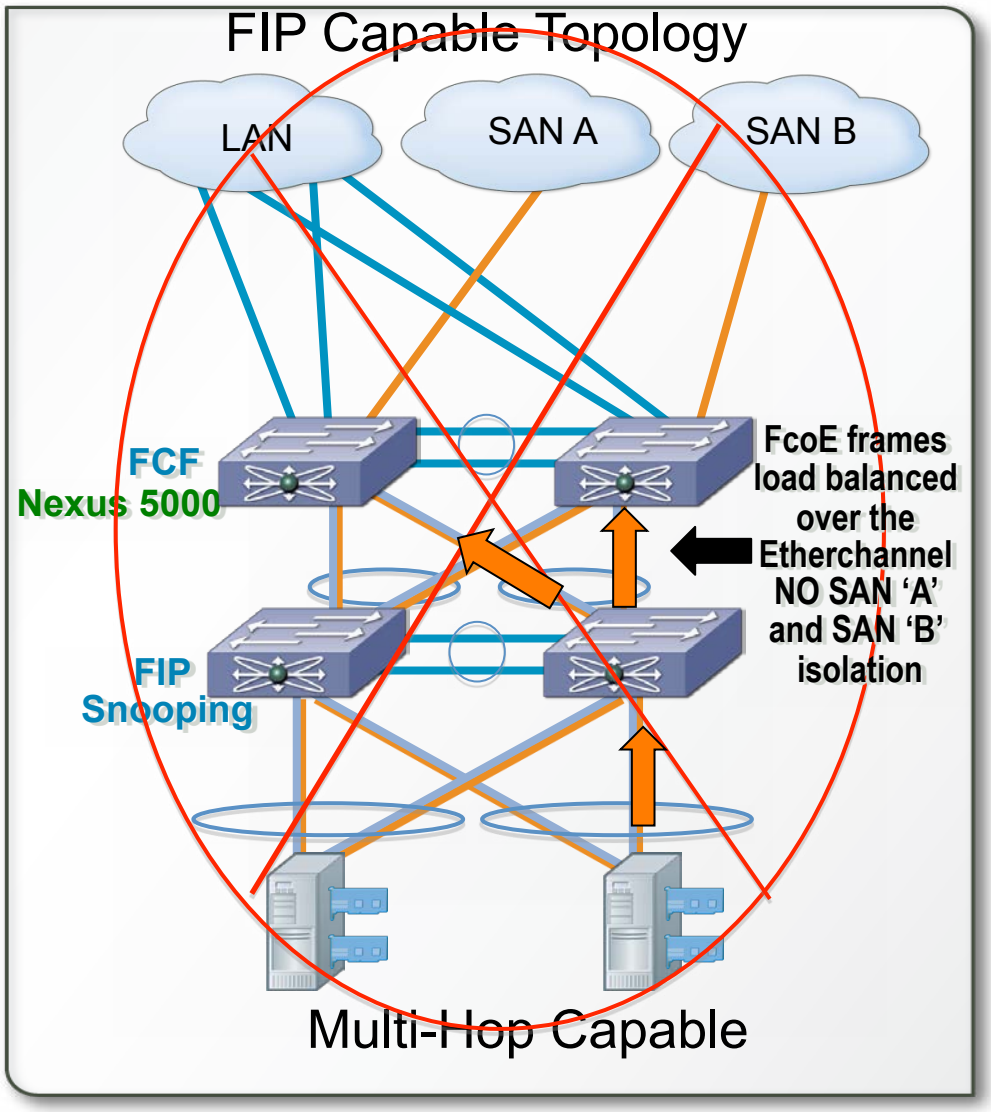



4G FC
10G Ethernet
10G Unified I/O

# Data Center Access Architecture

## FCoE Design Considerations

- Initial vPC is supported as it is possible to isolate the SAN ‘A’ and SAN ‘B’ traffic between the CNAs and first hop switches
- No ability to isolate SAN ‘A’ and SAN ‘B’ between the first and second tier of switches
- This is an *unsupported* configuration
- Not consistent with Fibre Channel High Availability design requirements



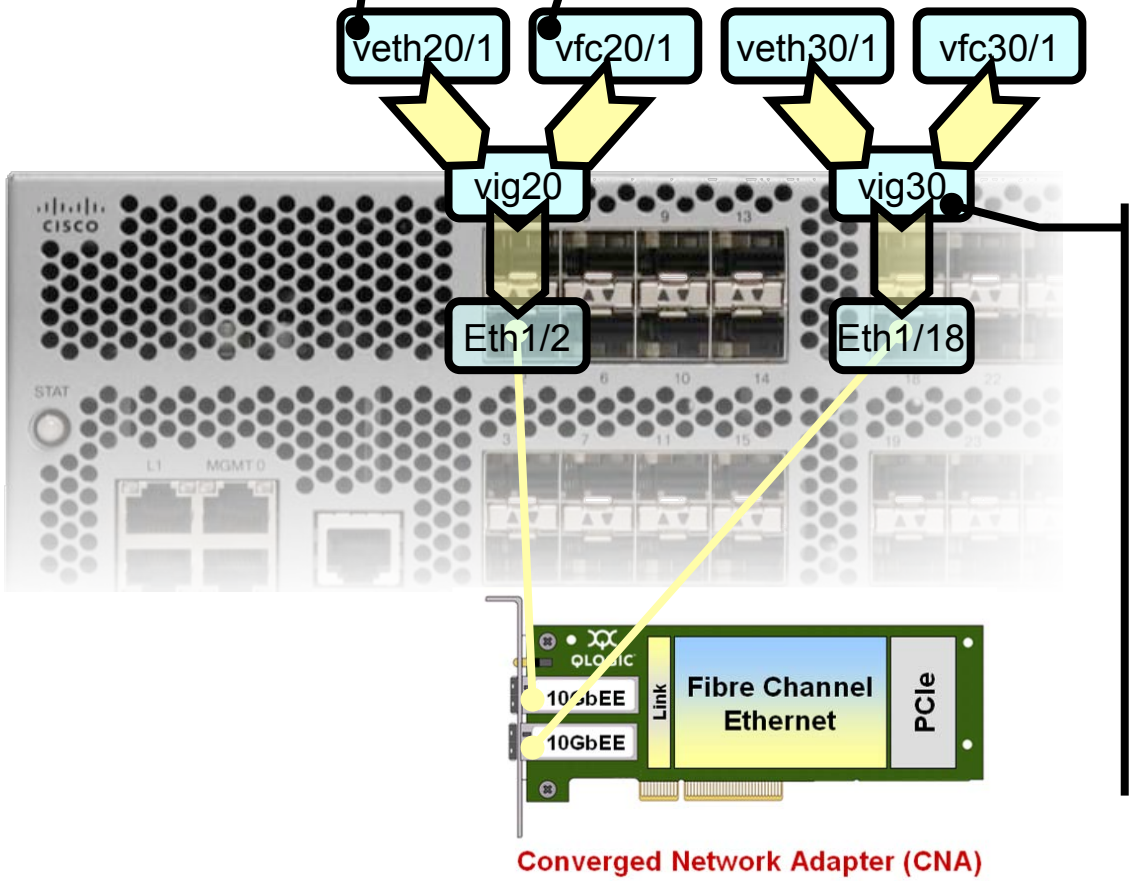

4G FC
10G Ethernet
10G Unified I/O

# FCoE configuration

# Virtual Interface terminology on Nexus 5000

**Virtual ethernet interface**: maps an ethernet vNIC on a server

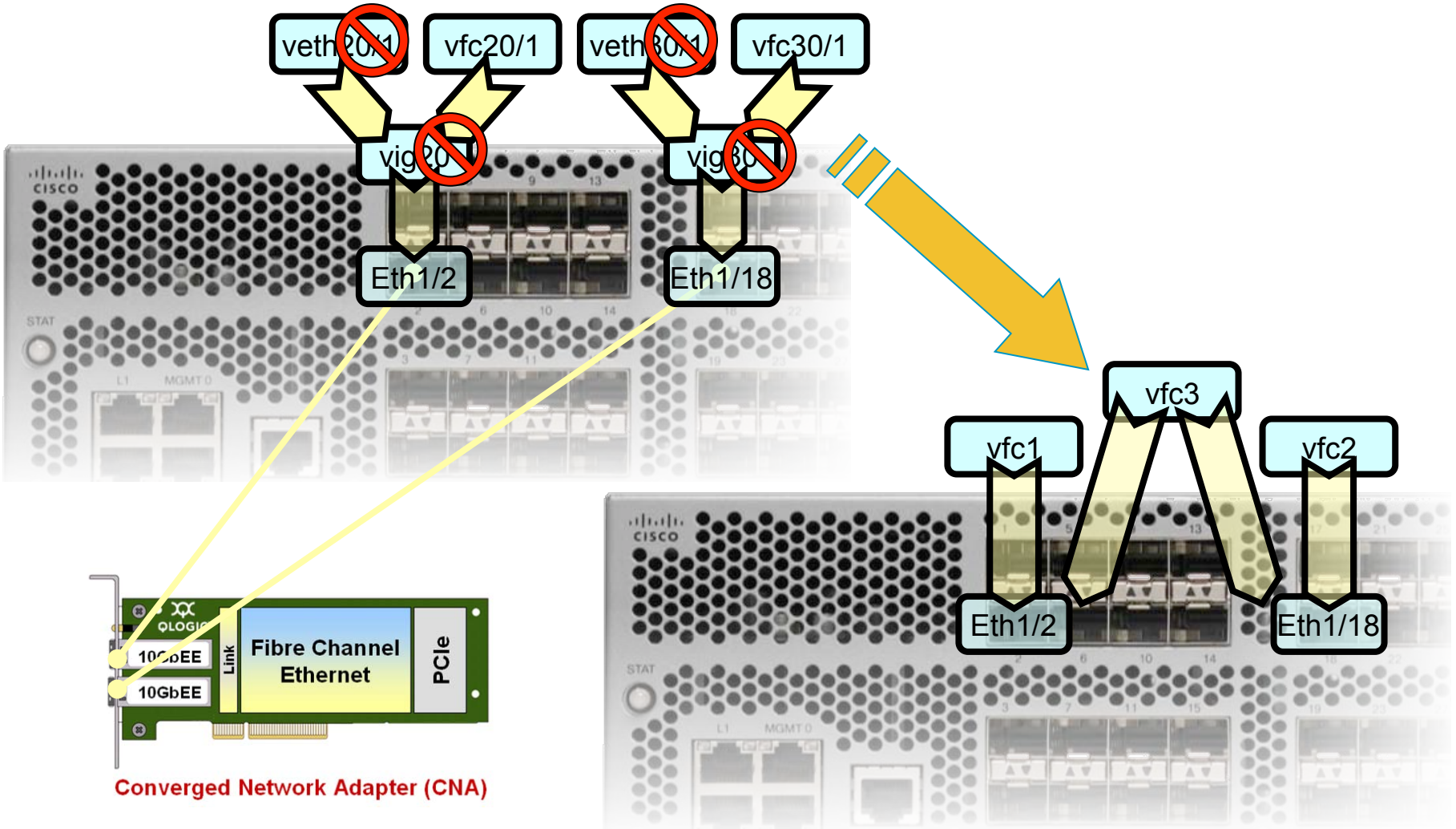
**Virtual FC interface**: maps an FC vHBA on a server



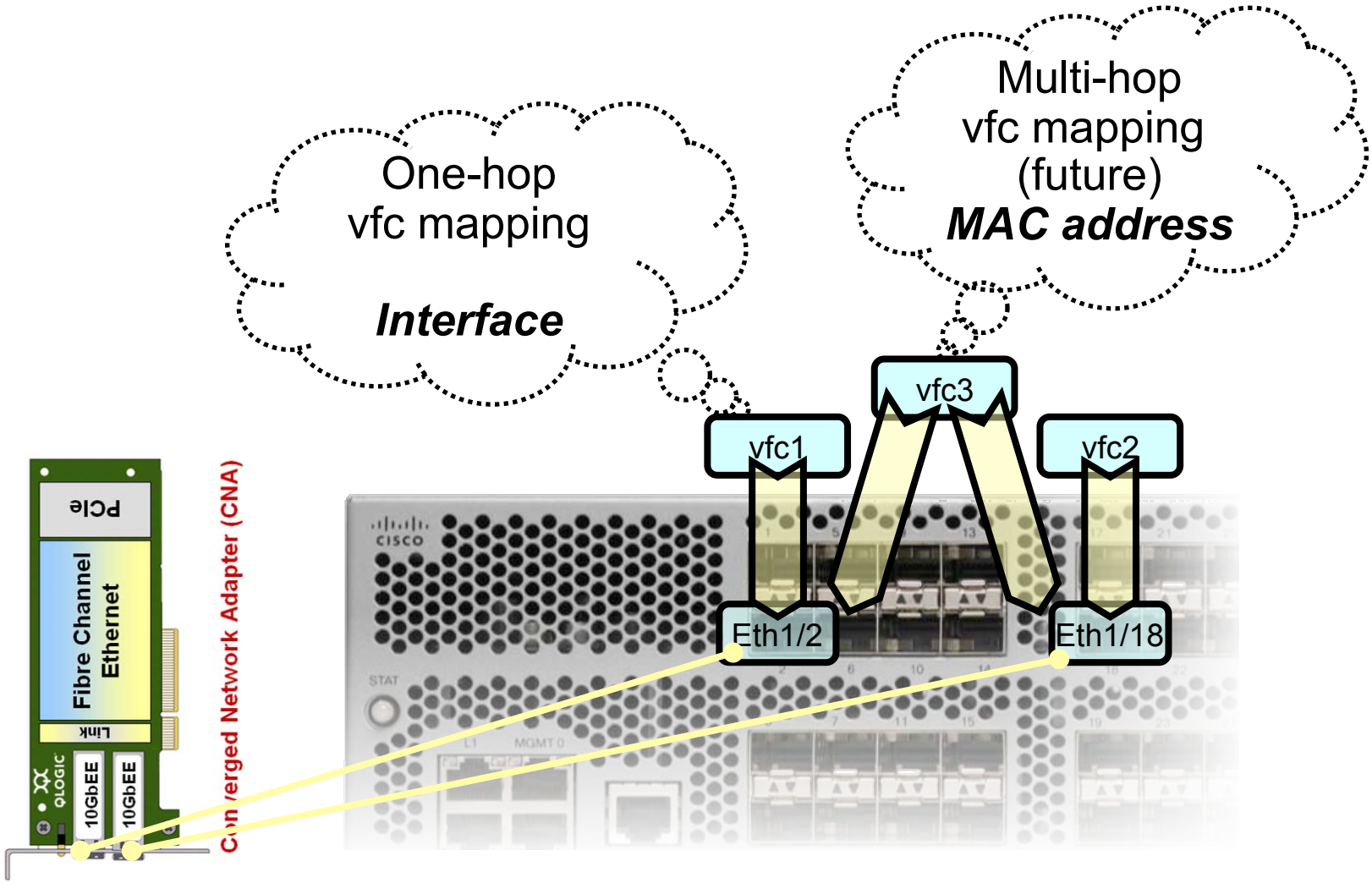

**Virtual Interface Group**

- Maps a logical adapter port connecting to the switch
- Groups veth and vfc interfaces
  - 1 veth and 1 vfc per vig
- Allows offline configuration for link level properties
- Configured to be bound to a physical interface (etherchannel in the roadmap)

# Getting ready for FIP

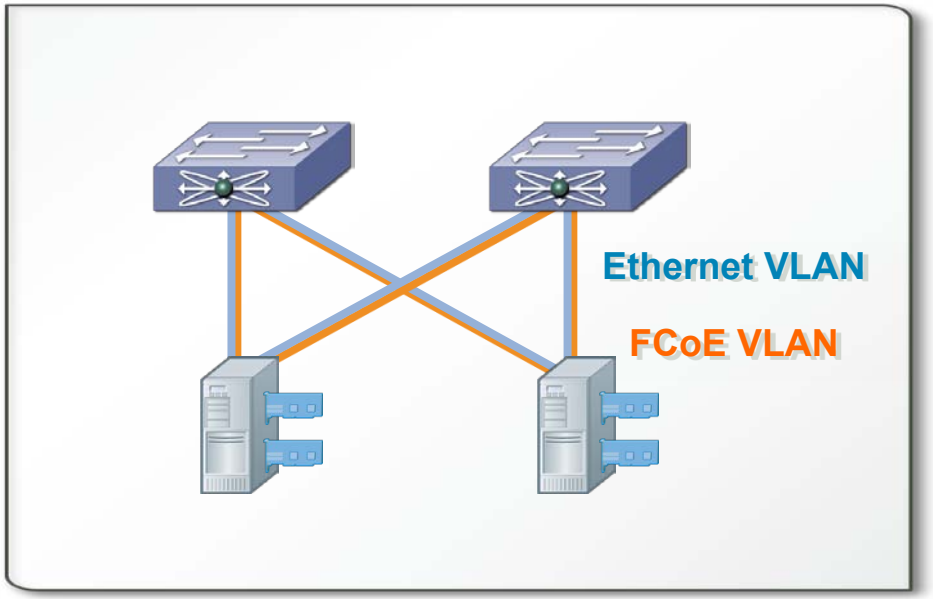


# New vfc mappings



# VSAN to VLAN mapping


```
N5K# configure terminal
N5K(config)# vlan 10
N5K(config-vlan)# fcoe vsan 1
N5K(config-vlan)# show vlan fcoe
VLAN       VSAN       Status
--------   --------   --------
10         1          Operational
```



# Management

Who owns what?

# How do I manage this??

- To a SAN Admin: just another FC edge switch

  All Fibre Channel services are available and configured the same

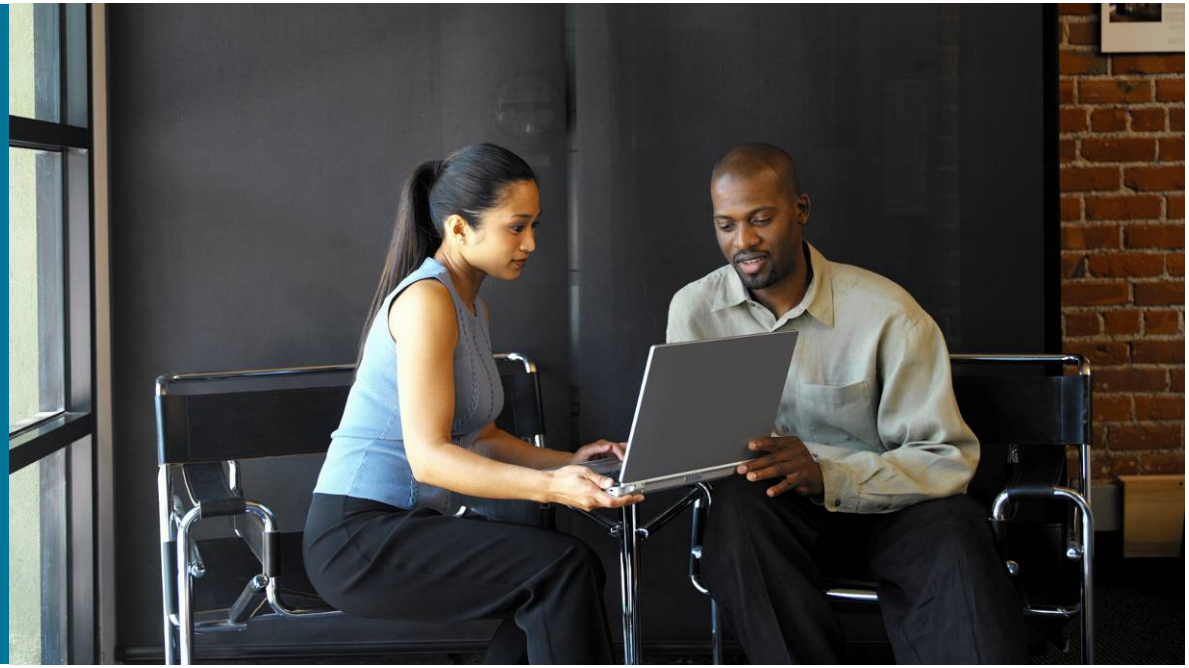
  - Managed by Fabric Manager/Device Manager
  - FCoE hosts as normal FC host
  - Normal zoning applies
  - Nexus can operate in Switch or NPV mode

- To a LAN Admin: just another L2 Ethernet access switch

  All Ethernet is configured the same

Roles Based Access Control
is used for guaranteed separation

# In Summary…



# Key takeaways

- *FCoE is an extension of Fibre Channel onto a Lossless Ethernet fabric*

  *From an Ethernet perspective – just another Ethernet packet*

  *From a Fibre Channel perspective – just another FC packet encapsulated in an Ethernet frame*

- SAN best practices remain valid in Unified I/O environments

- Unified I/O will be realized first at the edge layer, followed by the core layer

  *CAPEX savings are evident at the edge, OPEX savings are evident throughout…*

# Q & A…


CISCO