Unified Fabric Infrastructure

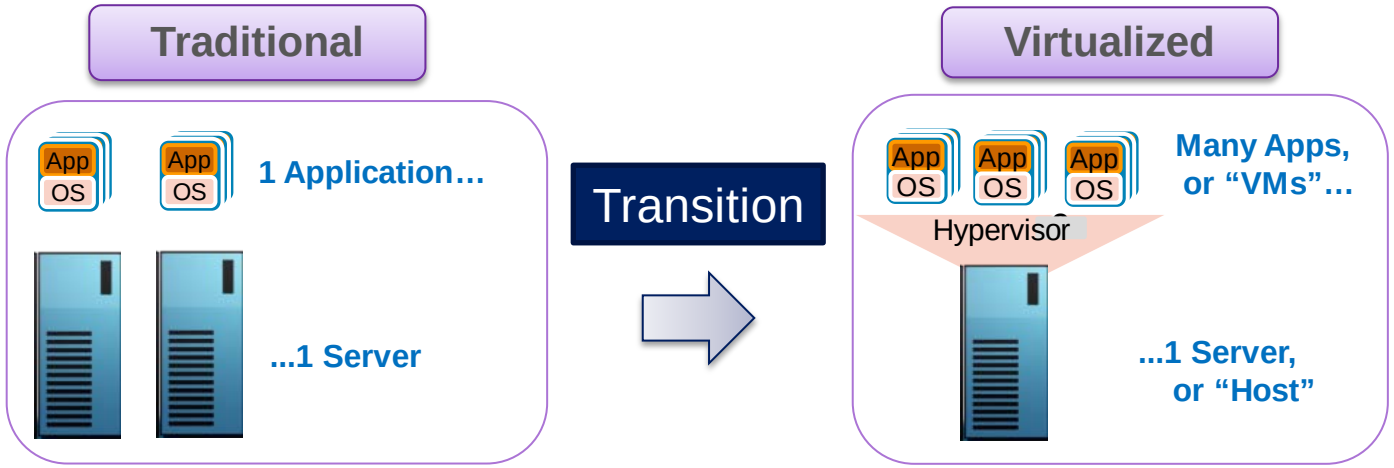
Gilles Chekroun
Distinguished Engineer – Data Center Virtualization Team - EMEAR

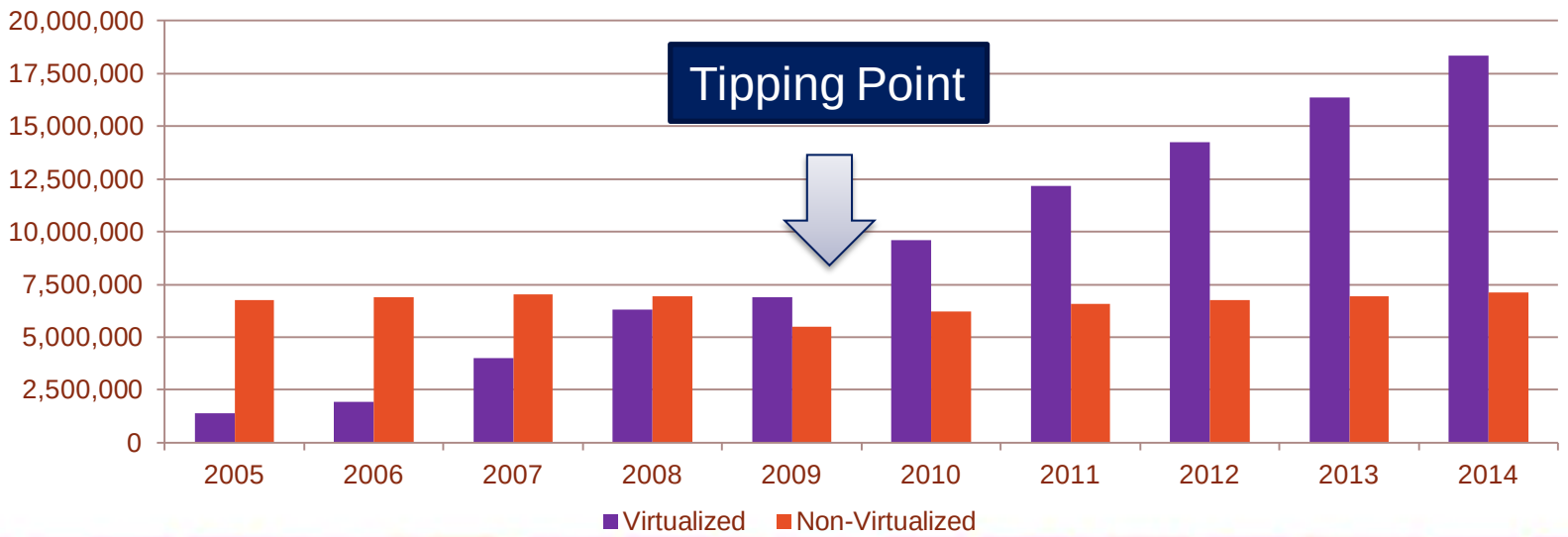
10-OCT-2012

Storage Networking Trends

The Evolving Data Center Architecture

- Virtualization Requires Network Based Storage





Source: IDC, Nov 2010

Why Network Consolidation?

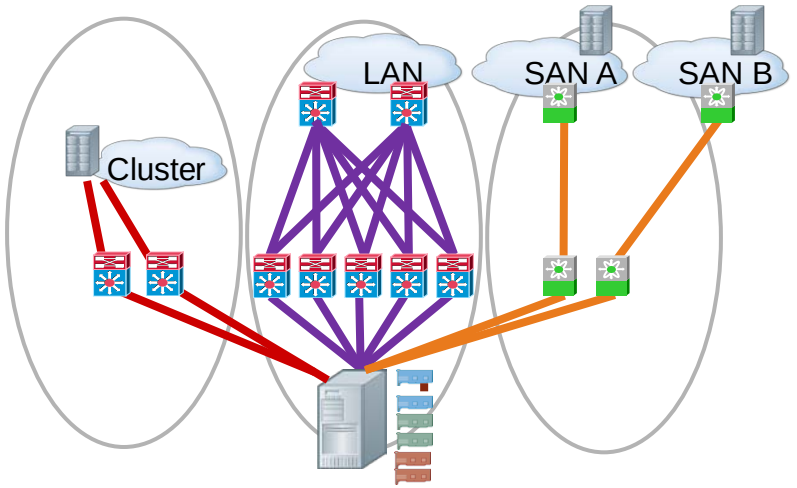
- Fewer physical objects in the Data Center, better asset utilization with the same workloads

  Conservation of space, power, cooling

- Fewer technologies, specialized expertise, greater economies of scale

10GbE Enables Consolidation of Networks, Reduction in Specialized Networks

Specialized High Performance Workload Network



Specialized Fibre Channel Storage Network

Why FCoE?

It's Ethernet!!

Ethernet Model Has Proven Benefits

Ethernet Economic Model

- Embedded on Motherboard
- Integrated into O/S
- Many Suppliers
- Mainstream Technology
- Widely Understood
- Interoperability by Design

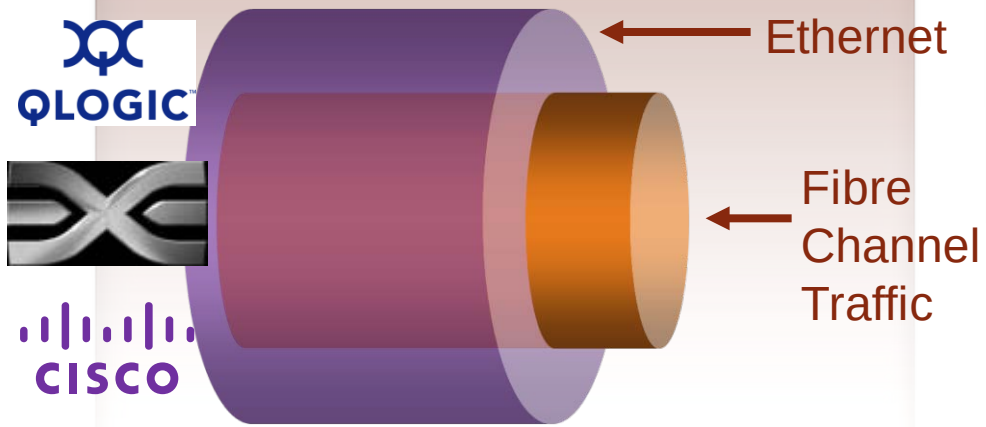
FC Economic Model

- Always a Stand-Up Card
- Specialized Drivers
- Few Suppliers
- Specialized Technology
- Special Expertise
- Interoperability by Test

FC over Ethernet (FCoE)

FCoE

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

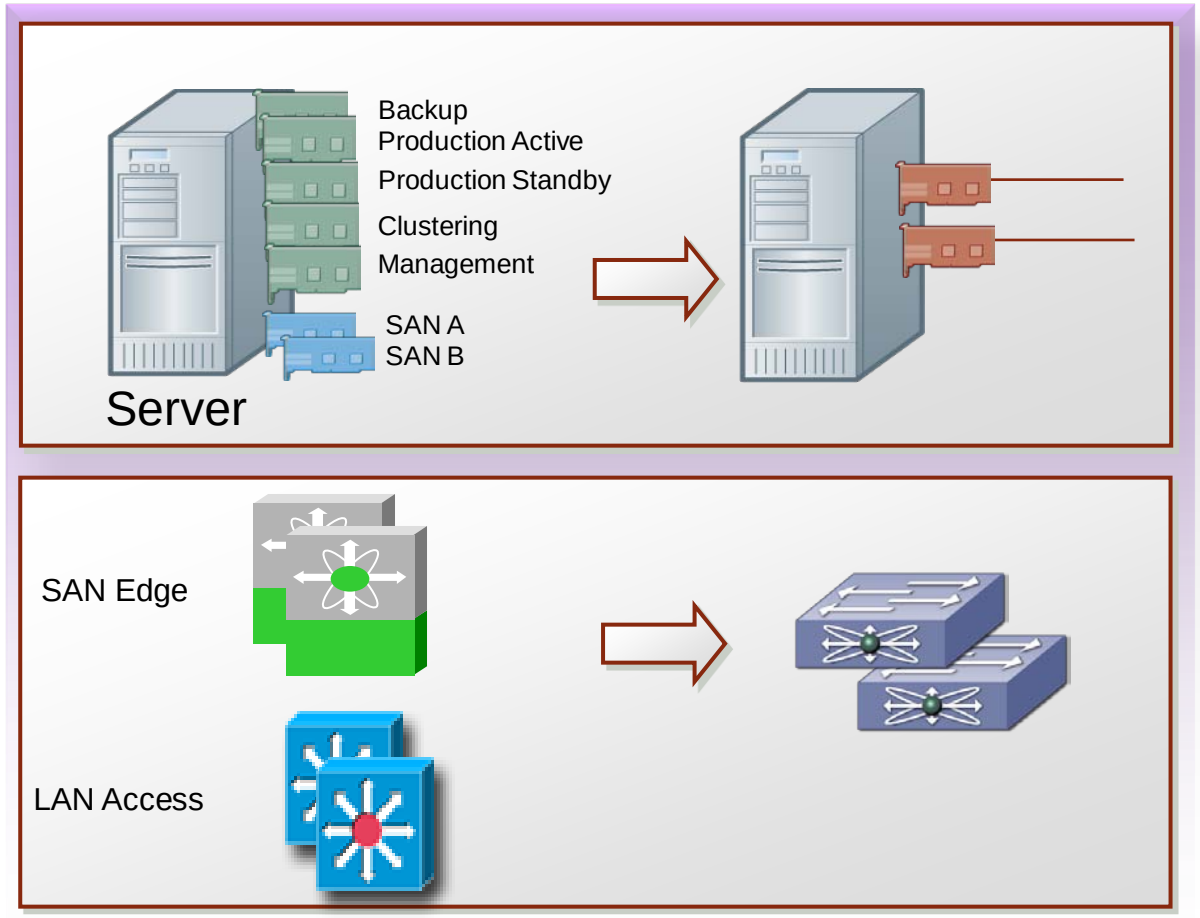


Benefits

- 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 of SAN remain consistent
- No Gateway

FCoE Enables Unified Fabric

- Enables Shared LAN and SAN Links for Access Layer TCO Benefits



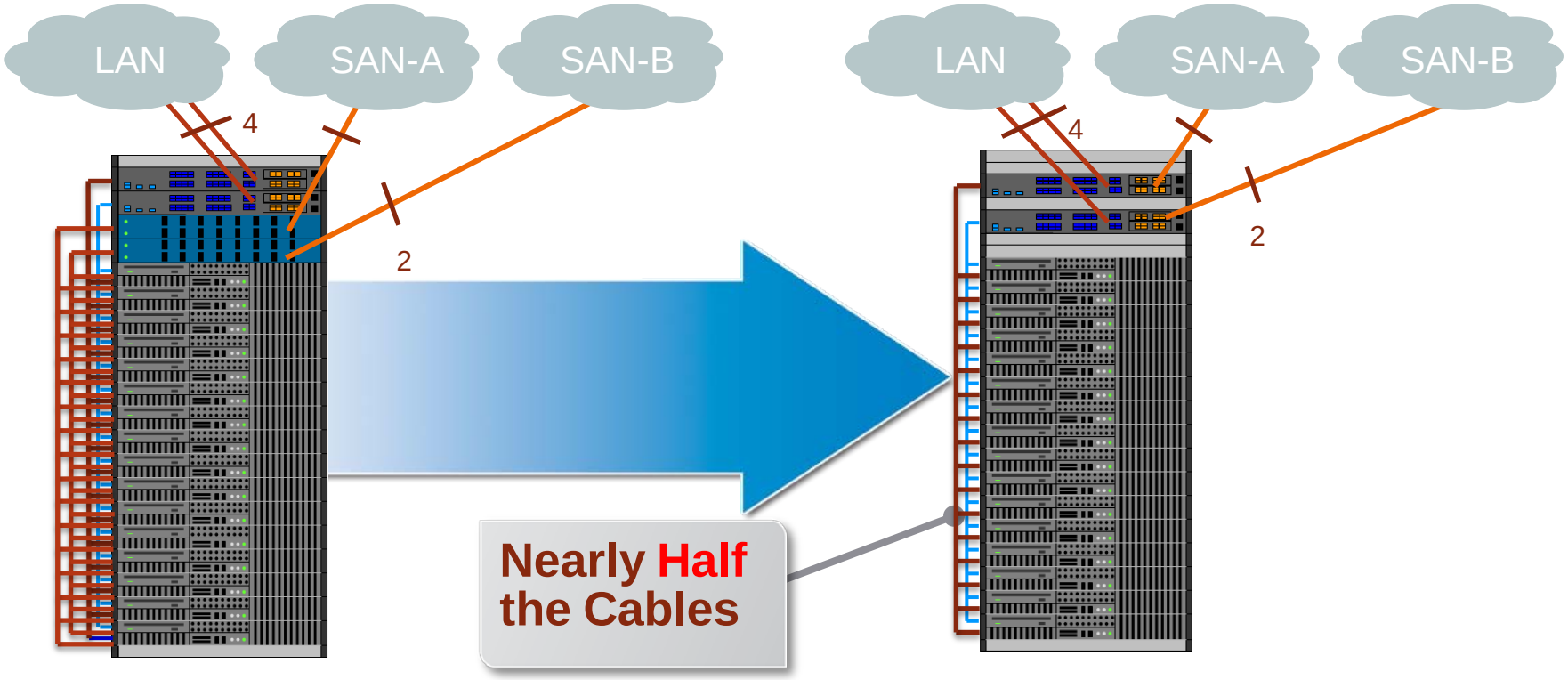
Consolidate I/O

- Eliminate adapter and switch ports
- Eliminate cables
- Avoid air dams
- Reduce cable maintenance
- Reduce power

Standardize I/O

- Faster server deployment
- Enhance server consolidation
- Enhance storage consolidation

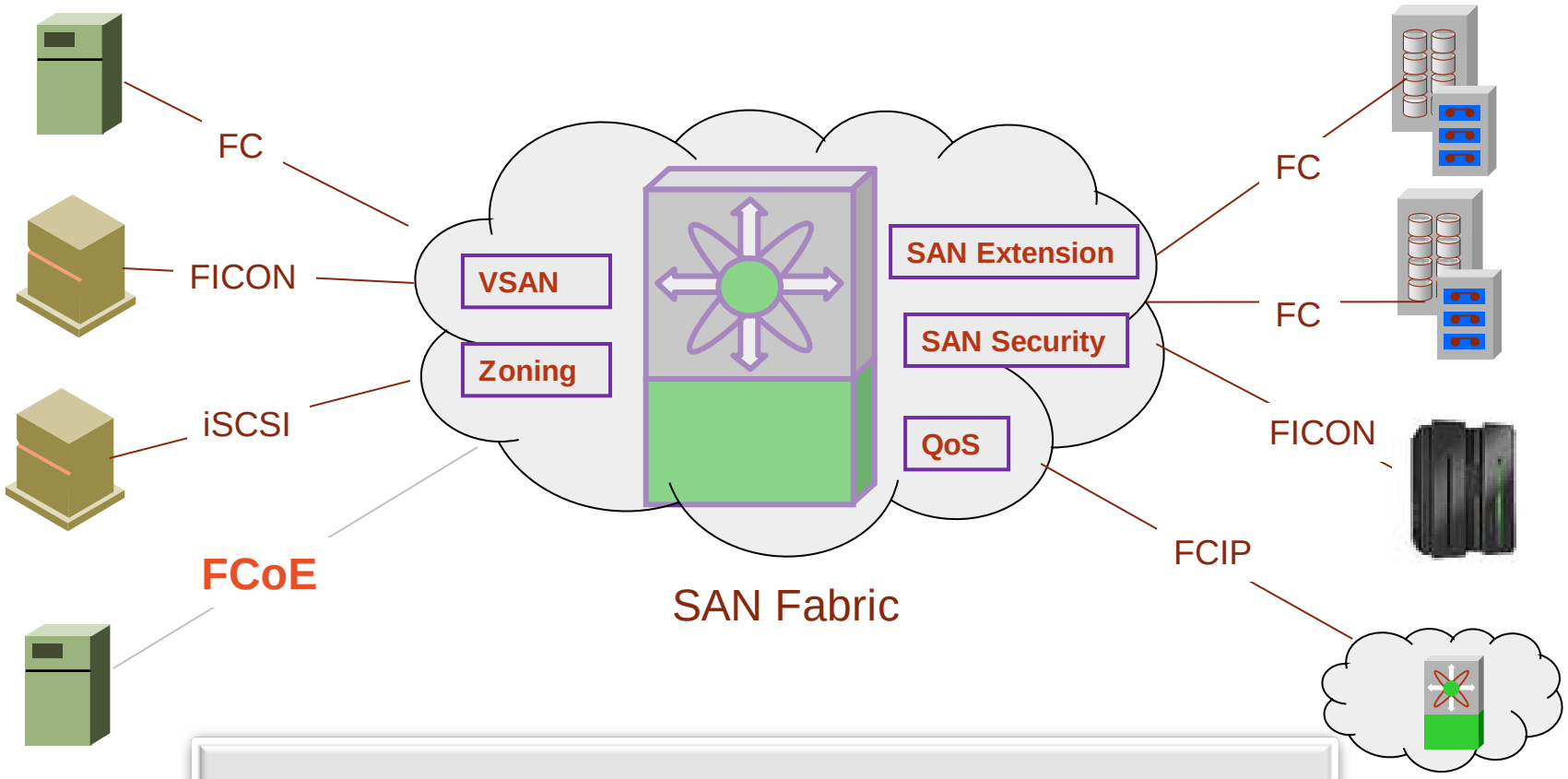
FCoE Cabling Reduction



16 Servers	Enet	FC	Total
Adapters	16	16	32
Switches	2	2	4
Cables	36	36	72
Mgmt. Pts.	2	2	4

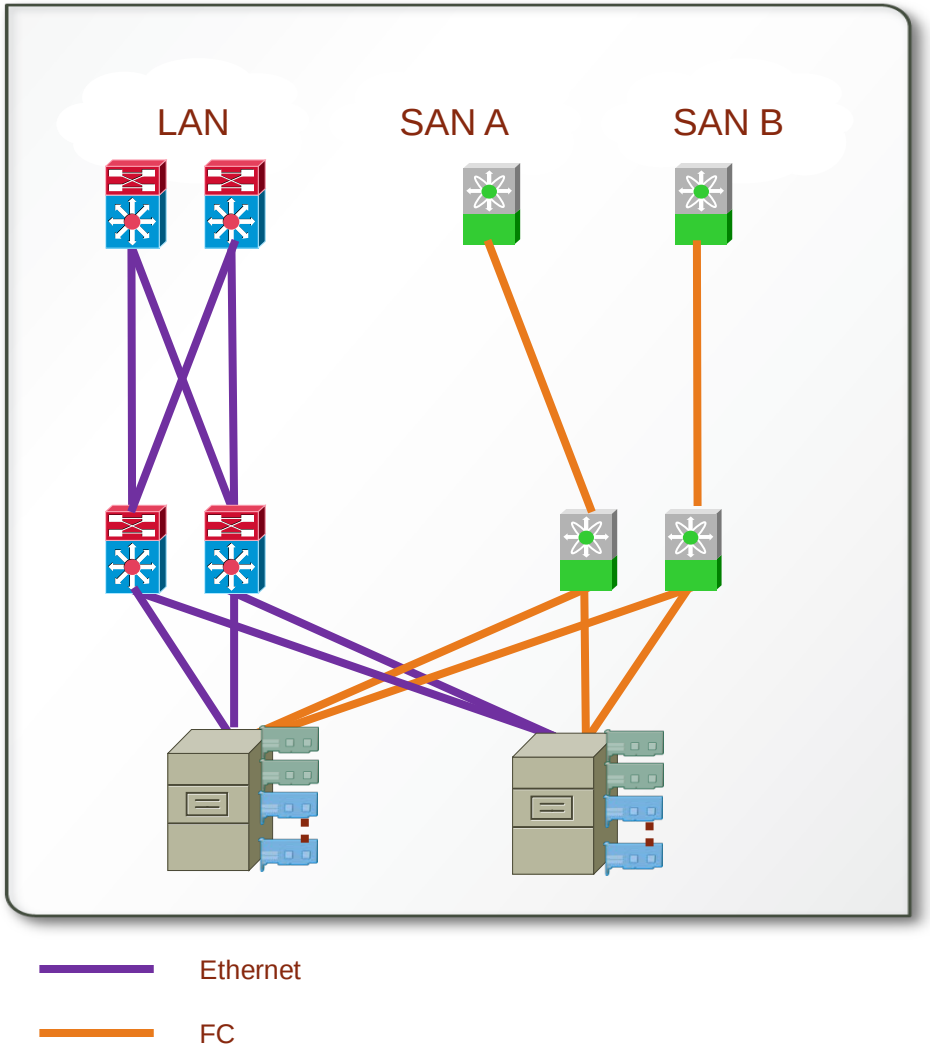
16 Servers	Enet	FC	Total
Adapters	16	0	16
Switches	2	0	2
Cables	36	4	40
Mgmt. Pts.	2	0	2

FCoE Connectivity Extends FC SANs



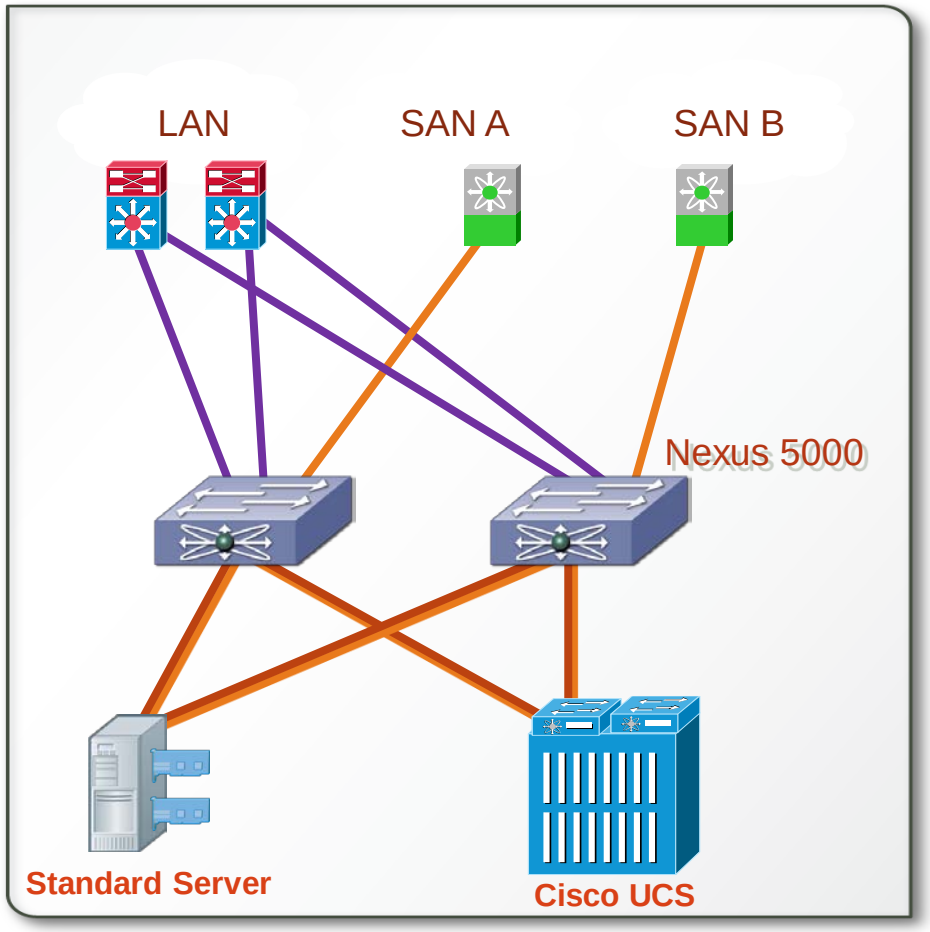
- Preserves FC investments
- Simplifies SAN-attach of servers

Before I/O Consolidation



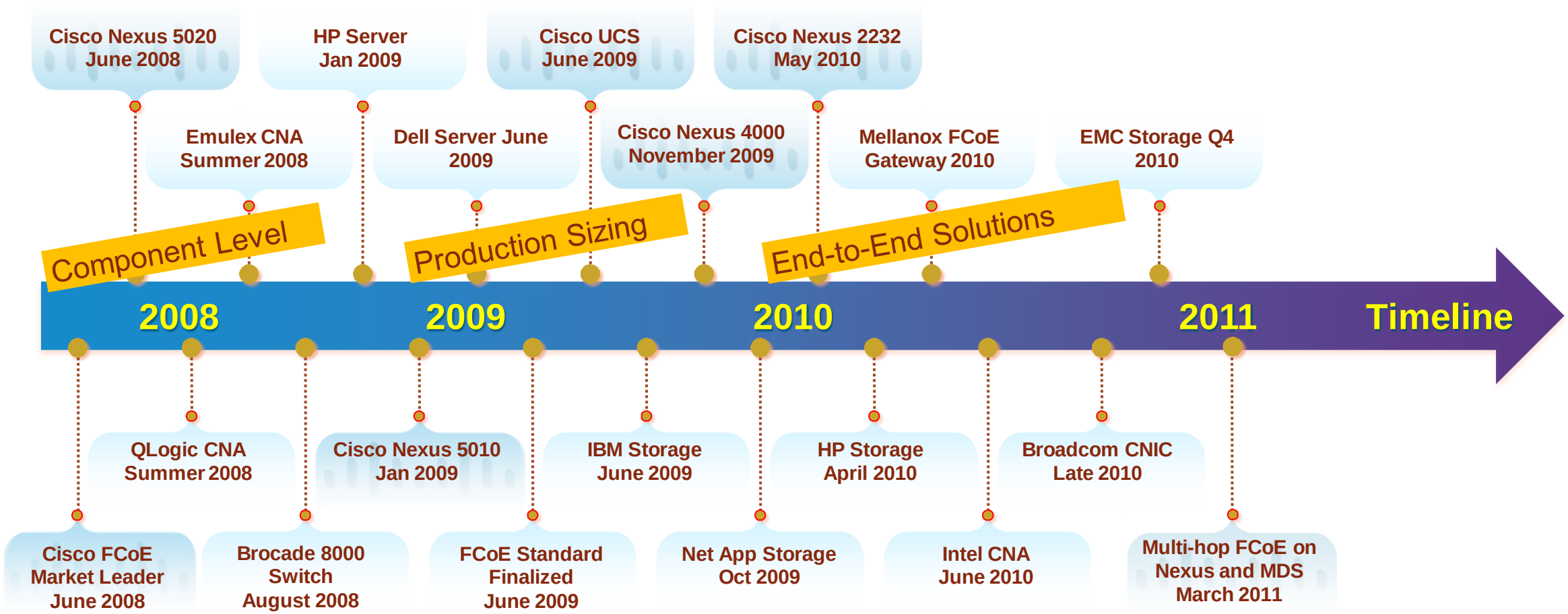
- Parallel LAN/SAN Infrastructure
- Inefficient use of Network Infrastructure
- 5+ connections per server—higher adapter and cabling costs
  - Adds downstream port costs; cap-ex and op-ex
  - Each connection adds additional points of failure in the fabric
- Longer lead time for server provisioning
- Multiple fault domains—complex diagnostics
- Management complexity

After I/O Consolidation



- Reduction of server adapters
- Simplification of access layer and cabling
- Gateway free implementation—fits in installed base of existing LAN and SAN
- L2 Multipathing Access—Distribution
- Lower total cost of ownership
- Fewer cables
- Investment protection (LANs and SANs)
- Consistent operational model

FCoE Ecosystem Timeline



Virtualization: VMware, Oracle Virtual Box, and Microsoft Hyper-V

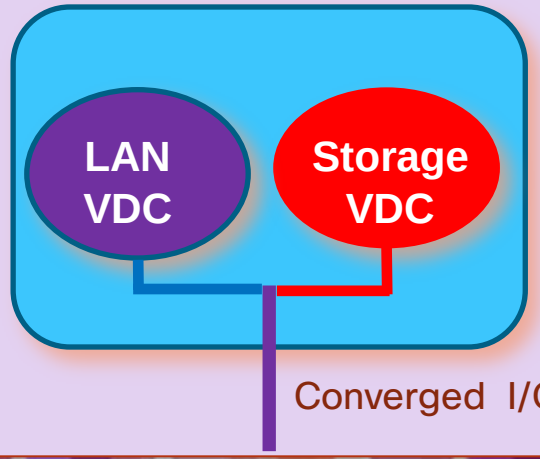
Operating Systems: Microsoft Windows Server, Red Hat Enterprise Linux, Oracle Enterprise Linux and Novell SUSE Linux Enterprise All Support FCoE

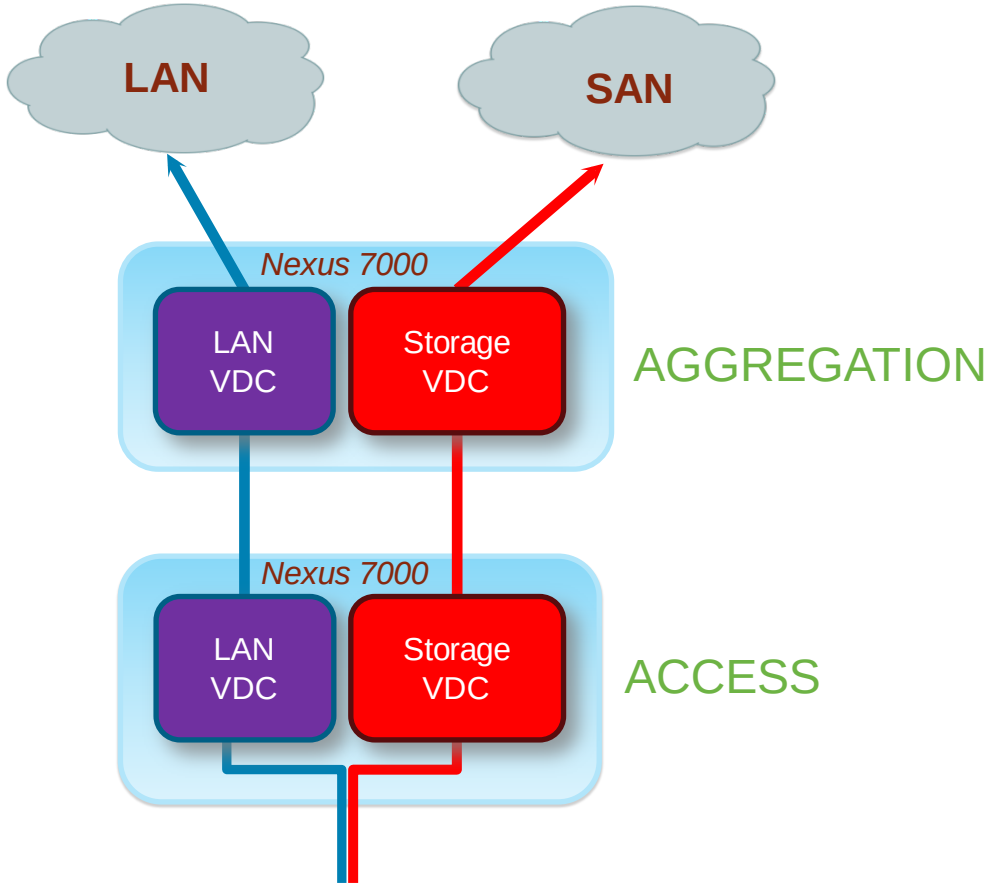
Storage VDC on the Nexus 7000

- A Virtual MDS

Dedicated Storage VDC – Converged Interfaces

- Model for host/target interfaces, not ISLs
- Separate VDC running ONLY storage related protocols
- Ingress Ethernet traffic is split based on frame ether type
- FCoE traffic is processed in the context of the Storage VDC





VDCs Offer Fault Isolation for Higher Availability

Cisco Unified IO Ecosystem

Scalability

SAN

MDS 9500

MDS 9200

MDS 9100

LAN

Nexus 1010

Nexus 3000

Nexus 1000V

LAN/SAN

Nexus 7000

Nexus 5000

Nexus 4000

Nexus B22

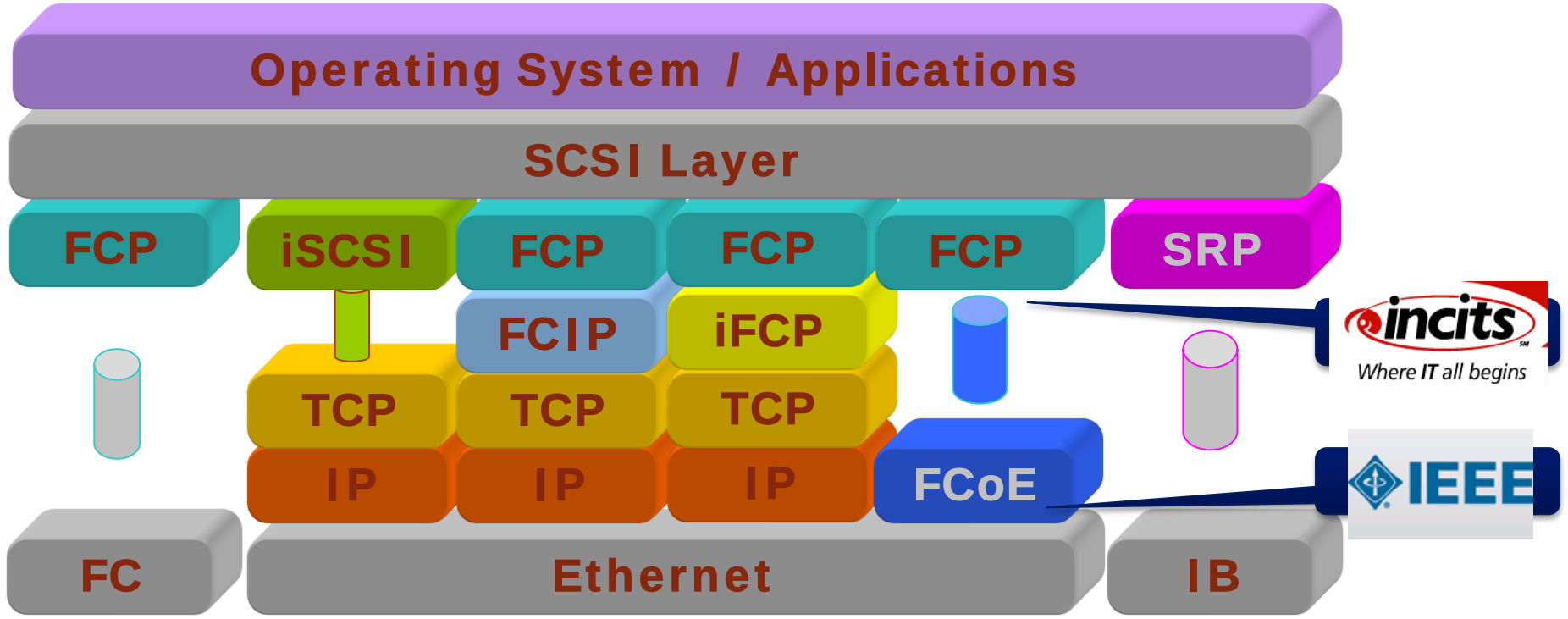
Nexus 2000

Cisco NX-OS: One OS from the Hypervisor to the Data Center Core

Convergence	VM-Aware Networking	10 GbE switching	Fabric Extensibility	Cloud Mobility

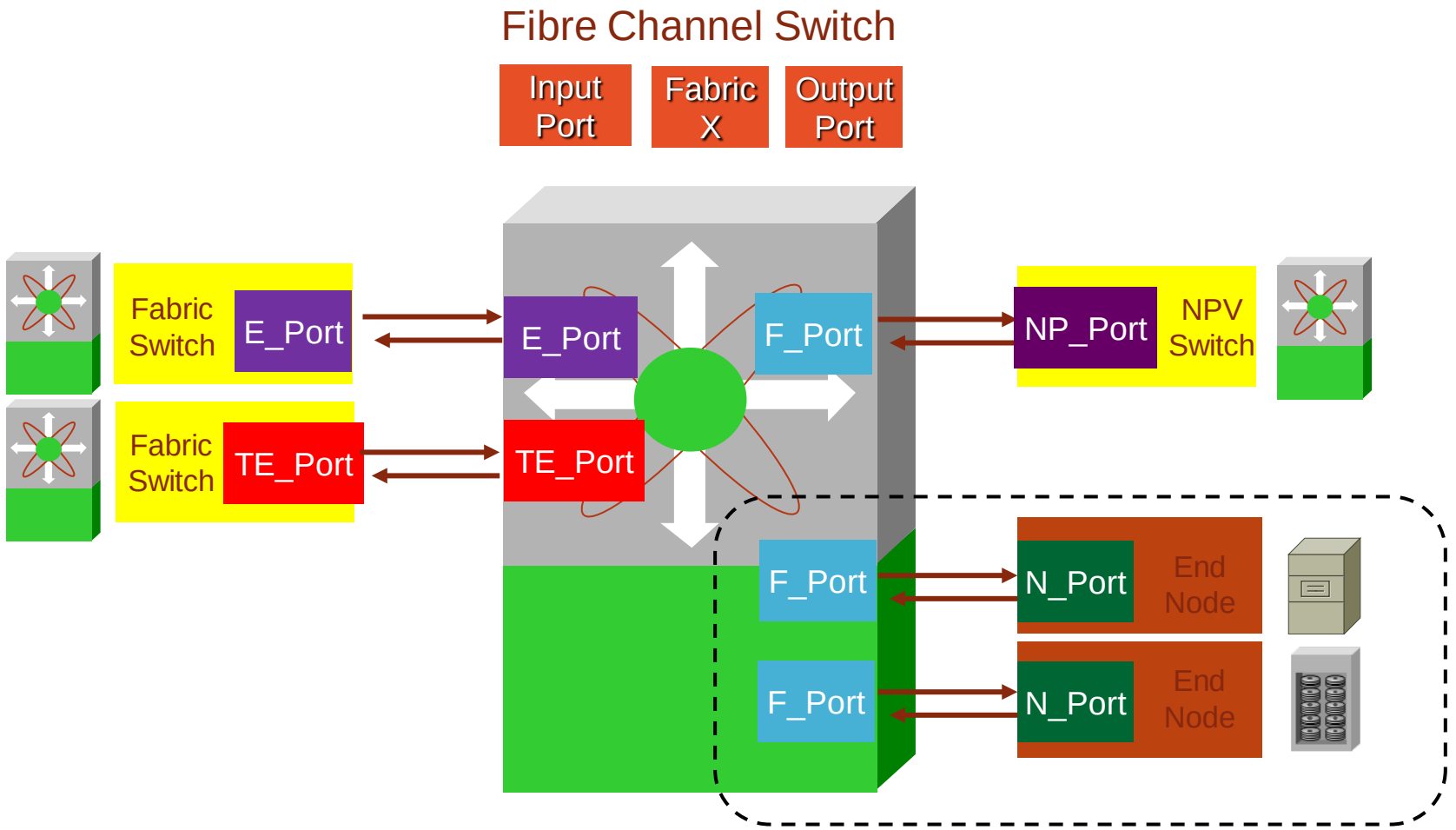
SAN Basics

Storage Protocol Technologies



FCoE Is Non Routable, Localized DC Transport Solution with Lower Protocol Overhead than FCIP or iSCSI

Fibre Channel Port Types

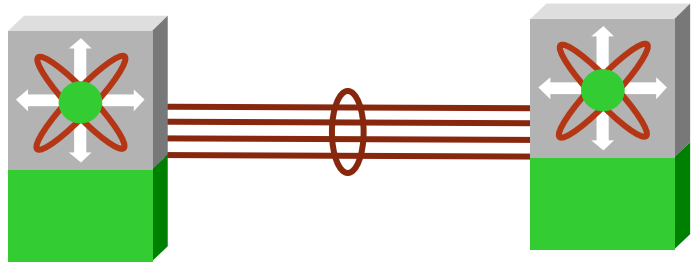


Fibre Channel Fabric Topology

- Port Trunking and Channeling

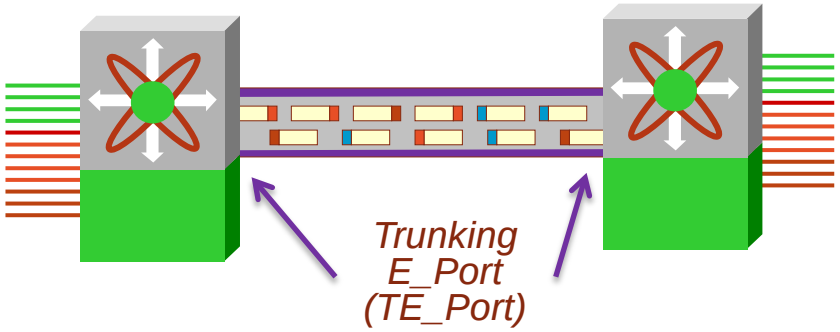
Port Channels

- Higher aggregate bandwidth
- Hardware-based load balancing
- Only supported on switch to switch connections (E_Port to E_Port and NP_Port to F_Port)



Trunking

- Trunking E_Port (TE_Port)
- Carries tagged frames from multiple VSANs
- Enhanced ISL (EISL) link

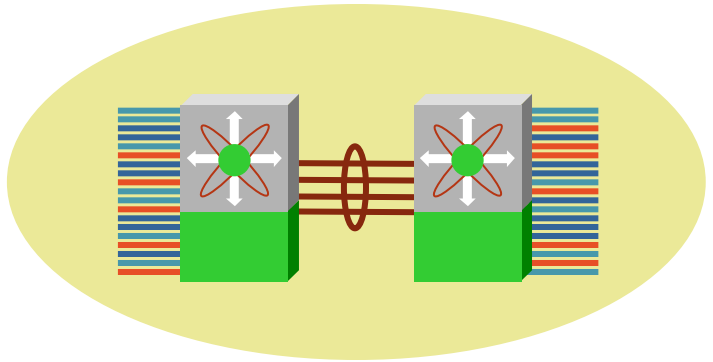


Standardization of ***Enhanced*** Capabilities is **less** common across vendors in the Fibre Channel market than You May Be Used to in the LAN market

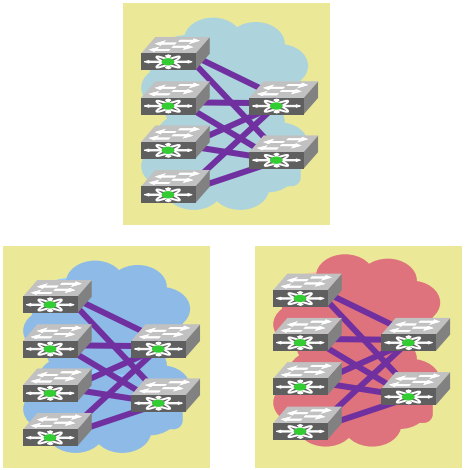
Virtual SANs (VSANs)

- VLANs or 802.1q for FC

A Virtual SAN (VSAN) Provides a Method to Allocate Ports within a Physical Fabric and Create Virtual Fabrics



Physical SAN Islands Are Virtualized onto a Common SAN Infrastructure

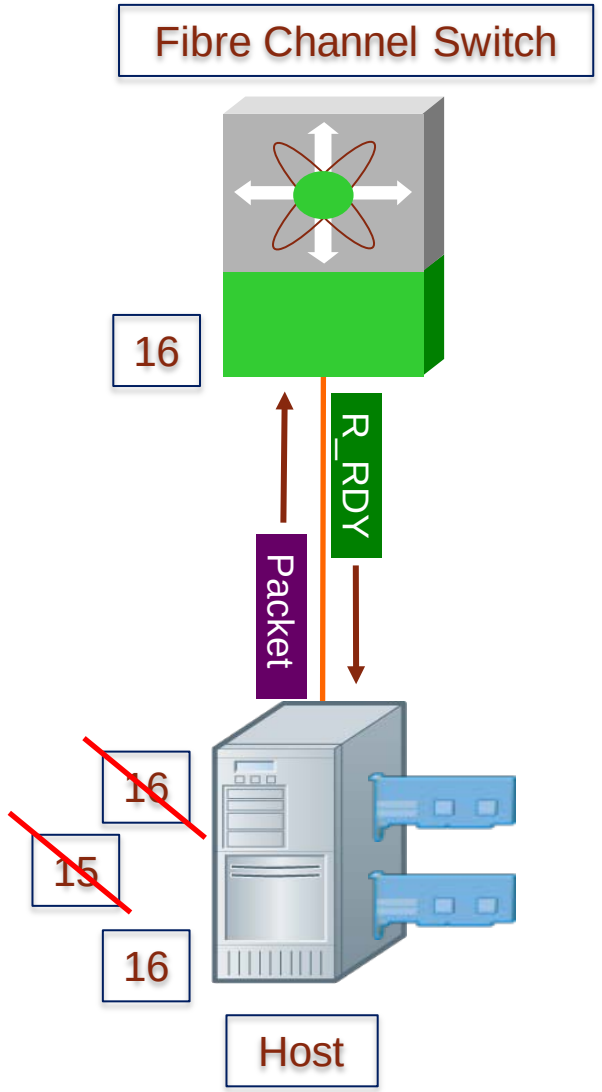


- Analogous to VLANs in Ethernet
- Virtual fabrics created from larger cost-effective redundant physical fabric
- Reduces wasted ports of a SAN island approach
- Fabric events are isolated per VSAN which gives further isolation for High Availability
- FC Features can be configured on a per VSAN basis.

Buffer to Buffer Credits

- Fibre Channel Flow Control

- Source regulated flow control
- B2B Credits used to ensure that FC transport is lossless
- # of credits negotiated between ports when link is brought up
- # Credits decremented with each packet placed on the wire
  - Independent of packet size
  - If # credits = 0, packet transmission is stopped
- # of credits incremented with each “transfer ready” received
- B2B Credits need to be taken into consideration as distance and/or bandwidth increases



Fabric Shortest Path First

- Fibre Channel Forwarding – Just Like OSPF

- FSPF routes traffic based on destination domain ID
- For FSPF a domain ID identifies a VSAN in a single switch

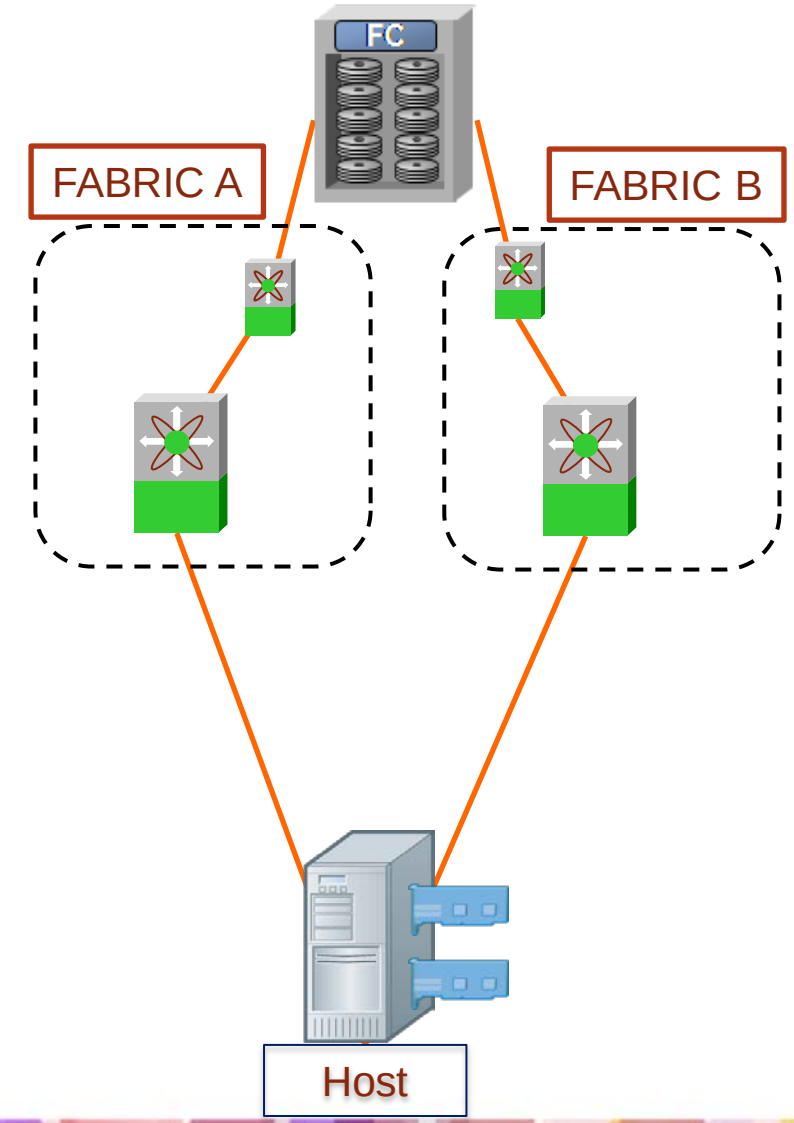
  This limits the max number of switches that can be supported in the Fabric to 239.

- FSPF performs hop-by-hop routing
- FSPF uses total cost as the metric to determine most efficient path
- FSPF supports equal cost load balancing across links
- Link cost can be manually adjusted

How Does a Host Access its' Storage?

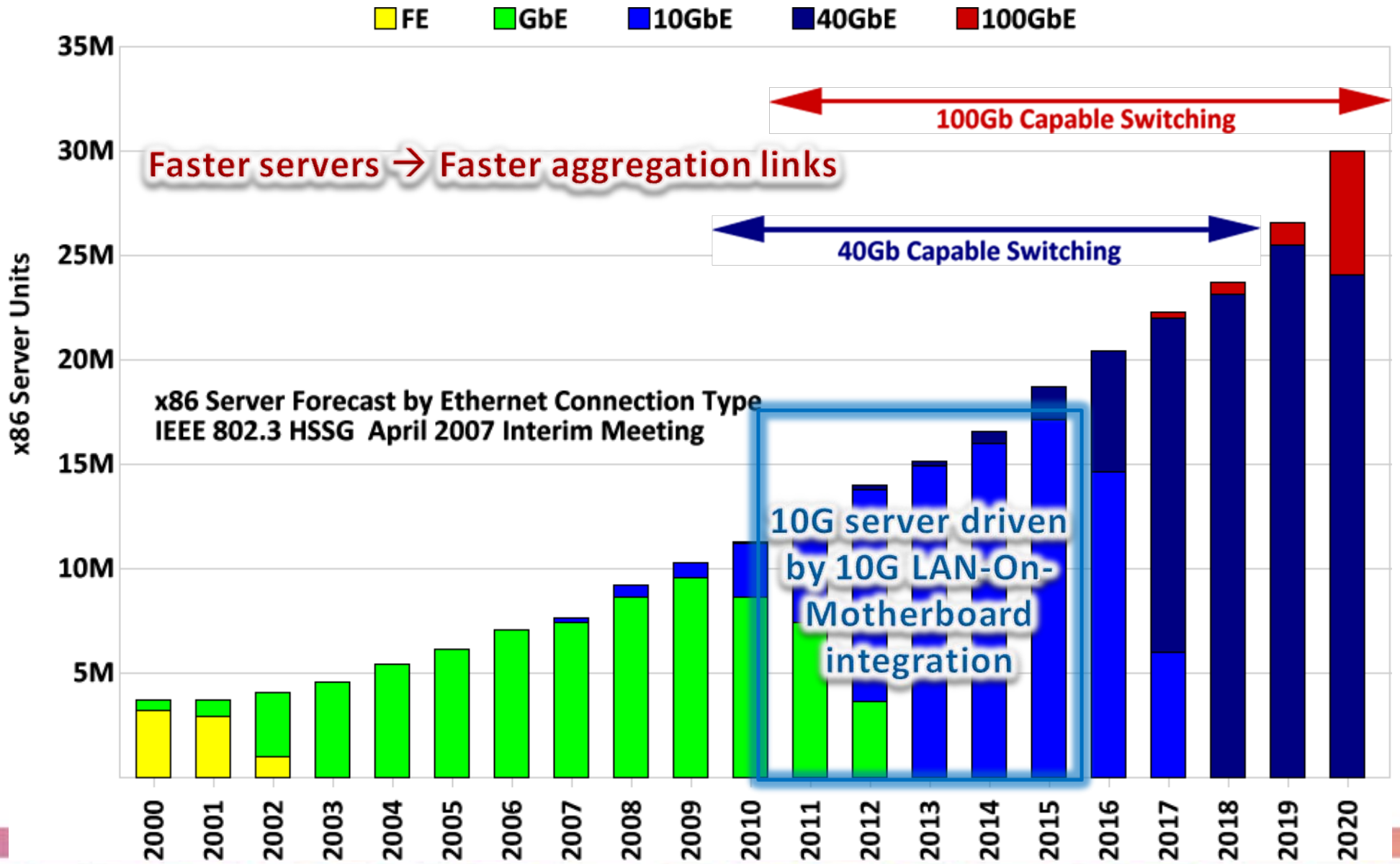
- In Beginning…

- Start with a host and a target that need to communicate
  - Typical Host has 2 HBA's (one per fabric) each with a WWN
  - Target has multiple ports to connect to multiple fabrics
- Connect them to a FC Switch
  - Port Type Negotiation
  - Speed Negotiation
- FC Switch is part of the SAN "fabric"
- Most commonly, dual fabrics are deployed for redundancy



Ethernet Enhancements

High-Speed Ethernet Server Adoption

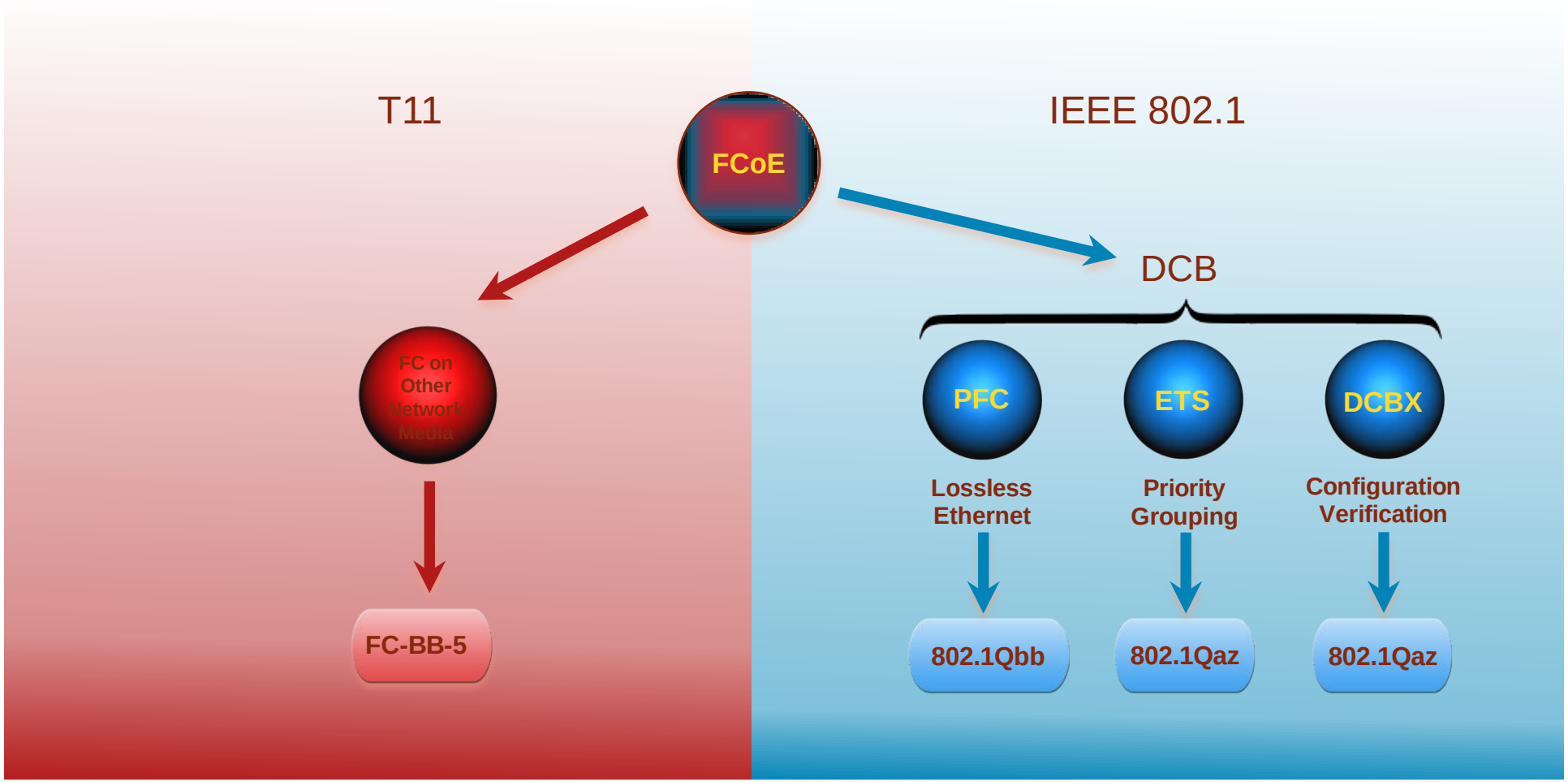


Why 10 Gigabit Ethernet to the Server?

- Multi core CPU architectures driving increased network bandwidth demands
- Virtual Machines driving increased I/O connections and I/O bandwidth per server
- Low latency 10GE affordability (even optics…)
- Increased adoption of NAS (NFS/CIFS) and iSCSI
- Consolidation of networks
  - Unified Fabrics with Ethernet
- Ubiquity of large scale Ethernet networks
  - Extensive range of management, diagnostic and troubleshooting tools
  - Massive base of manufacturers, suppliers and integrators
  - → Competition, price, services and innovation

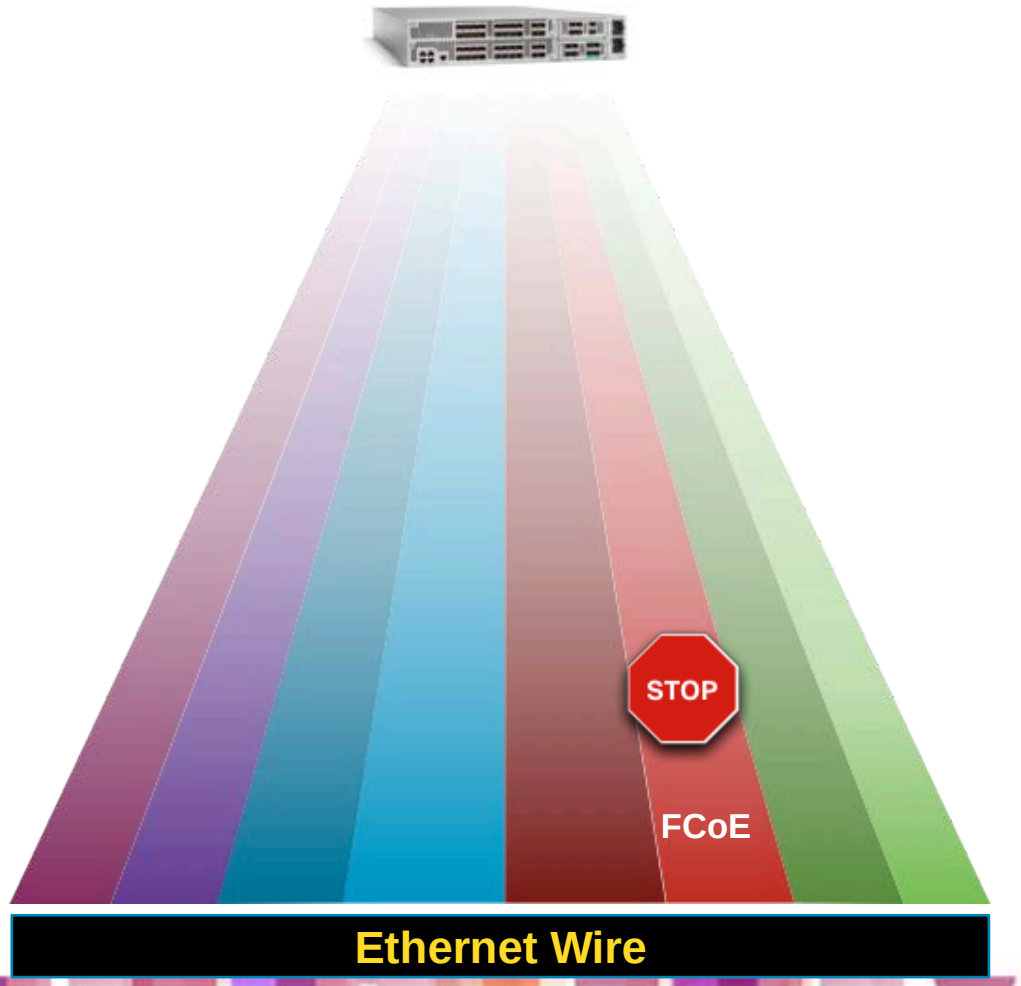
Standards for FCoE

- FCoE is **fully** defined in FC-BB-5 standard
- FCoE works alongside additional technologies to make I/O Consolidation a reality



PFC: Priority Flow Control

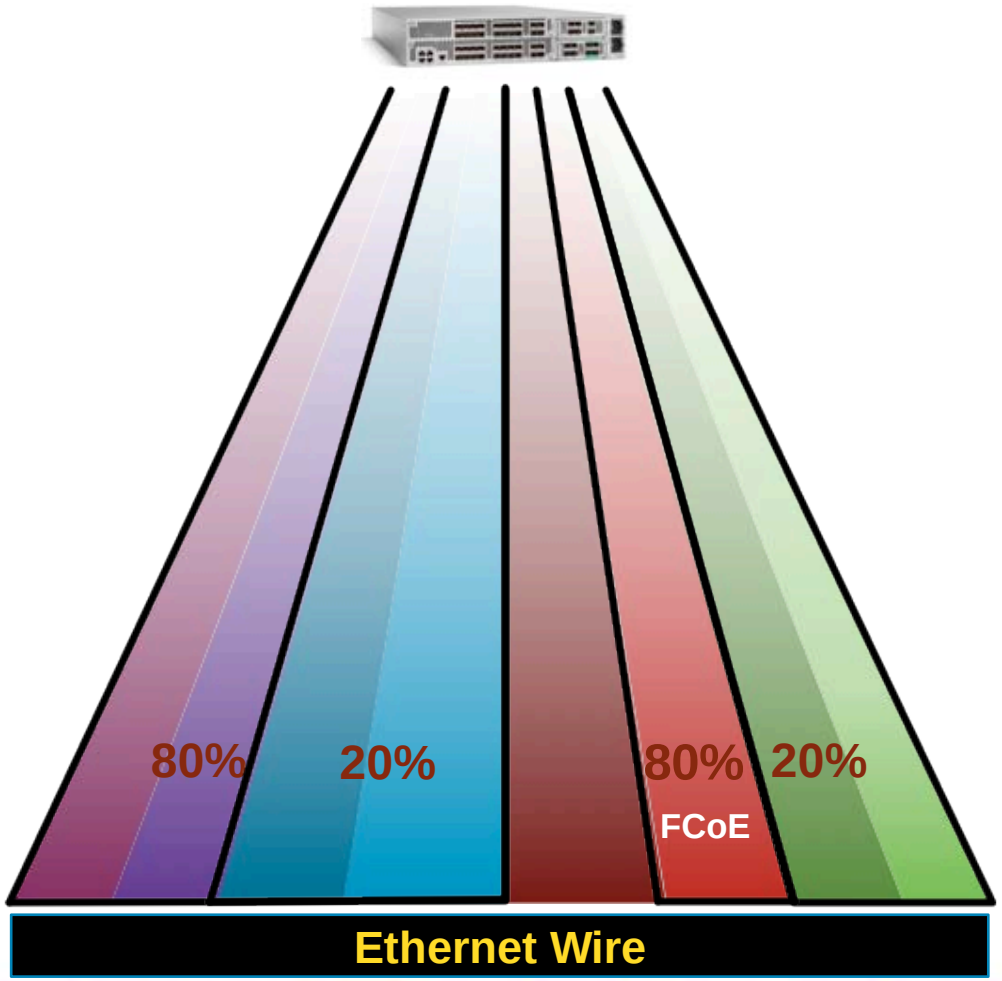
- IEEE 802.1Qbb



- VLAN Tag enables 8 priorities for Ethernet traffic
- PFC enables Flow Control on a Per-Priority basis using PAUSE frames (802.1p).
- Therefore, we have the ability to have lossless and lossy priorities at the same time on the same wire
  - Allows FCoE to operate over a lossless priority independent of other priorities

ETS: Enhanced Transmission Selection

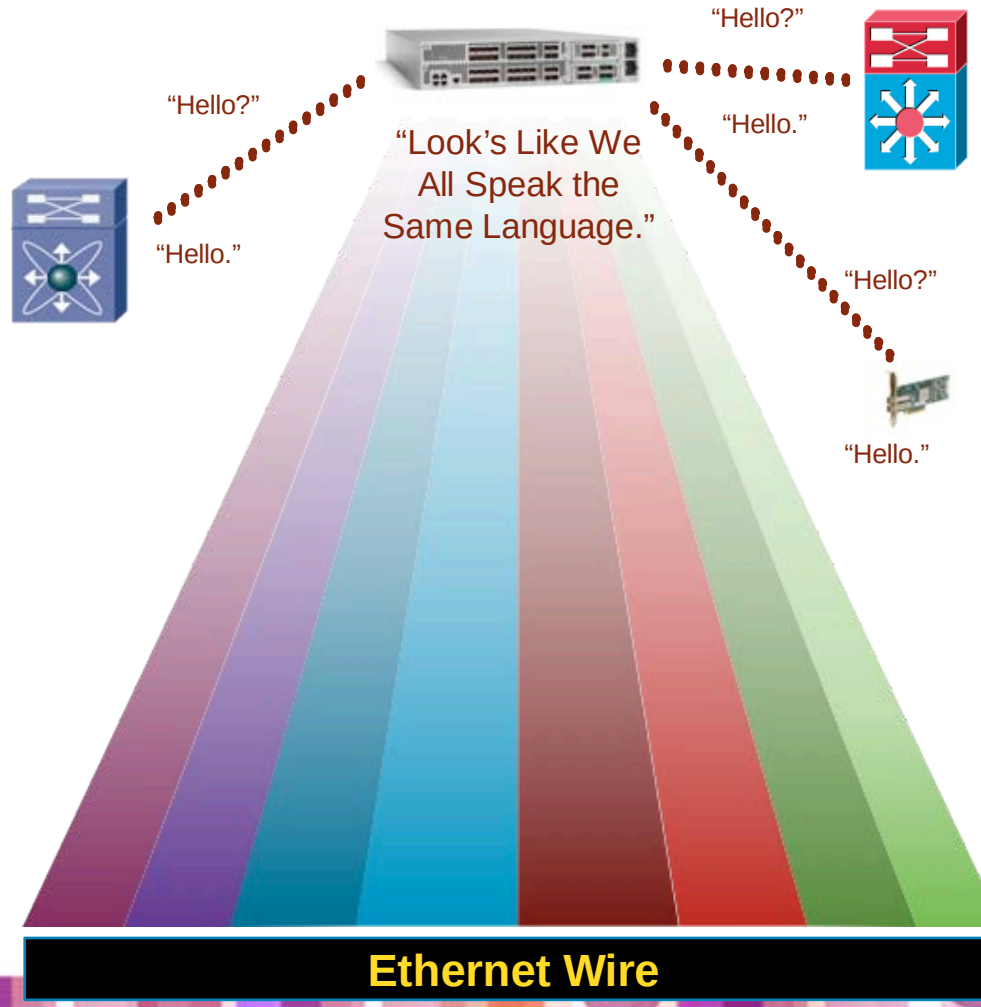
- IEEE 802.1Qaz



- Allows you to create priority groups
- Can guarantee bandwidth
- Can assign bandwidth percentages to groups
- Not all priorities need to be used or in groups

DCBX: Data Center Bridging eXchange

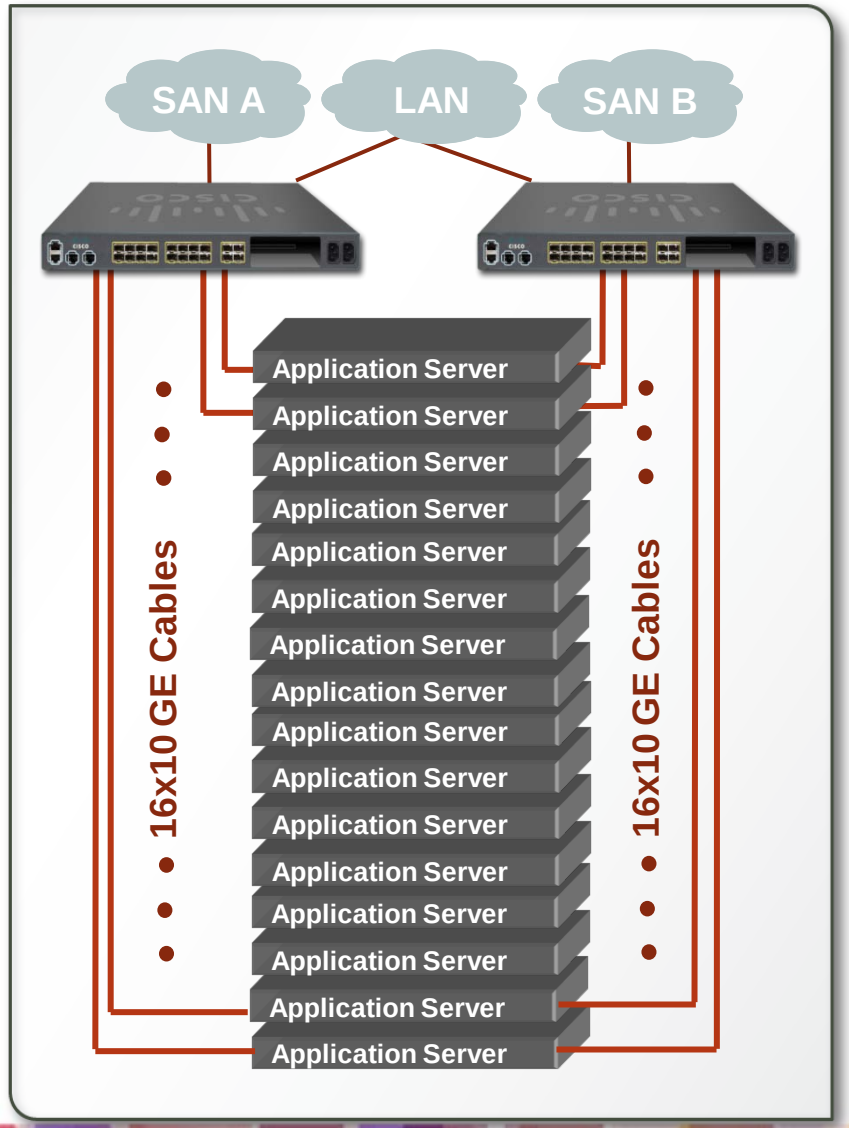
- IEEE 802.1Qaz



- Allows network devices to advertise their identities and capabilities over the network
  - Enables hosts to pick up proper configuration from the network
  - Enables switches to verify proper configuration
- Provides support for:
  - PFC
  - ETS
  - Applications (e.g., FCoE)

Cabling: Twinax Copper Cable

- Low power consumption
- Low cable cost
  - SFP+ CX-1 Copper (SFF 8431)
- Low transceiver latency
- Low error rate (10^{17})
- Thinner cable with higher bend radius
  - Supports 10GE passive direct attached up to 10 meters
- Easier to manage cabling solution reduces deployment time
- All copper cables are contained within the rack



Converged Network Adapters

LOM

First Generation Mid-2008

QLOGIC
EMULEX

- Multiple components
- Full height/length

Second Generation Mid-2009

QLOGIC
EMULEX

- Single chip
- Half height/length
- Less than half the power
- Same support as HBAs

10GE Today FCoE in Software

intel

Open-FCoE.org

And Other NIC Vendors

- Brings the economy of scale of Ethernet to SAN
- Field upgrade-able NICs

FCoE Basics

What Is FCoE?

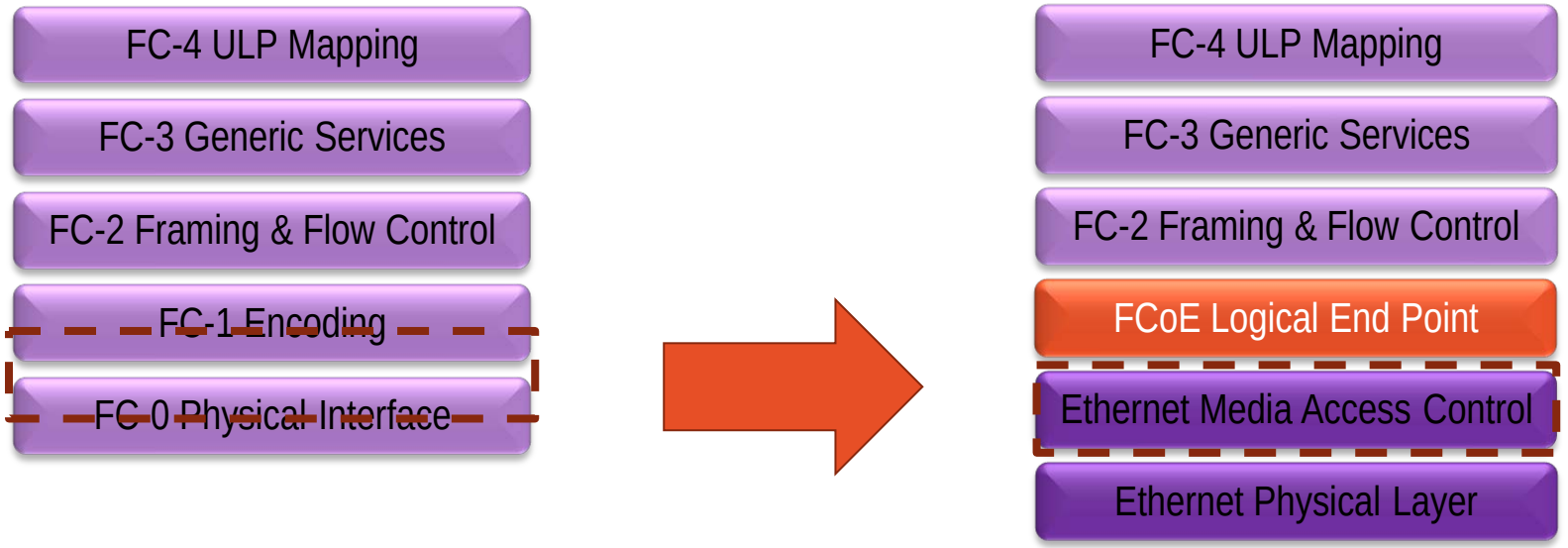
- It's Fibre Channel

- From a Fibre Channel standpoint it's

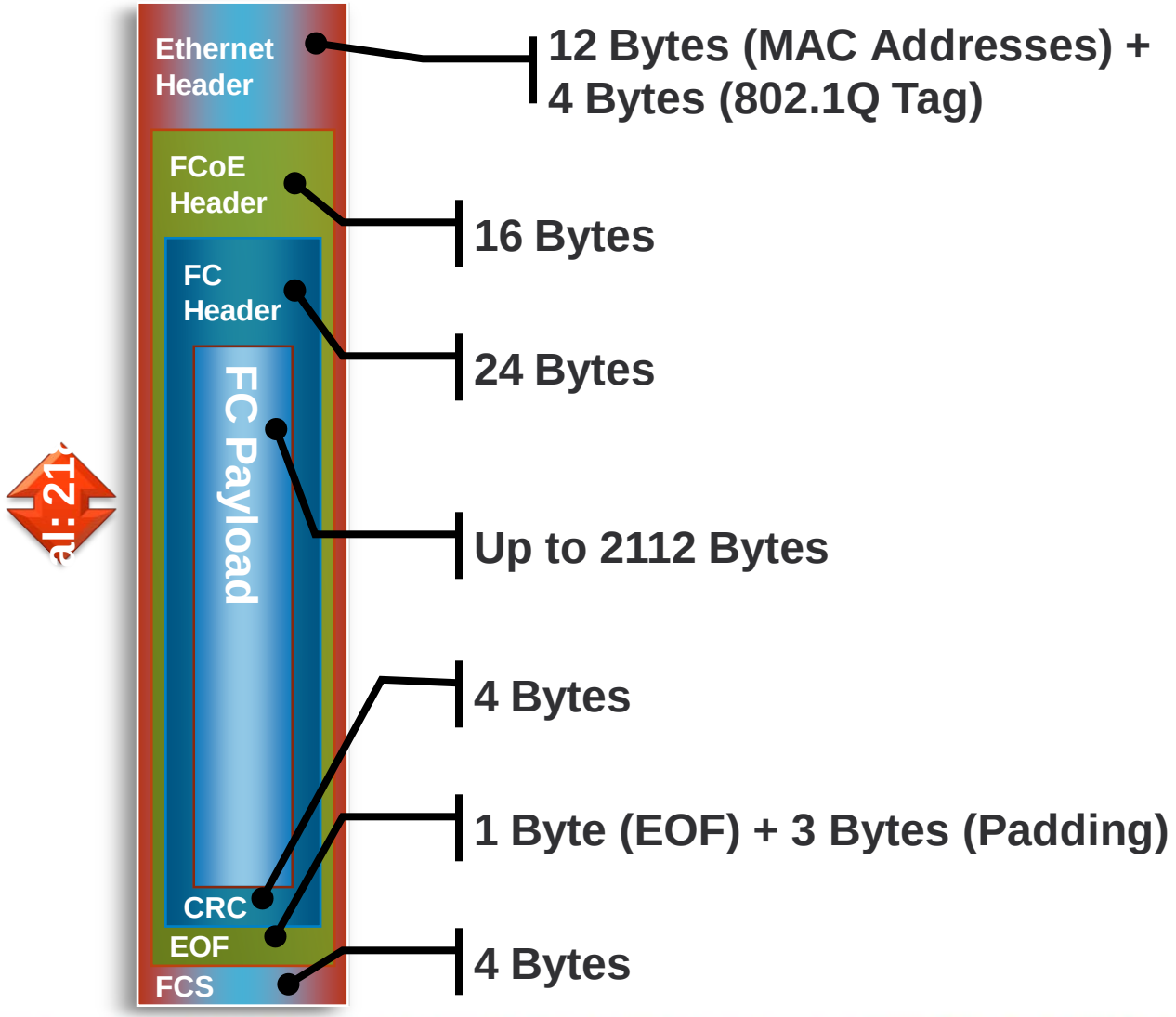
  FC connectivity over a new type of cable called… Ethernet

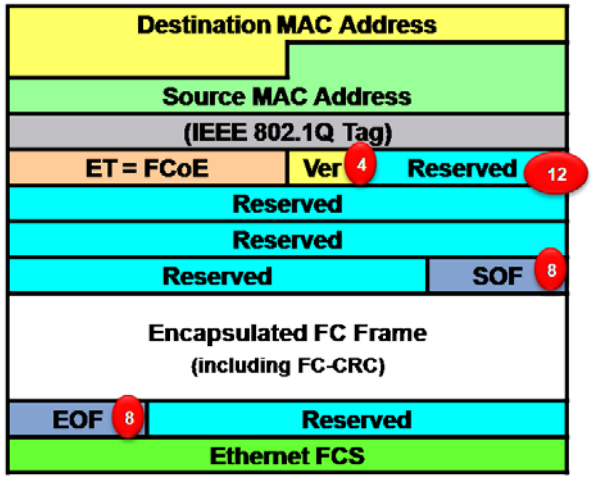
- From an Ethernet standpoints it's

  Yet another ULP (Upper Layer Protocol) to be transported



FCoE Frame



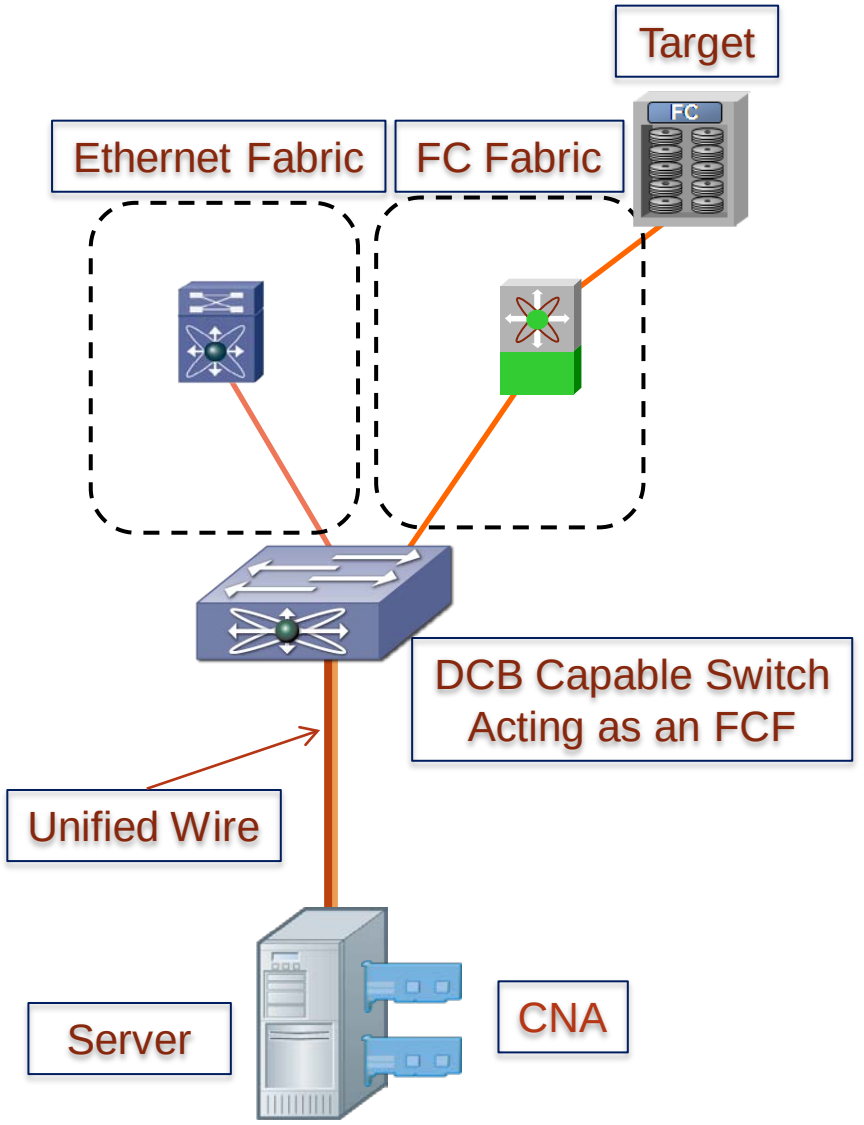


FCoE Standard (FC-BB-5)
Requires Jumbo Support;
2.5KB = “Baby Jumbo”

FCoE Designs

FCoE Designs: Same Model as FC

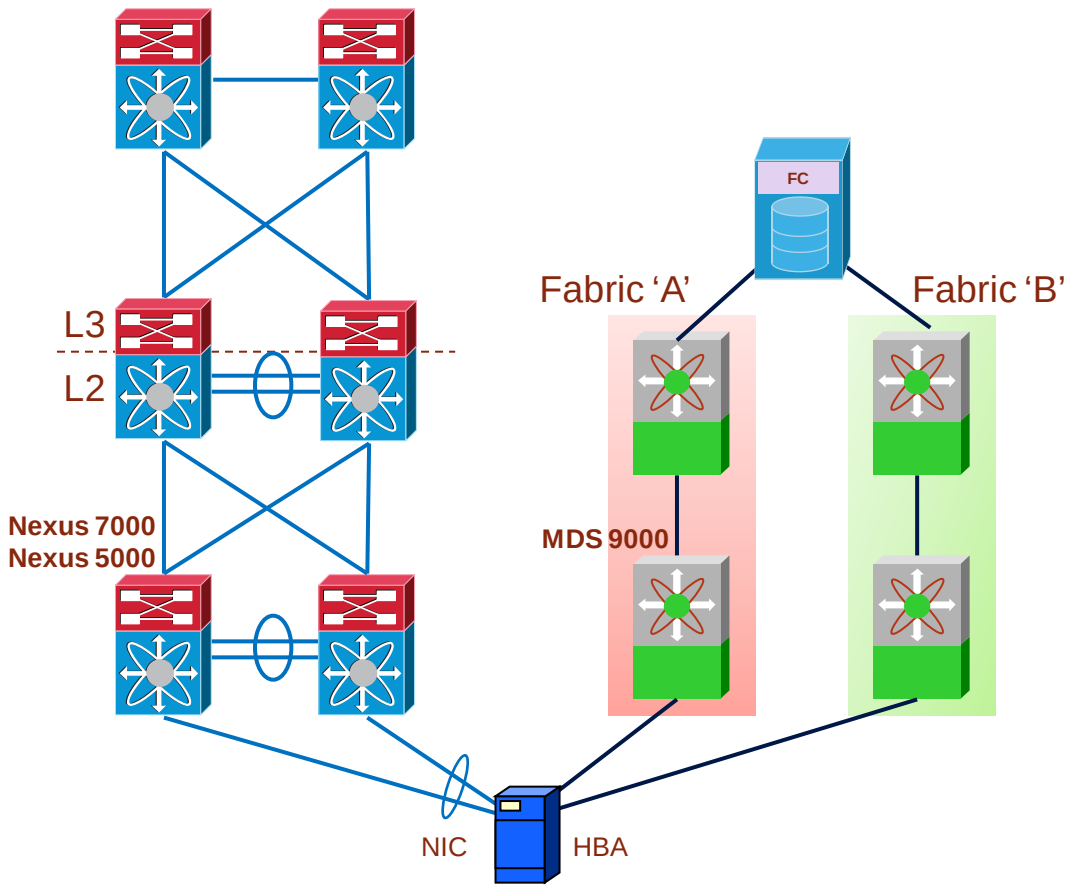
- Same host to target communication
  - Host has 2 CNA's (one per fabric)
  - Target has multiple ports to connect to fabric
- Connect to a DCB capable switch
  - Port Type Negotiation (FC port type will be handled by FIP)
  - Speed Negotiation
  - DCBX Negotiation
- Access switch is a Fibre Channel Forwarder (FCF)
- Dual fabrics are still deployed for redundancy



Traditional Data Center Design

- Ethernet LAN and Fibre Channel SAN

- Physical and Logical separation of LAN and SAN traffic
- Additional Physical and Logical separation of SAN fabrics

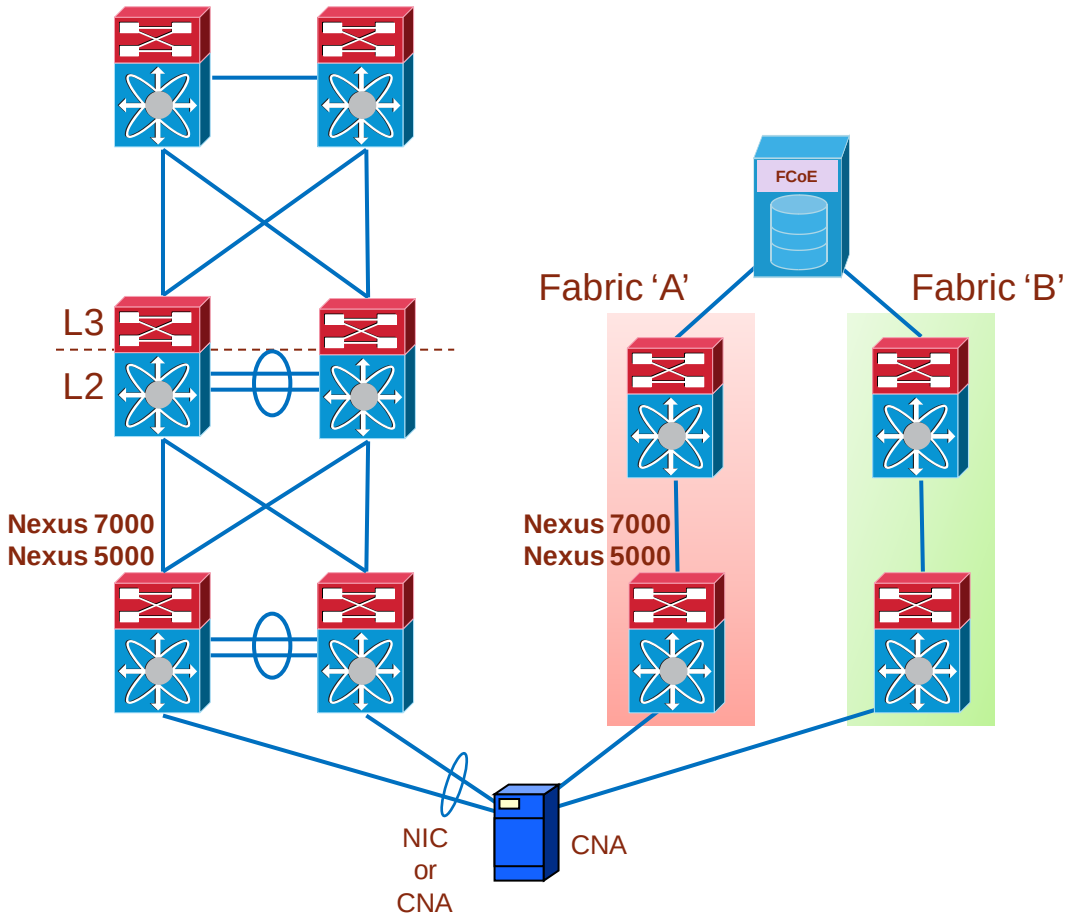


Isolation — Convergence

Data Center Design with Eth-SAN

- Ethernet LAN and Ethernet SAN

- Same topologies as existing networks, but using Nexus Unified Fabric Ethernet switches for SANs
- Physical and Logical separation of LAN and SAN traffic
- Additional Physical and Logical separation of SAN fabrics
- Ethernet SAN Fabric carries FC/FCoE & IP based storage (iSCSI, NAS, …)
- Common components: Ethernet Capacity, common parts sparing and Cost

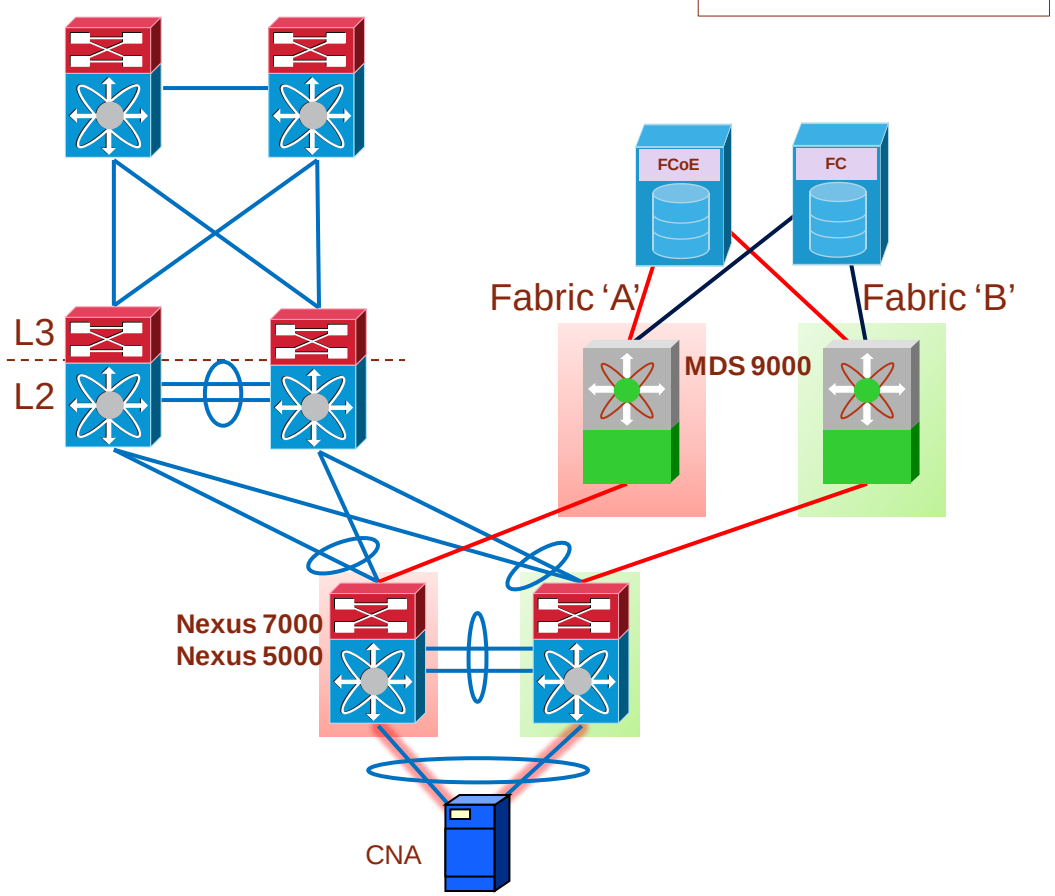


Isolation — Convergence

Converged Access

- Sharing Access Layer for LAN and SAN

- Shared Physical, Separate Logical LAN and SAN traffic at Access Layer
- Physical and Logical separation of LAN and SAN traffic at Aggregation Layer
- Additional Physical and Logical separation of SAN fabrics
- Storage VDC for additional management and operation separation
- Higher I/O, HA, fast re-convergence for host LAN traffic

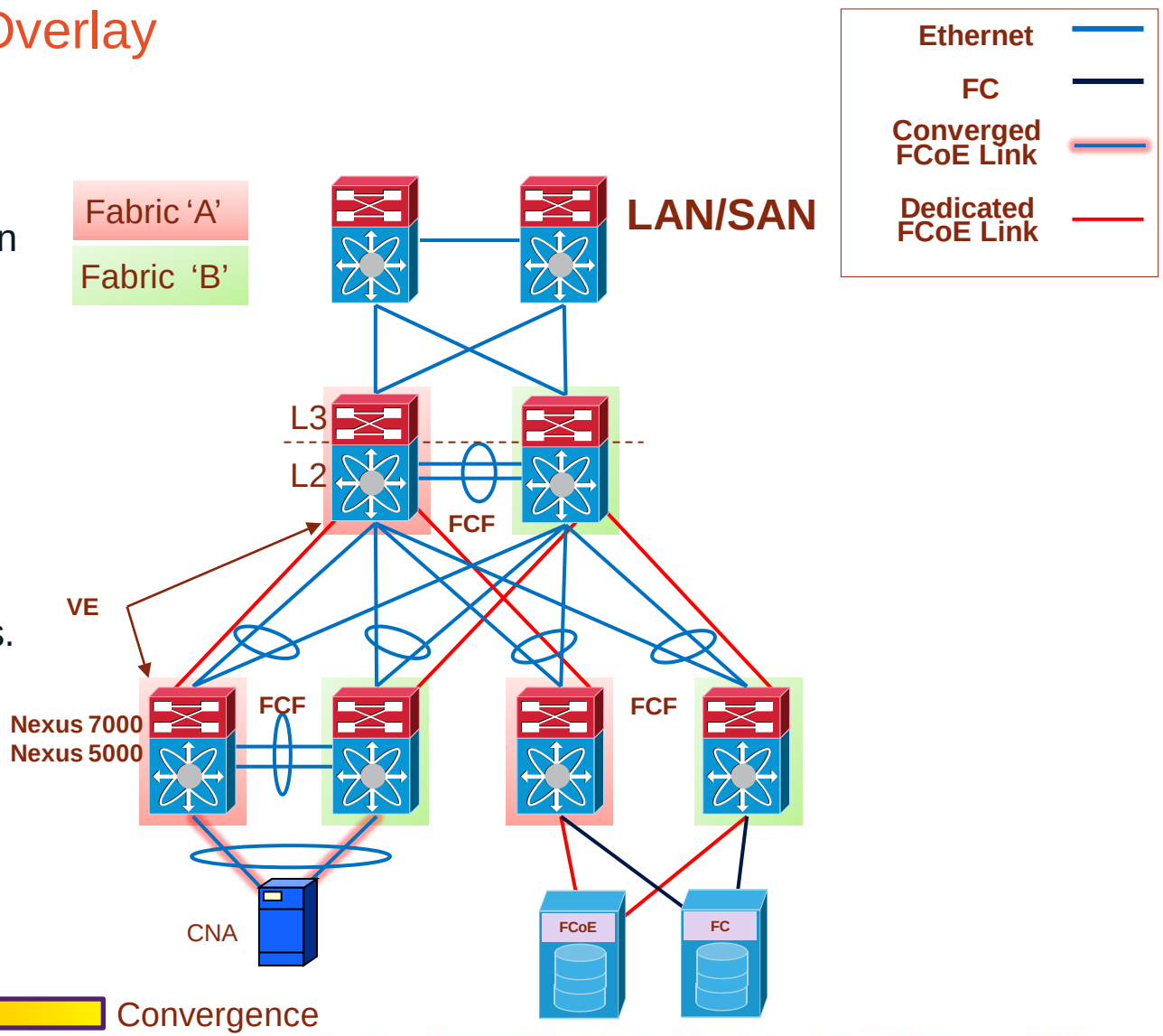


Isolation — Convergence

Converged Network Fabrics with Dedicated Links

Maintaining Dual SAN Fabrics with Overlay

- LAN and SAN traffic share physical switches
- LAN and SAN traffic use dedicated links between switches
- All Access and Aggregation switches are FCoE FCF switches
- Dedicated links between switches are VE_Ports
- Storage VDC for additional operation separation at high function agg/core
- Improved HA, load sharing and scale for LAN vs. traditional STP topologies
- SAN can utilize higher performance, higher density, lower cost Ethernet switches for the aggregation/core

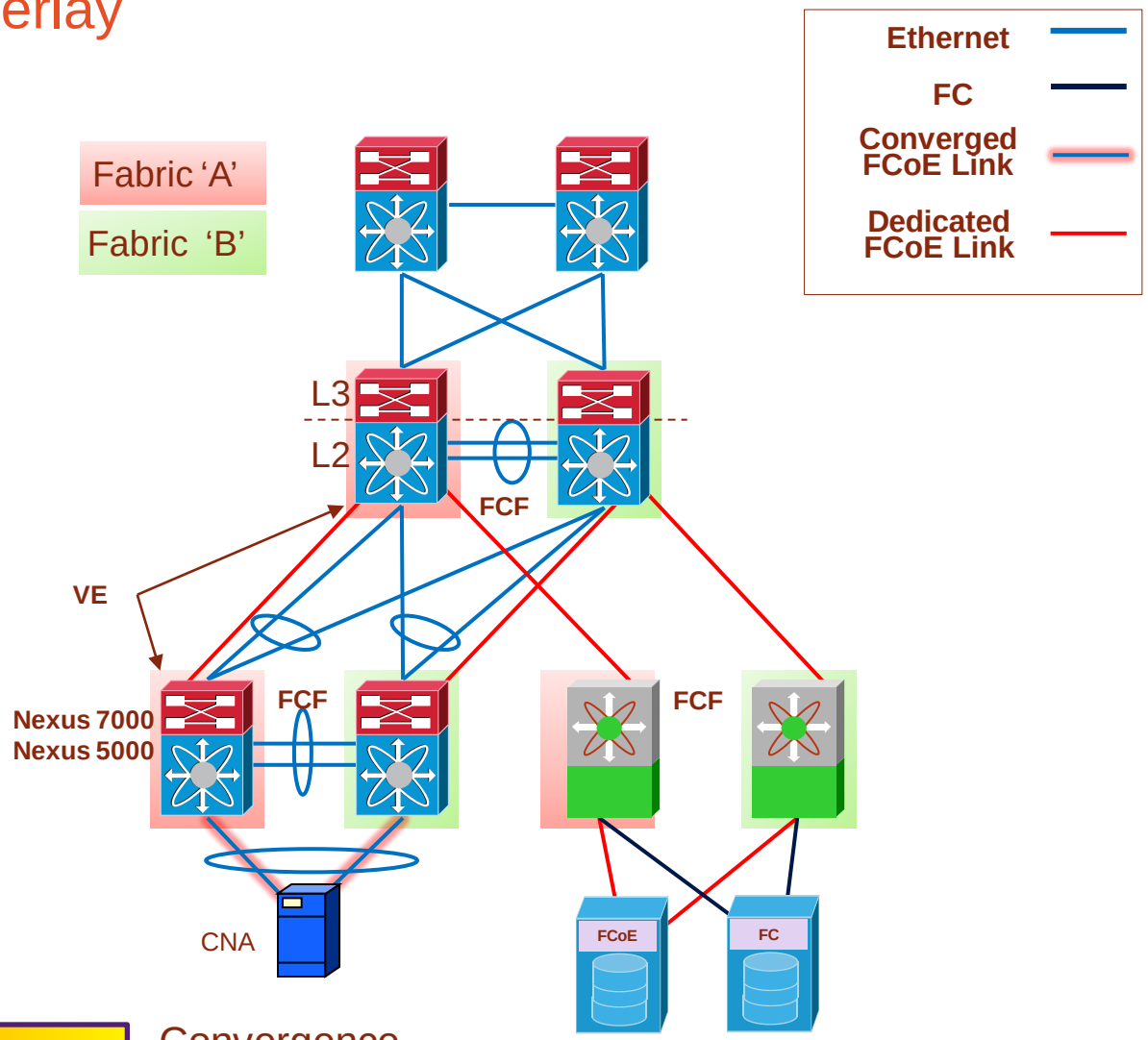


Isolation — Convergence

Converged Network Fabrics with Dedicated Links

Maintaining Dual SAN Fabrics with Overlay

- LAN and SAN traffic share physical switches
- LAN and SAN traffic use dedicated links between switches
- All Access and Aggregation switches are FCoE FCF switches
- Dedicated links between switches are VE_Ports
- Storage VDC for additional operation separation at high function agg/core
- Improved HA, load sharing and scale for LAN vs. traditional STP topologies
- SAN can utilize higher performance, higher density, lower cost Ethernet switches for the aggregation/core

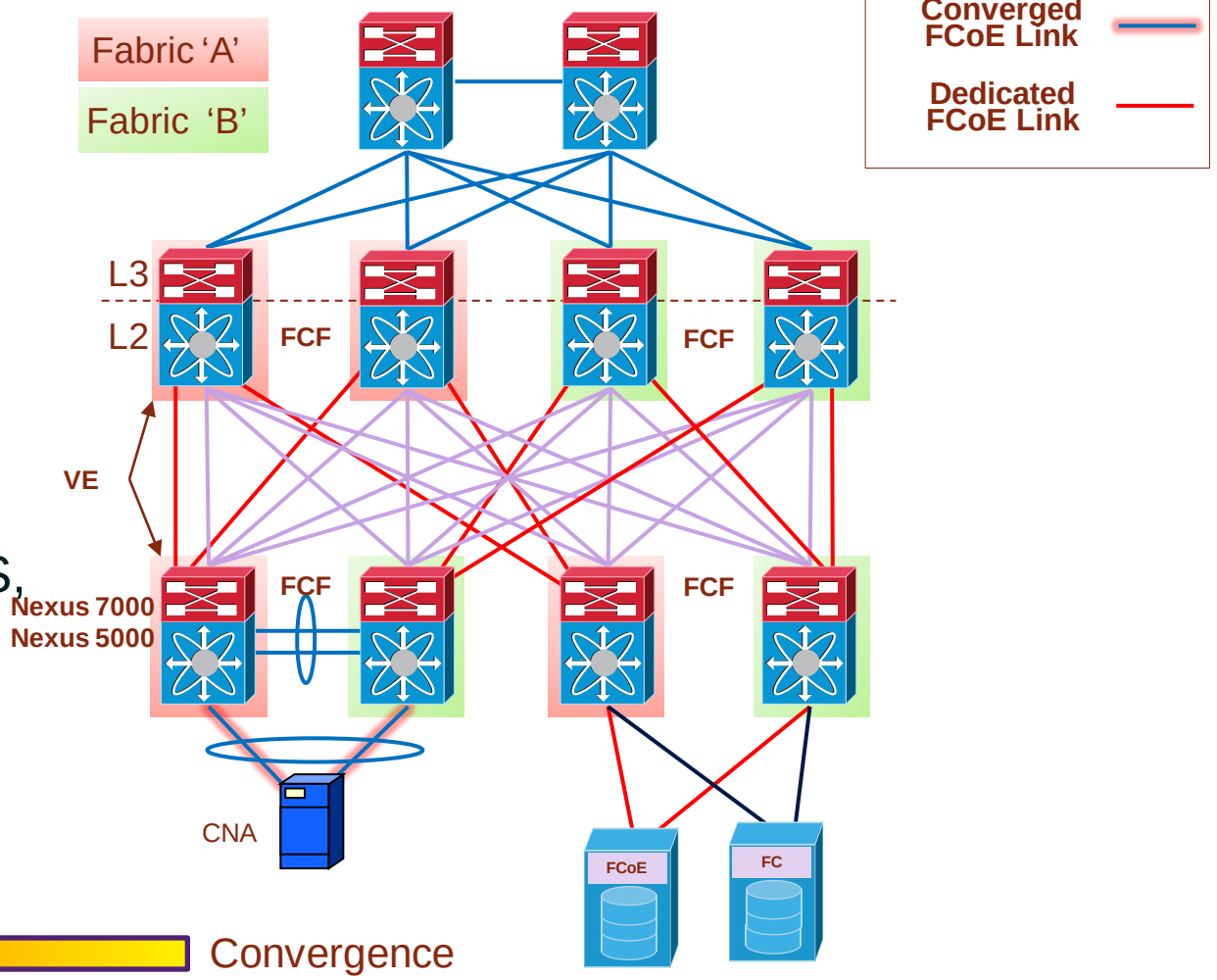


Isolation — Convergence

Converged Network with Dedicated Links

- Maintaining Dual SAN Fabrics with FabricPath

- FabricPath enabled for LAN traffic
- Dual Switch core for SAN A & SAN B
- All Access and Aggregation switches are FCoE FCF switches
- Dedicated links between switches are VE_Ports
- Storage VDC for additional operation separation at high function agg/core
- Improved HA and scale over vPC (ISIS, RPF, … and N+1 redundancy)
- SAN can utilize higher performance, higher density, lower cost Ethernet switches

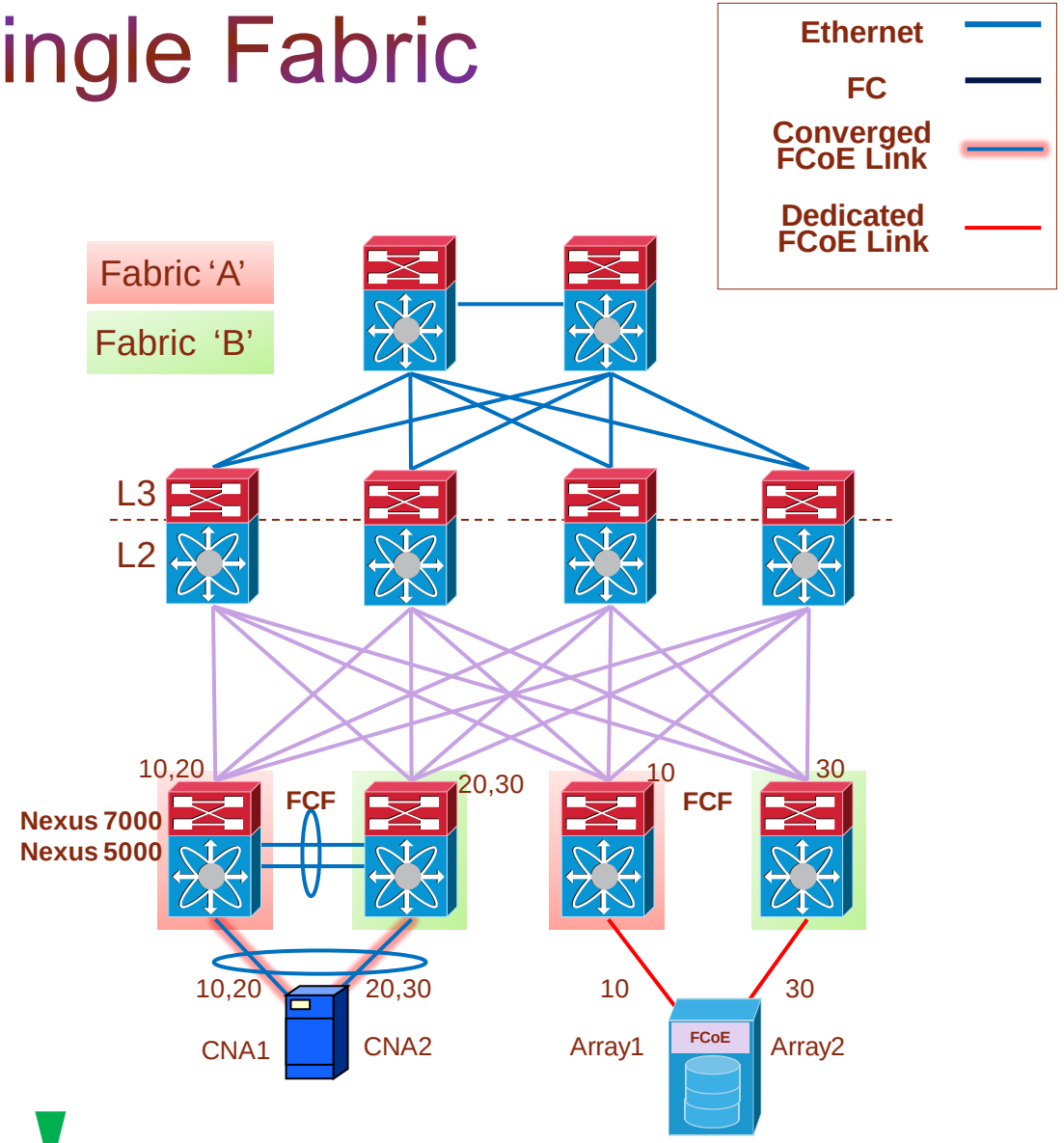


Isolation — Convergence

Converged Network – Single Fabric

- SAN Separation at the Access Switch

- LAN and SAN traffic share physical switches and links
- FabricPath enabled
- All Access switches are FCoE FCF switches
- VE_Ports to each neighbor Access switch
- Single process and database (FabricPath) for forwarding
- Improved (N + 1) redundancy for LAN & SAN
- Sharing links increases fabric flexibility and scalability
- Distinct SAN ‘A’ & ‘B’ for zoning isolation and multipathing redundancy



*Not currently supported

Isolation — Convergence

Q&A
Wrap Up

Please fill evaluation forms

Thank you.