#SHAREorg

SHARE
Technology · Connections · Results

Datacenter Networking Convergence Trends and Directions in the Datacenter August 2012

Tony Almeida
talmeida@cisco.com

12081: Datacenter Networking Convergence - Trends and Directions

SHARE in Anaheim 2012

Abstract

This trends and directions session will discuss where Datacenter networking is going. Included in this discussion with be FC/FICON and the evolution of these protocols with things like HPF, Next Generation Ethernet to allow consolidation of SAN and LAN traffic to a single interconnect (DCE), as well as other technologies on the horizon.

Both zSeries and Open environments will be discussed.

IT Infrastructure is Reaching a Breaking Point

85% Idle

In distributed computing environments, 85% of computing capacity sits idle

70¢ per 1$

70% on average is spent on maintaining infrastructure rather than adding new capabilities

1.5x

Explosion of data driving 54% growth in storage every year

54%

54% of all network outages were due to manual error

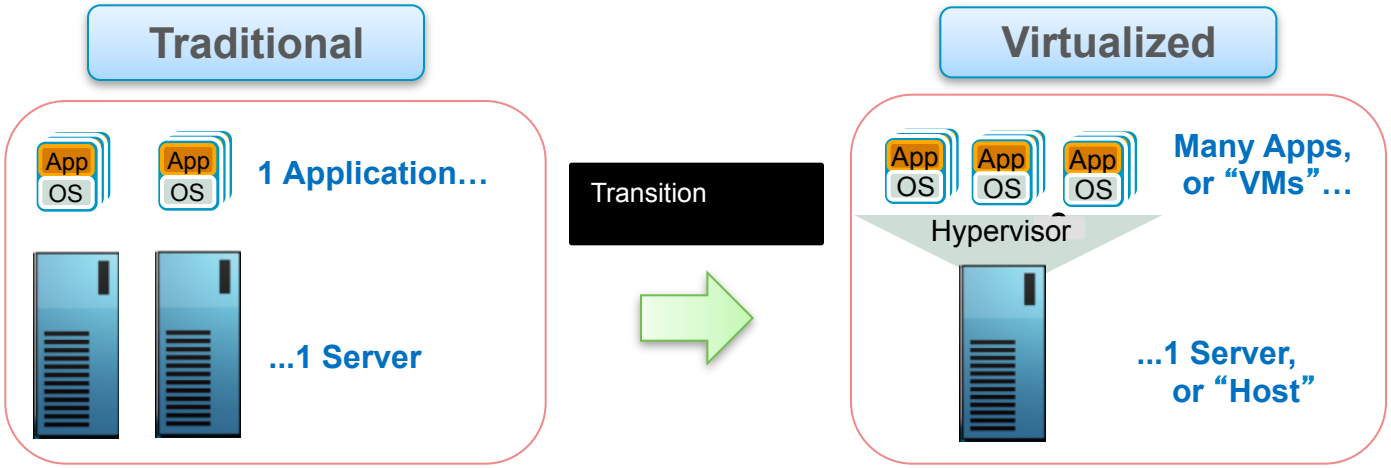
Business Drivers

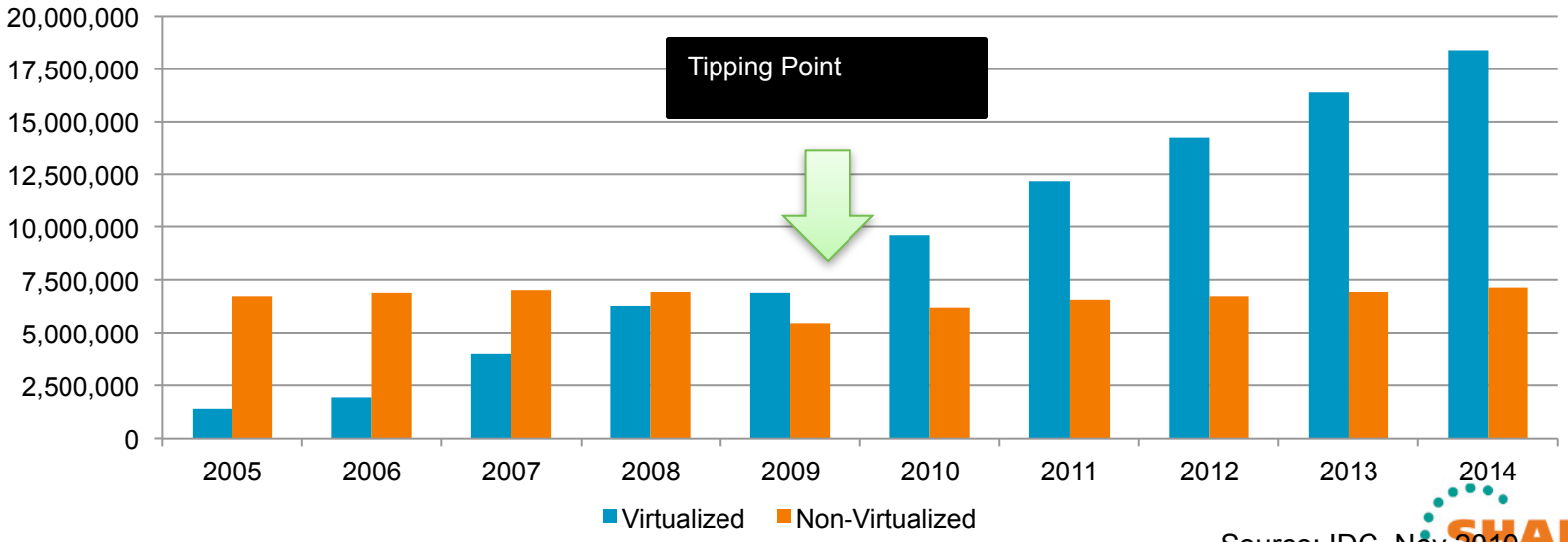
- Value Optimization
  - Address Overall Cost Problems
  - Manage Growth Costs
  - Reduce indirect costs (cabling, cabinets)
- Datacenter Efficiency
  - Limit physical expansion
  - Enable Air-side Economization
  - PUE Targets in Future
- IT as a Service
- Integrated Access Networking
  - Delivered with the Cabinet
  - Eliminate multiple infrastructures in 1 cabinet
  - Reduced Cost and Complexity
- Speed and Agility
  - Reduce time from order to delivery
  - Avoid server implementation surprises
  - Simplify capacity planning

Technical Goals

- Enterprise Workload
  - 10gbps Converged Access Network
  - Increase Compute Density
  - Eliminate extra switches/cabling
  - Reduced Server cost
- Integrate with SAN
  - Access Network combined
  - Keep dedicated SAN Core for Simplicity
- Shrink Infrastructure Footprint
  - Eliminate most patch grids
  - Reduce dedicated network cabinets
- Simplified Deployment
  - Cabinet –level Server Deployment
  - Reduce cabinet cabling
- Facilitate Choice
  - UNIX (AIX) or x64
  - Rack-mount or Blade
  - SAN or no-SAN
  - Dense Deployment (32-48/cab)
  - Sparse Deployment (6-8/cab)
- Legacy Support
  - Selectable Options for Gigabit Ethernet
  - Low density support (6-8 servers/cab)
  - Replace existing infrastructure w/o recabling
- Scale
  - Support scale from tens of devices to thousands
  - Manage costs in smaller deployments

The Evolving Data Centre Architecture
Technology Disruptor - Virtualization





Source: IDC, Nov 2010

What's Driving the Evolution of the Data Center Environment

- Need for better High Availability and lower Fate Sharing
- Need to achieve Higher Scalability
- Need to accommodate diverse workloads concurrently
- Need to further simplify operational models
- Need better Network **visibility**
- Need to prepare for:

Unified Fabric

Why Storage on an Ethernet Fabric?

Ethernet Model has Proven Benefits

Ethernet Economic Model

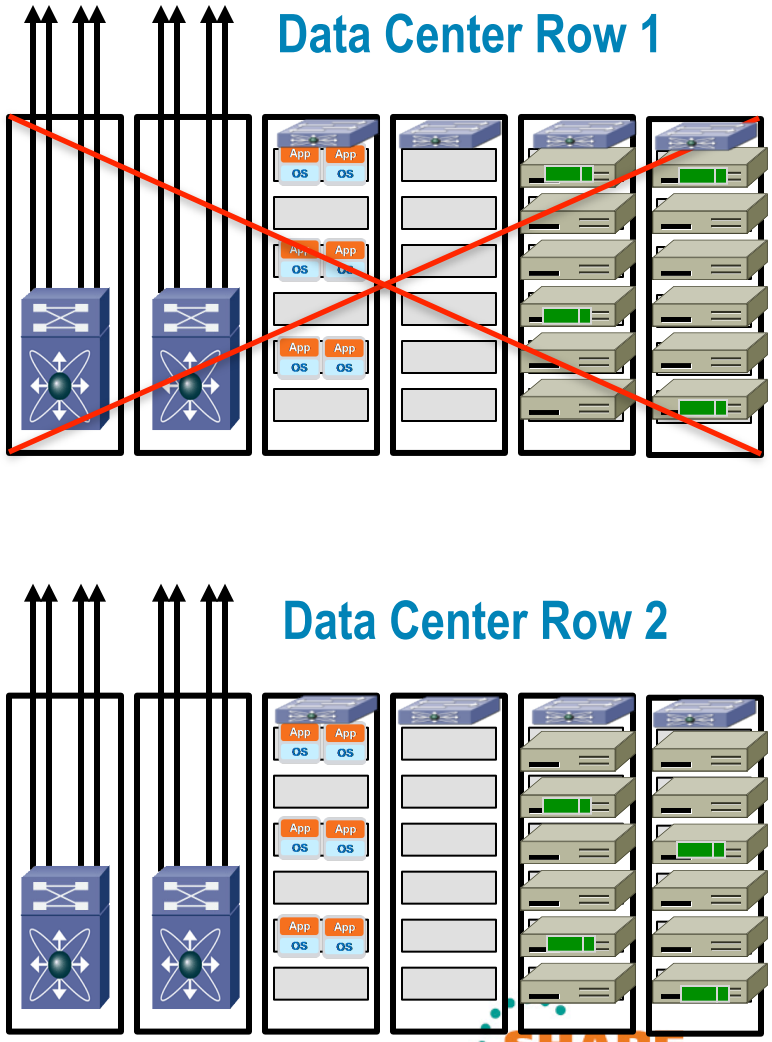
- Embedded on Motherboard
- Integrated into O/S
- Many Suppliers
- Mainstream Technology
- Widely Understood
- Interoperability by Design

FC Economic Model

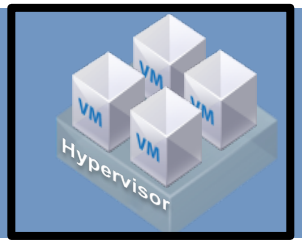
- Always a stand alone Card
- Specialized Drivers
- Few Suppliers
- Specialized Technology
- Special Expertise
- Interoperability by Test

The Evolving Data Centre Architecture
Challenges for the Classical Network Design

- Hypervisor based server virtualization and the associated capabilities (vMotion, …) are changing multiple aspects of the Data Center design
- Where is the server now?
  - Where is the access port?
  - Where does the VLAN exist?
  - Any VLAN Anywhere?
  - How large do we need to scale Layer 2?
- What are the capacity planning requirements for flexible workloads?
- Where are the policy boundaries with flexible workload (Security, QoS, WAN acceleration, …)?



Key Data Center Storage Trends ..

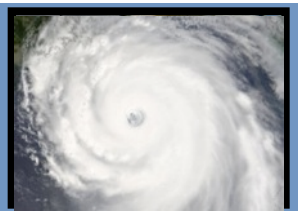


Growth in Server Virtualization

⇨ 86% of Virtual Servers use external storage - ESG, 2008

TCO Spirals Upwards Despite Virtualization

⇨ Density, power, cooling, cabling, management complexity



BC/DR and Compliance Requirements

⇨ Exacerbated by storage growth of 60% / year

Need More Agility and Flexibility

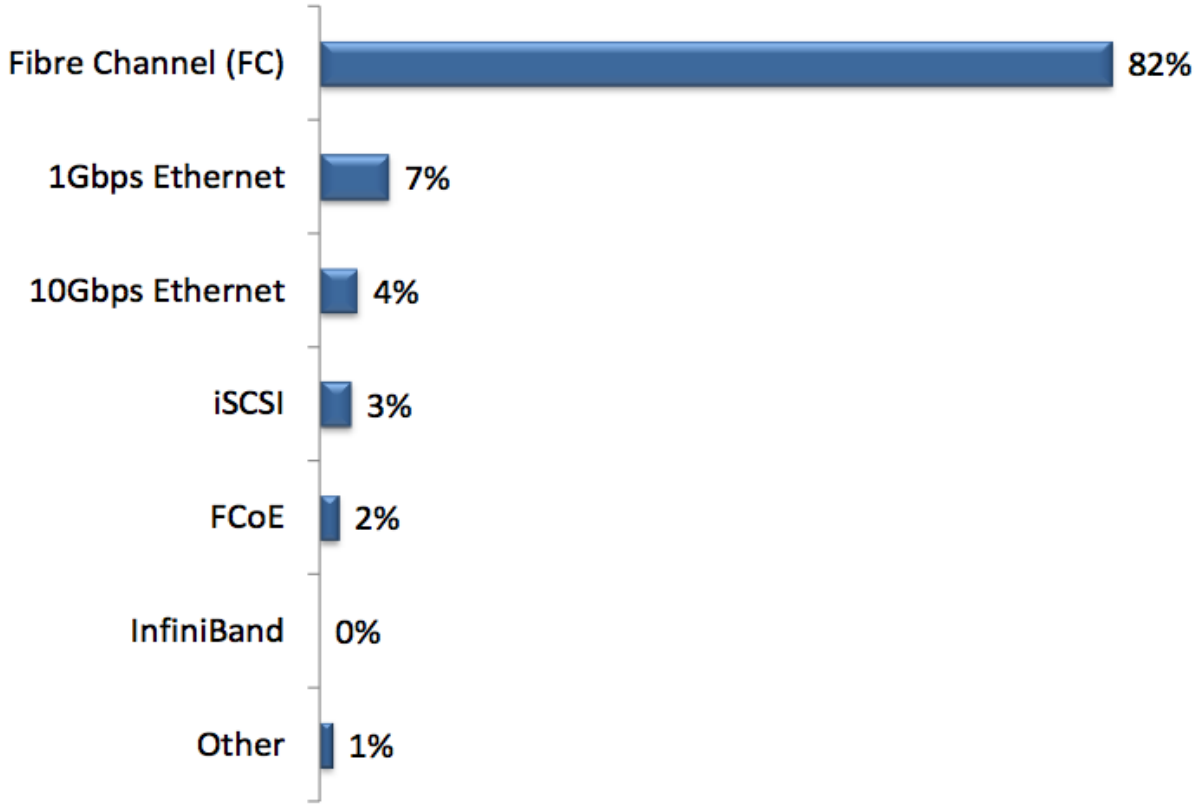
⇨ 2/3 of data center TCO is OpEx

… Legacy SAN Architectures No Longer Suffice

FC Dominates the Backbone Storage Network

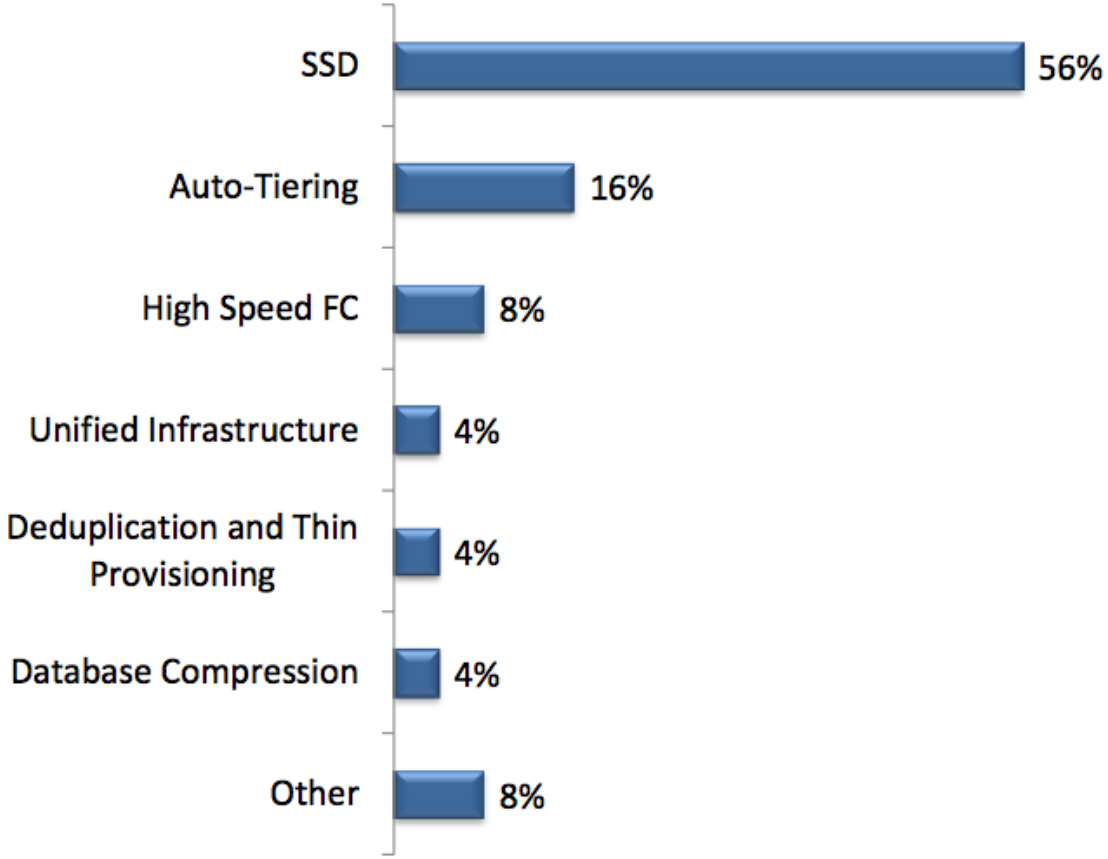
What is the predominant storage network backbone you use?

- By asking for the predominant storage network, we find where the heart of storage networking lies. With 82% selecting FC, the answer is clear.
- From our latest technology roadmap, non-FC storage network technologies are used in greater percentages than appear in this chart. FCoE is *in use* by 8% of respondents, and 10Gbps Ethernet (used for storage) is *in use* by 31%.



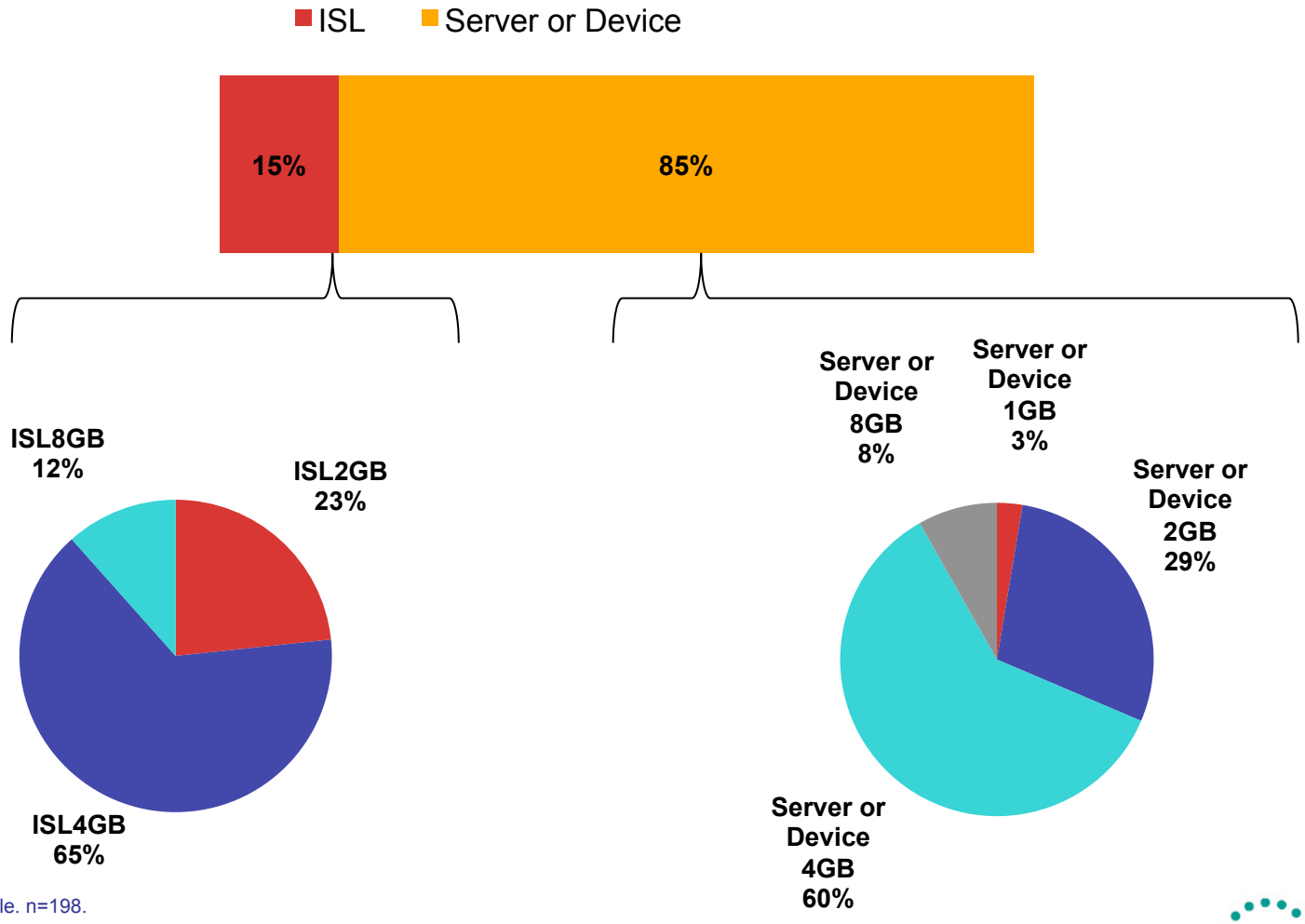
High-speed FC Not Seen As Solving Application Performance

Which storage technology shows the greatest promise for improving application performance?



Full Sample: FC Switch Ports – Types and Usage
8GB Still Arriving; Bulk on 4GB

Of these total FC switch ports, break out the types and usage as a percentage:

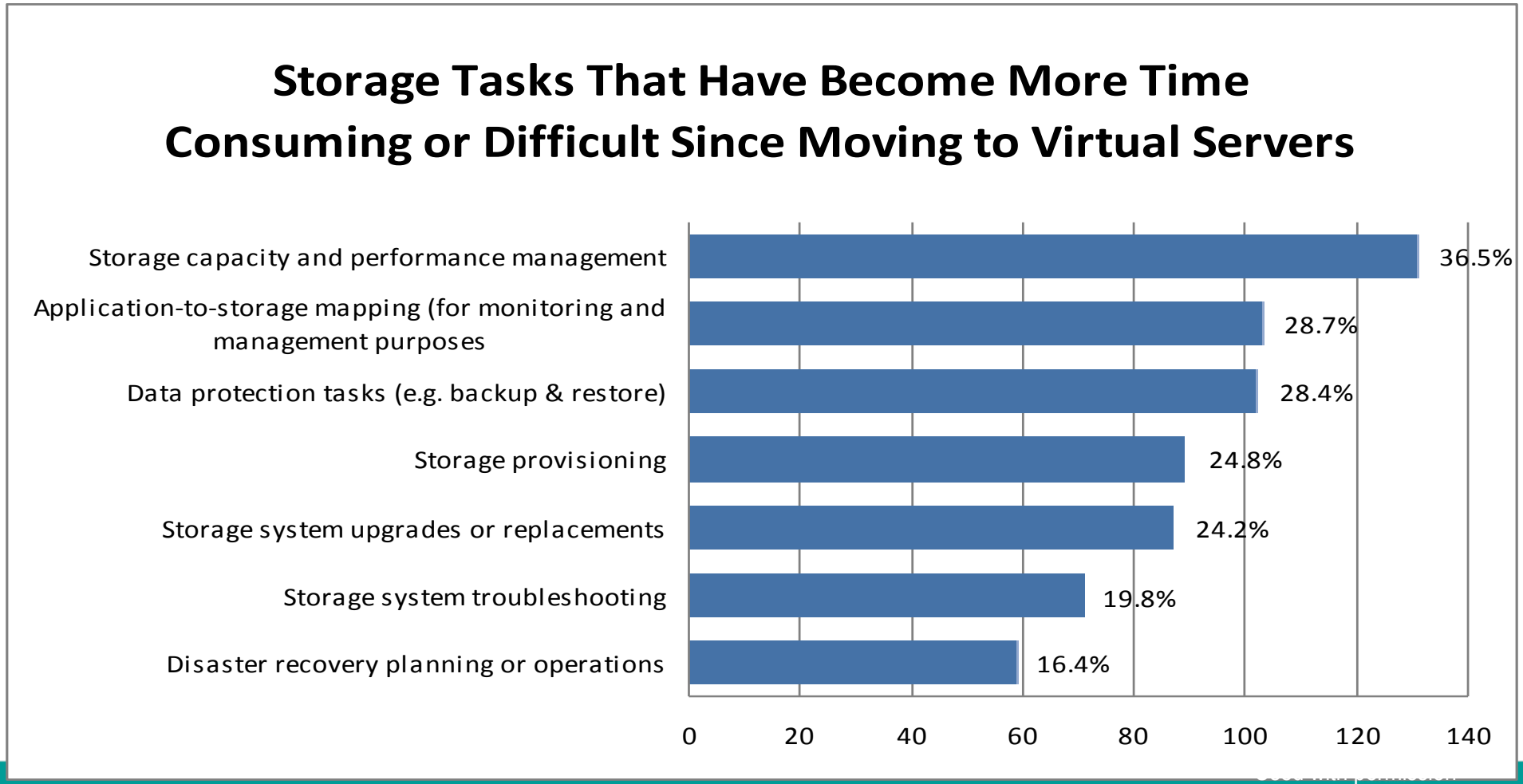


(5/31/11): Full Sample. n=198.

Source: The InfoPro, Q2 CY 2011

Storage Tasks That Are More Time Consuming or Difficult Because of Server Virtualization

- Since moving to virtual servers, 1 or more storage tasks have become more time consuming or difficult for nearly 75% of respondents
- User management (permissions, access control), storage security mgmt, and daily maint/admin tasks were each listed by approx. 10% of users

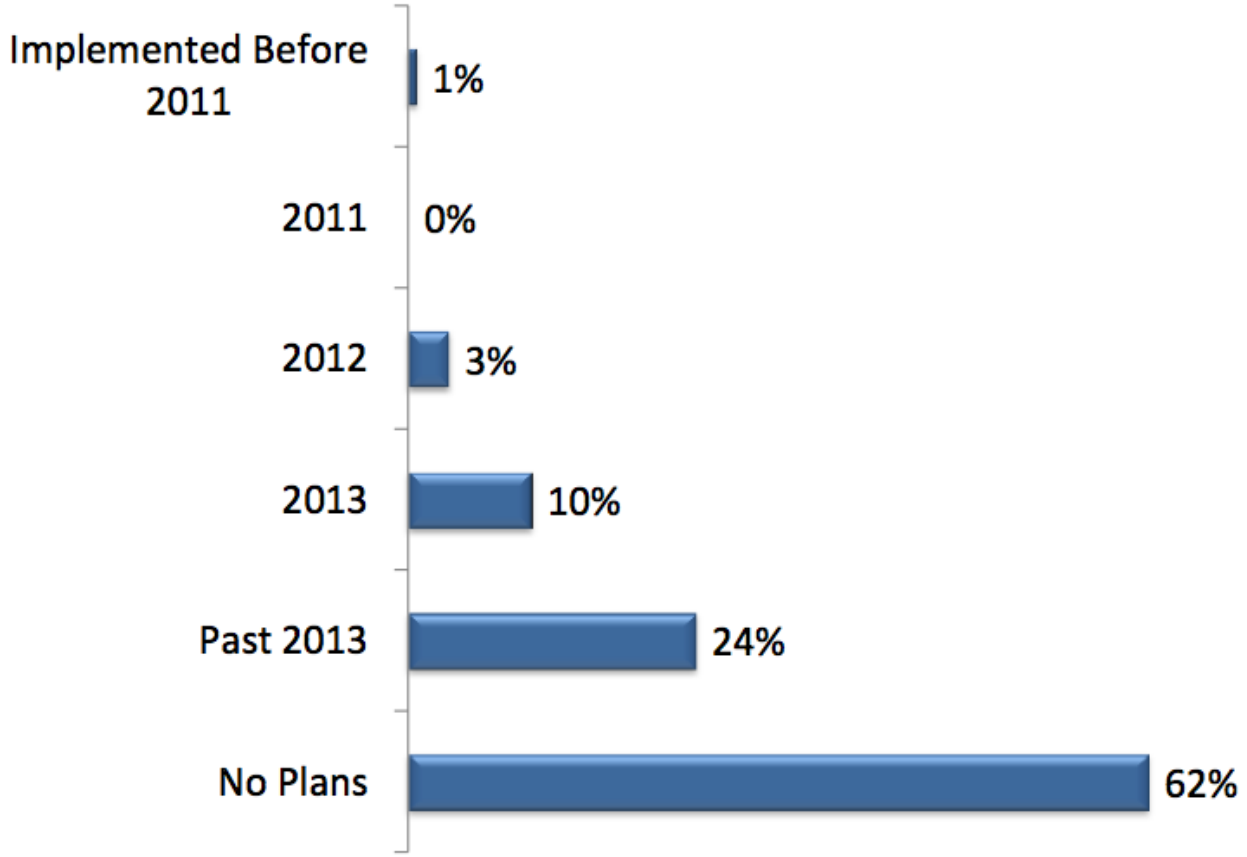


16Gbps Switch Implementation Is a Few Years Off

In which year do you expect to implement 16Gbps at the switch?

Narratives show that the challenges facing 16Gbps adoption include high cost, absence of compelling need, and abundance of non-FC choices:

- "When 16Gbps gets to same price as 8Gbps, perhaps. Do not need the bandwidth right now." – *storage pro at a large industrial/manufacturing enterprise*
- "[16Gbps is] too expensive right now, but we are early adopters when the price is right." – *storage pro at a large telecom/technology enterprise*
- "16Gbps depends upon when the vendors stop selling 8Gbps. We have no real need." – *storage pro at a large healthcare/pharmaceuticals enterprise*
- "We are moving away from FC and would not deploy 16Gbps." – *storage pro at a large energy/utilities enterprise*
- "When 16Gbps becomes prevalent, we will already be using FCoE." – *storage pro at a large education organization*
- "If we go to iSCSI, we won't implement 16Gbps." – *storage pro at a large consumer goods/retail enterprise*



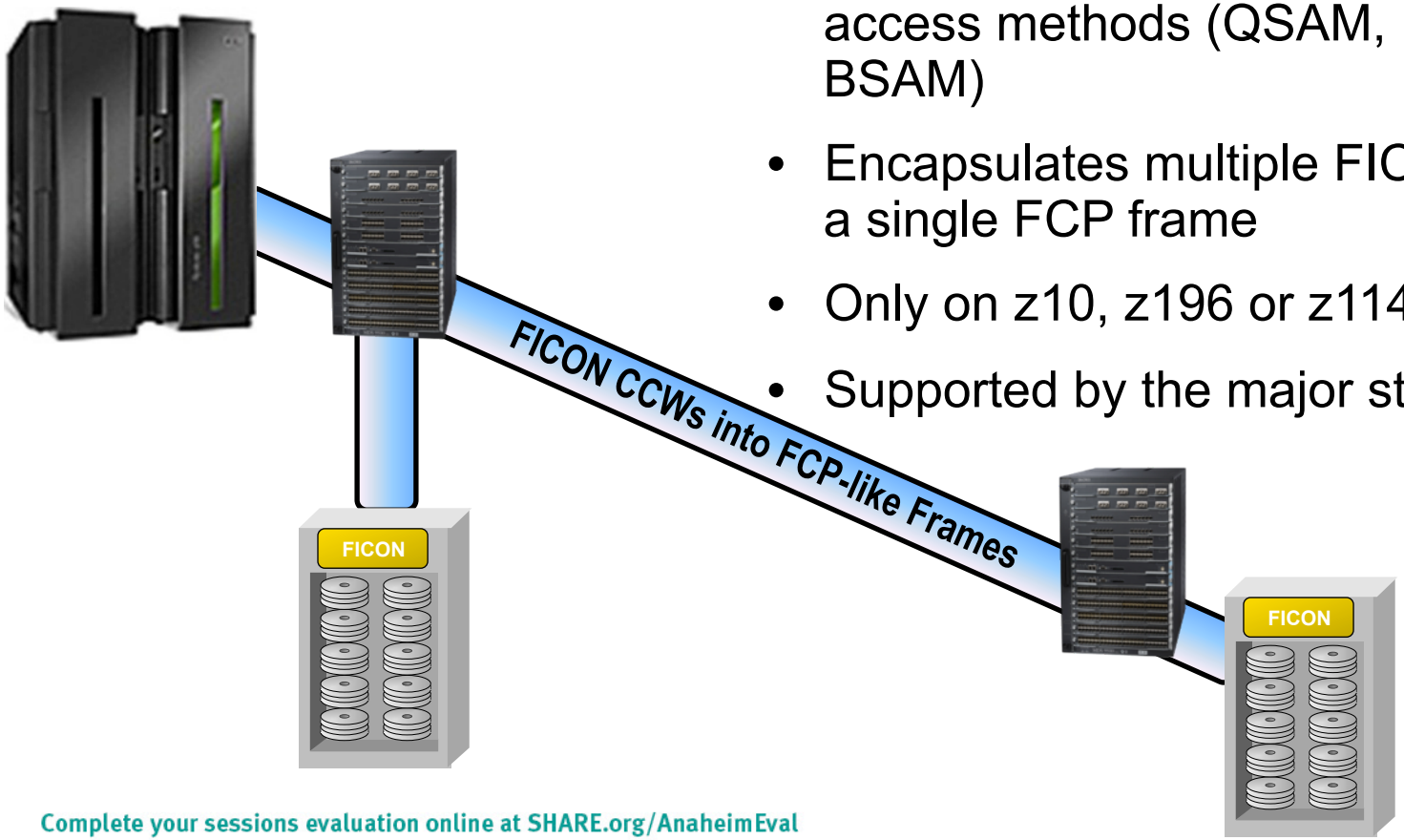
What about zSeries environments?

- High Performance FICON (zHPF) has allowed for increases in bandwidth to meet demands of current business applications as well as being able to support existing infrastructures
- FICON Express 8S offers increases to zHPF (8S only available on z196 and z114)
- Operates at 8G while also being able to auto-negotiate to 2 and 4 Gb/s for older connections
- Available on z10, z196 and z114
- 8S Increases performance 77% in IOPs and 108% in MB/sec

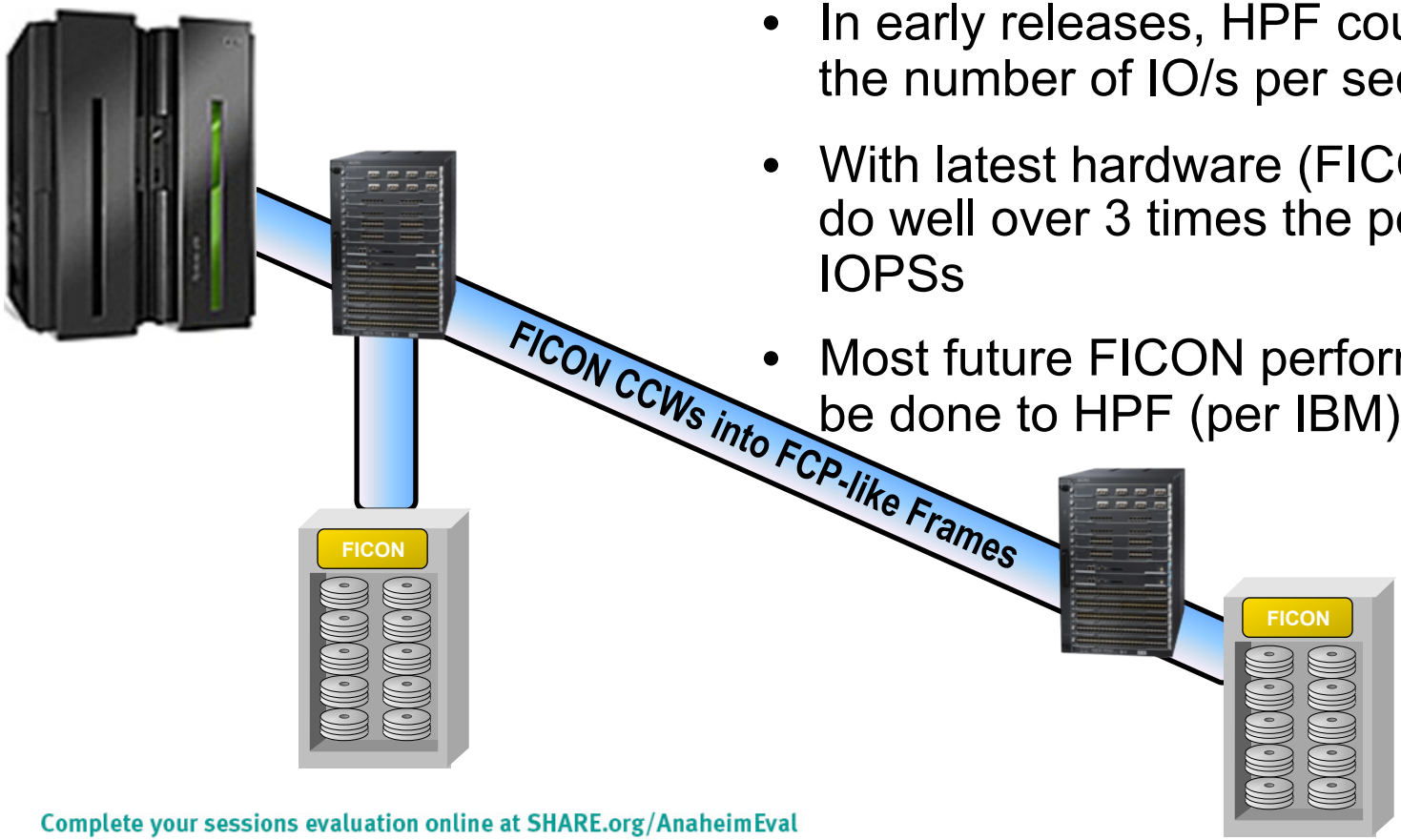
zHPF – High Performance FICON

First introduced in 2008

- Originally used for extended format data sets managed by Media Manager or by EXCP (DASD only)
- Recently enhanced to support further access methods (QSAM, BPAM, and BSAM)
- Encapsulates multiple FICON CCWs into a single FCP frame
- Only on z10, z196 or z114
- Supported by the major storage vendors

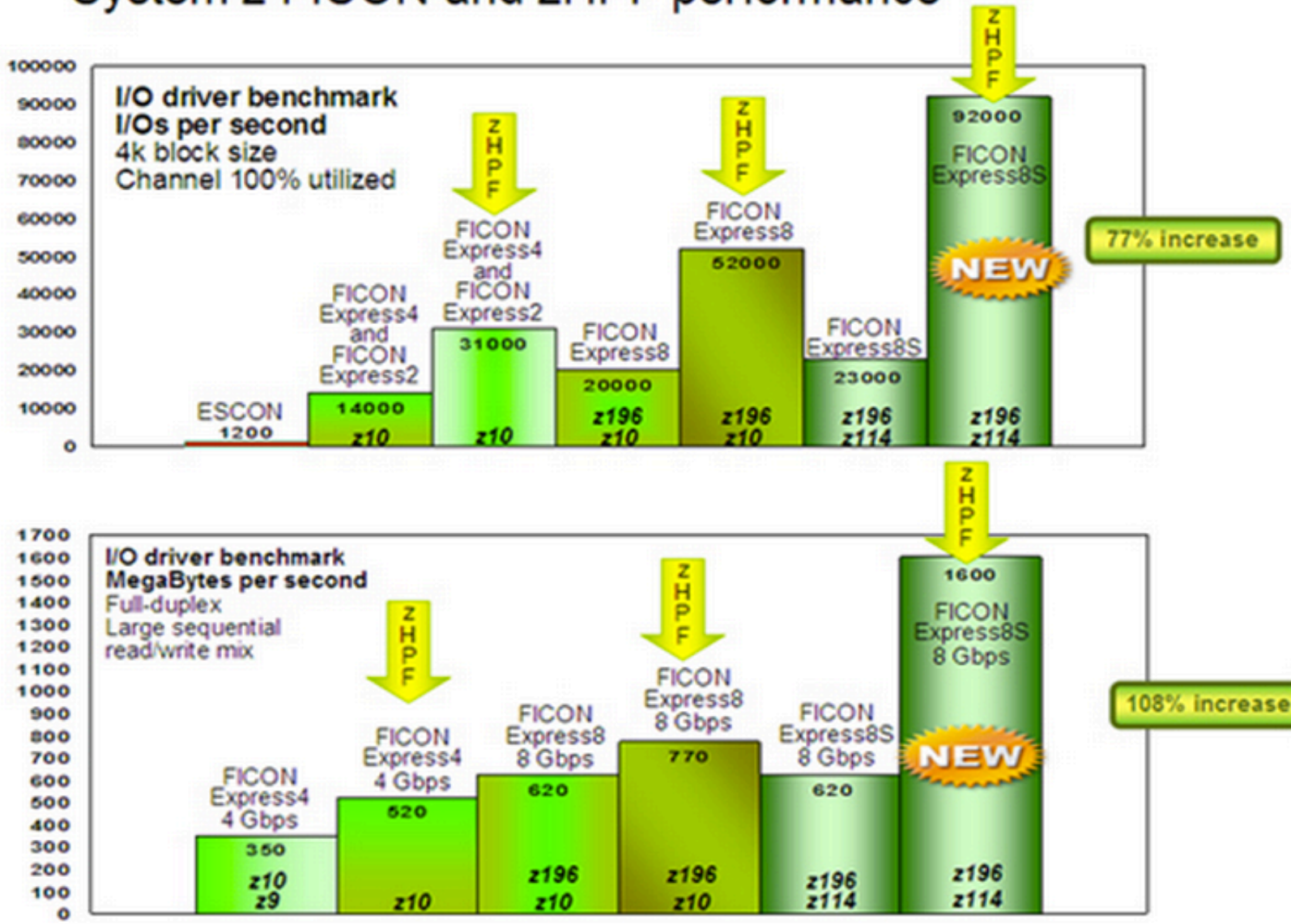


zHPF – High Performance FICON



- Real purpose is to gain the performance boost of open systems but maintain FICON reliability
- In early releases, HPF could do 2 times the number of IO/s per sec (IOPS)
- With latest hardware (FICON 8S), it can do well over 3 times the performance in IOPSs
- Most future FICON performance work will be done to HPF (per IBM)

System z FICON and zHPF performance



Unified Fabric

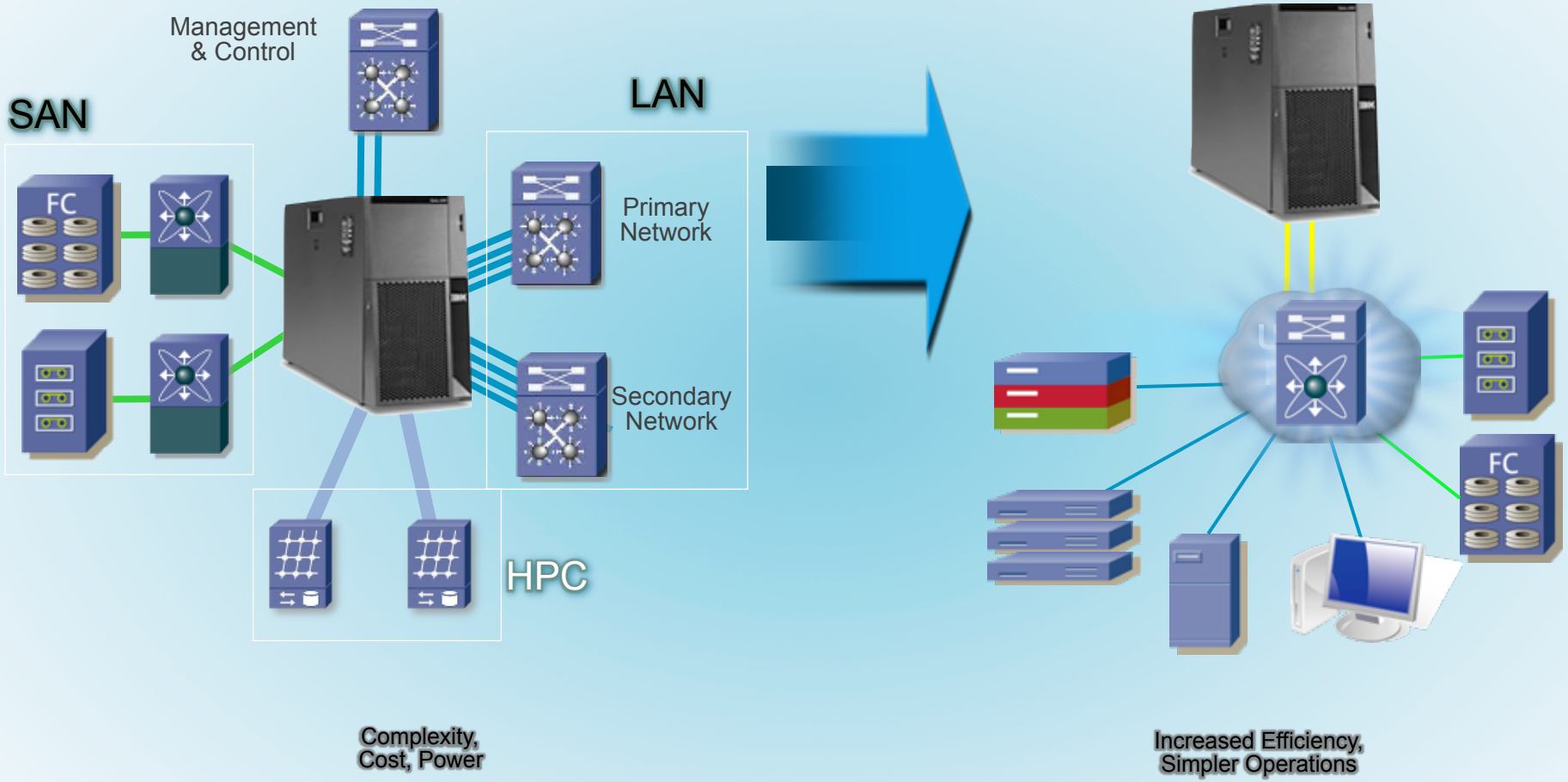
What is it? Really?

Ethernet
DCB
Fibre Channel
FCoE only

Core
Fabric 'A'
Fabric 'B'
FCoE
NAS
iSCSI
FC
L3
L2
Aggregation
Core
Shared Access
Access
Edge
CNA
NX-OS

Consolidating the Data Center Fabric

Many networks, One Infrastructure

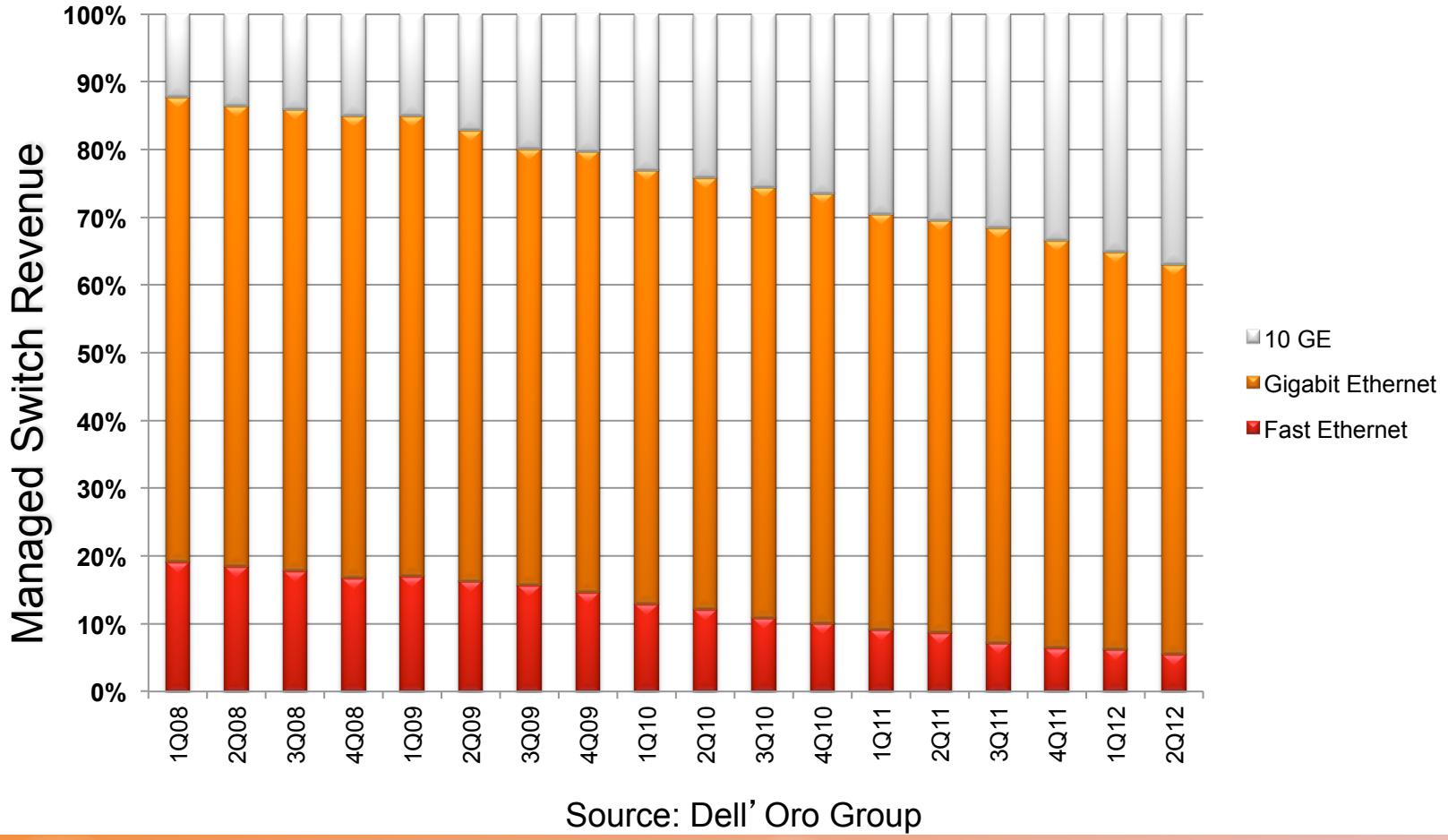


10 Gigabit Ethernet to the Server
Impacting DC access layer cabling architecture

- Multicore CPU architectures
- Virtual Machines driving Increased I/O bandwidth per server
  - increased business agility
- Increased network bandwidth demands
- Consolidation of networks
  - Segmentation & Unified Fabrics / UIO

Future Proofing - Network, Cable Plant

Tremendous opportunity with 10GE
Fastest growing switching technology



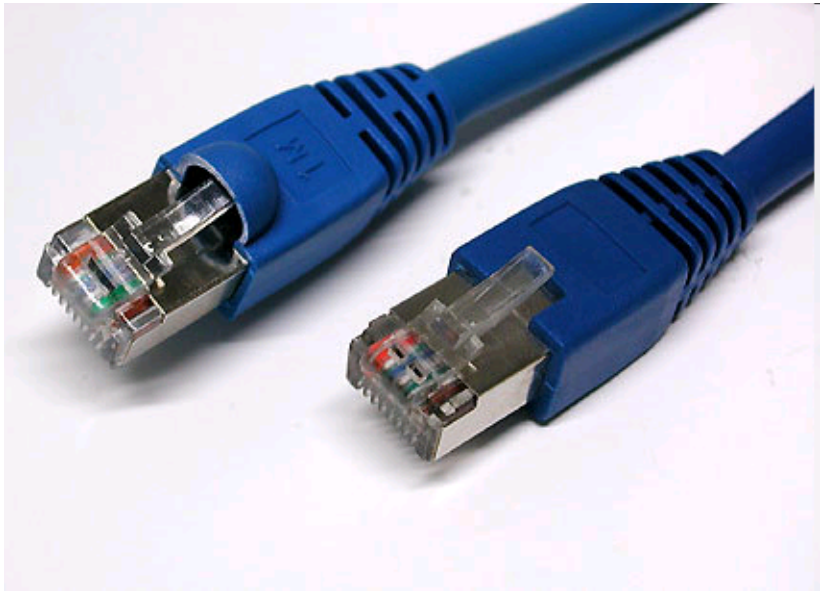
Source: Dell'Oro Group

10GE set to become ubiquitous!

What is 10GBASE-T ?

- A standard released in 2006 to provide 10 Gbps connections over unshielded or shielded twisted-pair cables over distances of up to 330 feet (100 meters)

- A Key objective to provide cost-effective and highly scalable 10 Gigabit Ethernet implementation over structured copper cabling infrastructure



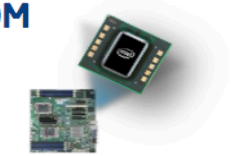
10G on LOM – a game changer

10GbE Deployment Costs

SFP+ Top of Rack

Fiber	Direct Attach
$243 per port	$243 per port
	
$778	$135
	
$256 per port	$256 per port
	
$1277 per port TOR	$634 per port TOR

10GBASE-T Top of Rack

10GBASE-T NIC	10GBASE-T LOM
$305 per port	$305 per port
	
$5	$5
	
$256 per port	LOM
	
$566 per port TOR	$310 per port TOR

Source: List price as of Feb 2011 from the following sources
1 Cisco Twinaxial cable SFP+ 10ft – Newegg.com
2 Cisco 2232TM - CDW.com
3 Cisco 2232TF (10GBASE-T switch) – Technoweb
4 Intel X520 and X540 NICs – Newegg.com
5 Multimode cable, Cisco SFP+ transceiver, Fiber cable – CDW.com

Challenges with 10GBASE-T

- Bit Error Rate (BER) characteristics
- Power
- Latency

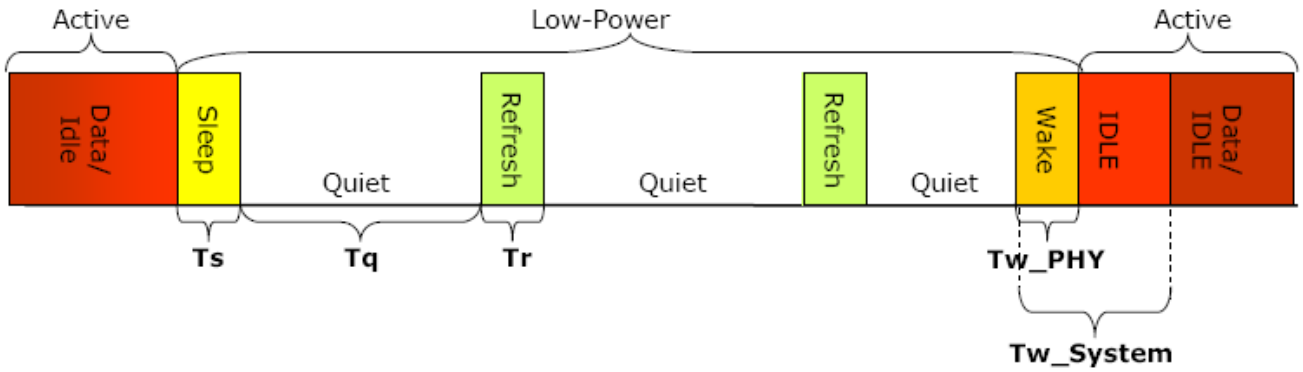
Technology	BER	Distance	PHY Power (each side)	Transceiver Latency
SFP+ CU Copper	~10^{-18}	10m	~0.1W	~0.25µs
SFP+ SR short reach	~10^{-18}	300m	1W	~0.1µs
SFP+ LR long reach	~10^{-18}	10km	1W	~0.1µs
10GBASE-T – 40nm	Est 10^{-15}	100m	2W-3W	~3µs ~3µs

The concept of Low Power Idle

Energy Efficient Ethernet

- Concept: transmit data as fast as possible, return to Low-Power Idle
- Saves energy by cycling between Active and Low Power Idle
- Power reduced by turning off unused circuits during LPI
- Energy use scales with bandwidth utilization

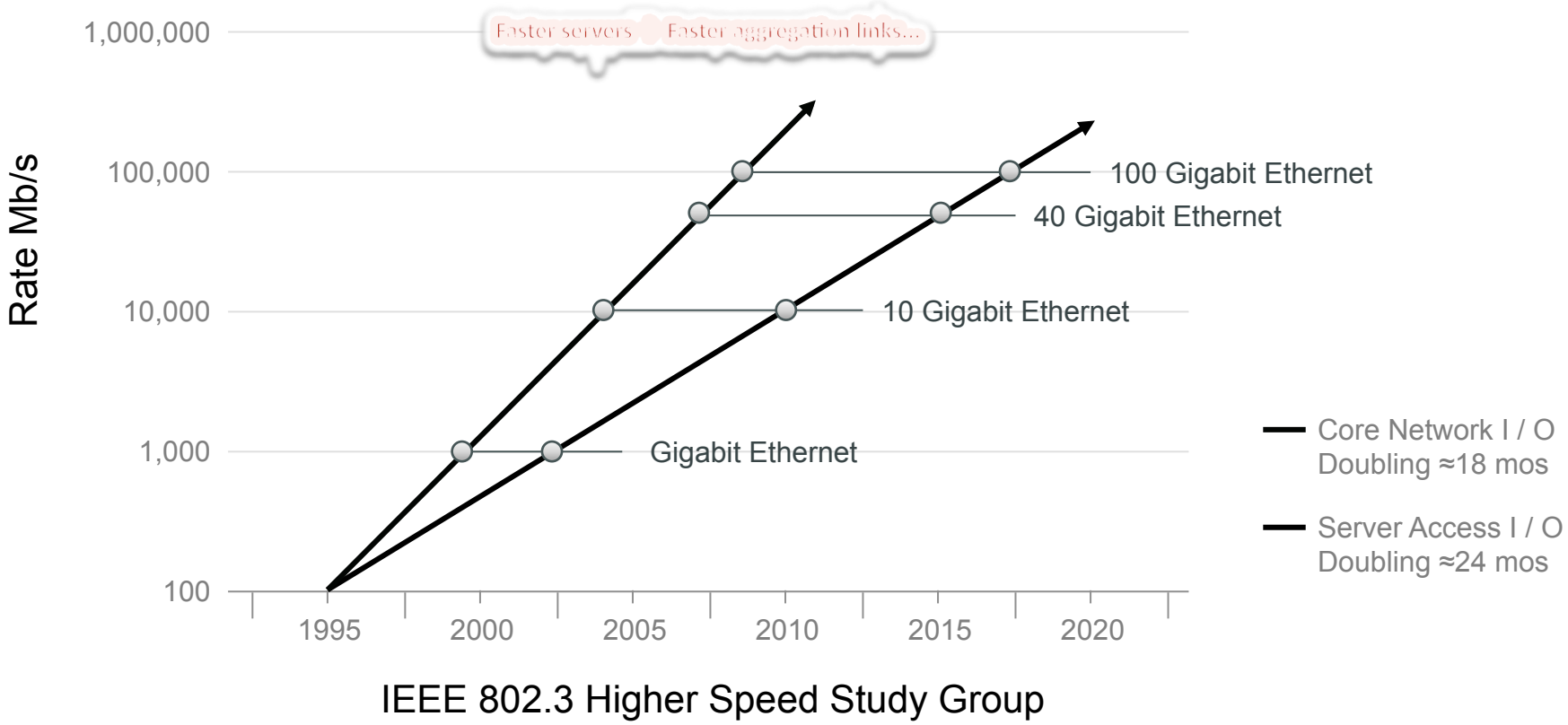
Term	Description
Sleep Time (Ts)	Duration PHY sends Sleep symbols before going Quiet.
Quiet Duration (Tq)	Duration PHY remains Quiet before it must wake for Refresh period.
Refresh Duration (Tr)	Duration PHY sends Refresh symbols for timing recovery and coefficient synchronization.
PHY Wake Time (Tw_PHY)	Duration PHY takes to resume to Active state after decision to Wake.
System Wake Time (Tw_System)	Wait period where no data is transmitted to give the receiving system time to wake up.



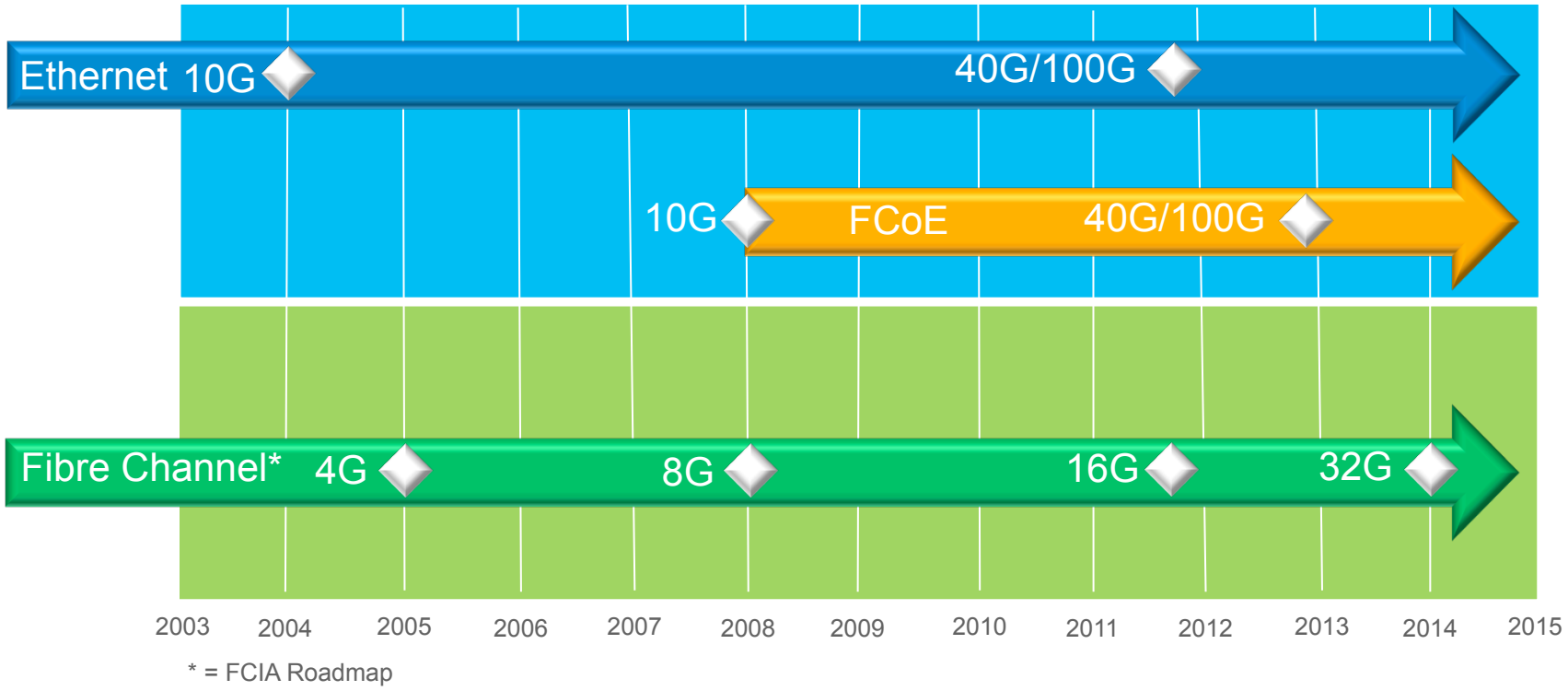
FCoE on 10GBase-T

- BER characteristics improving with newer generations of PHYs
  - 40nm PHYs (2012) seem better than 65nm PHYs (2011)
- FCoE support need ~10^{-15} – No single standard
- Working with the ecosystem to define requirement and test
  - Adapter vendors: QLogic, Emulex, Intel, Broadcom
  - Storage vendors: EMC, NetApp
- BER testing underway for following - no FCoE support at FCS
  - Nexus 2232TM-E
  - Nexus 5596T
  - Nexus 7000 F2-Series Copper

High-Speed Ethernet Trends (Source IEEE)



Protocol Roadmaps



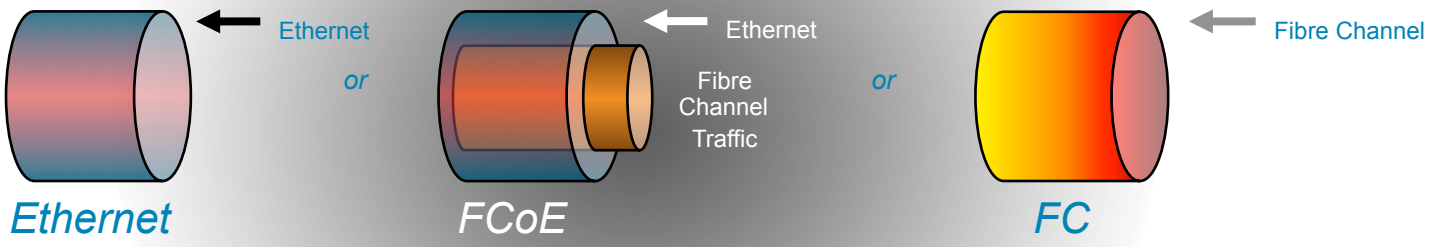
Ethernet is set to surpass Fibre Channel on throughput

Unified Ports

Unified Port

Ultimate Flexibility for Server Access Connectivity

- One port for all types of server IO
- Flexibility of use enables one standard chassis for all data center I/O needs



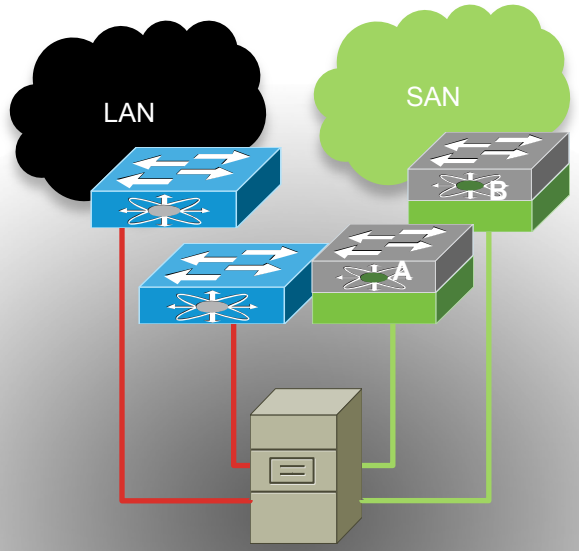
Features

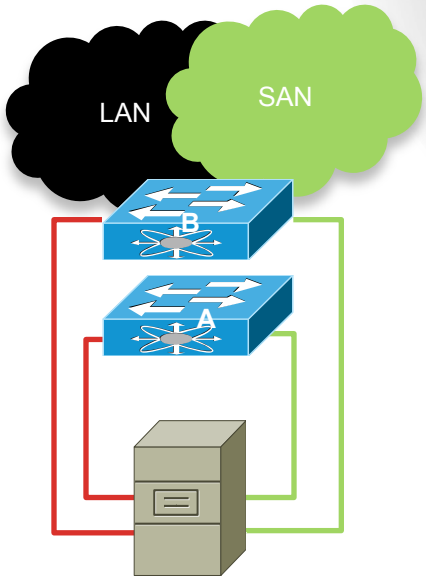
- Any Nexus 5500UP port can be configured as 1/10GE, DCB (lossless Ethernet), FCoE on 10GE (dedicated or converged link) or 8/4/2/1G native Fibre Channel

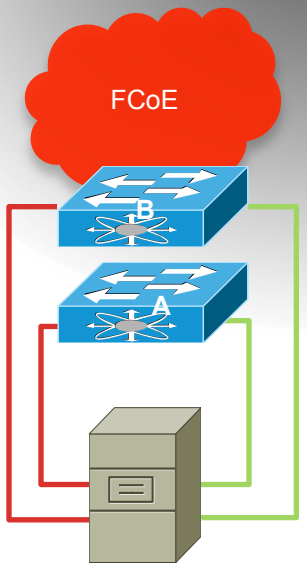
Benefits / Use-cases

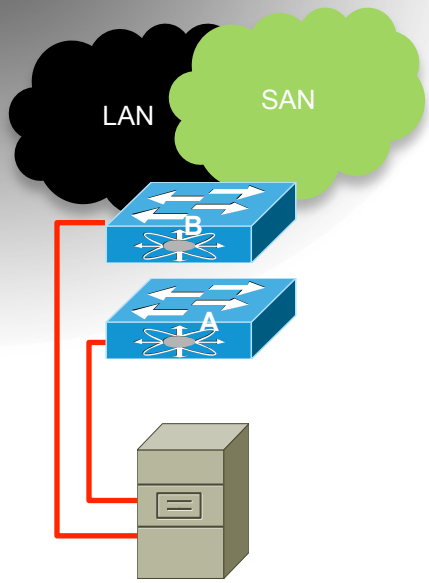
- Deploy Nexus 5500UP as a data center switch standard capable of all important I/O
- Mix FC SAN to host as well as switch and target with FCoE SAN
- Implement with native Fibre Channel today, enables smooth migration to FCoE in the future

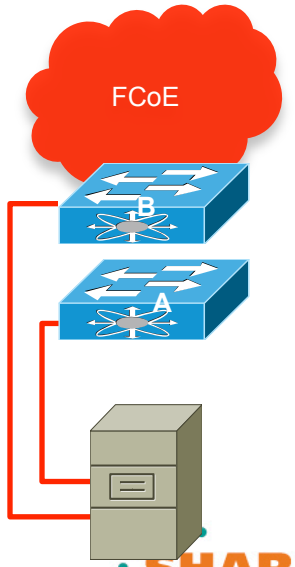
Flexibility of Unified Port











Cisco Storage Networking Portfolio

SHARE
Technology · Connections · Results

Enterprise and Service Provider Solutions

Small/Medium Business Solutions

FCoE

FC

Nexus 4000

Nexus 2232

New

New

Nexus 5548UP

Nexus 5596UP

Nexus 7010, 7018

Ethernet & FCoE

FC & FCoE

MDS 9506, 9509, 9513

MDS 9148

MDS 9222i

IBM FC Blade Switch

MDS 9124

Cisco DCNM

Cisco NX-OS

Thank You!

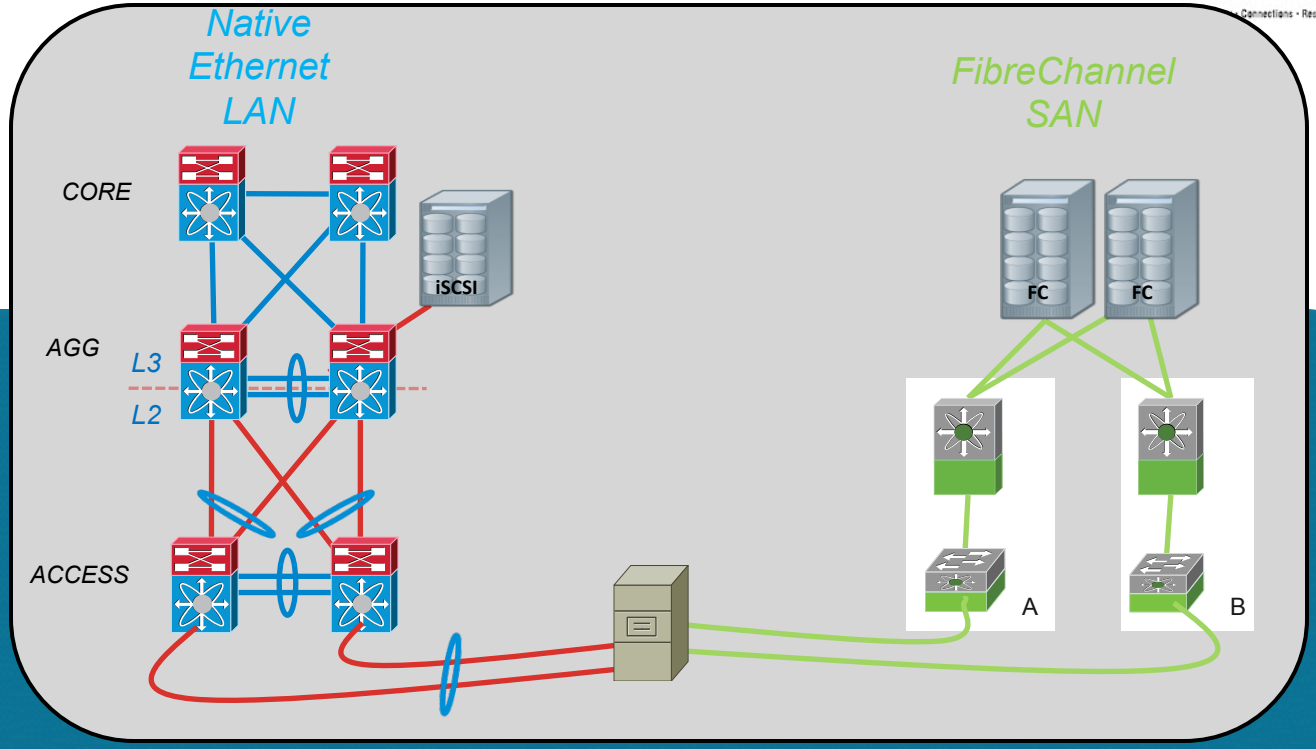
12081


CISCO

Deployment Scenarios

Explore the possibilities in FCoE deployments.

Data Center Network Today



Segregated
Purpose Built
Dually Managed
Limited in scale

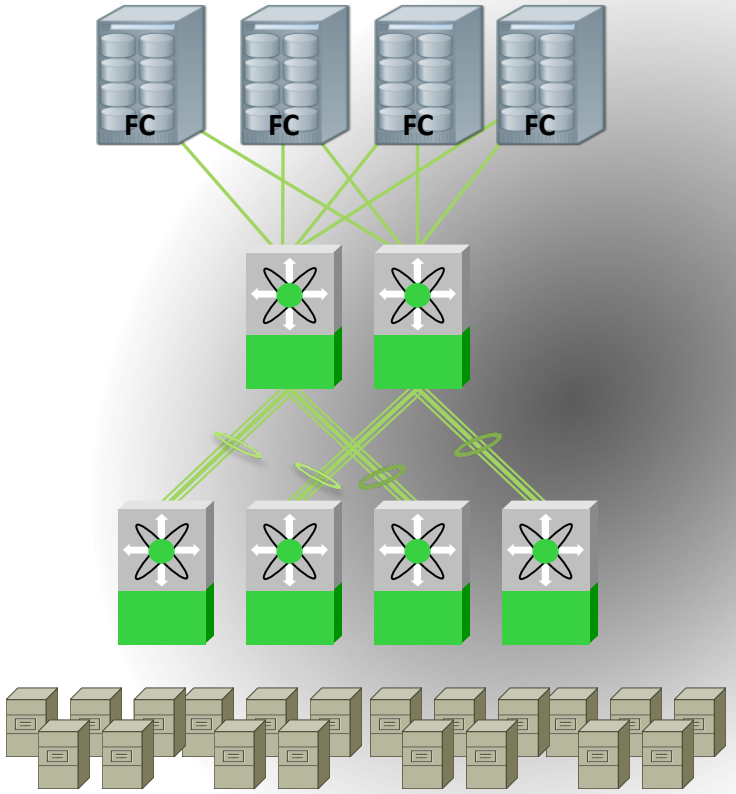
Ethernet is Dominant
Applications are Loss Tolerant
Out of Order Frames Tolerated

FibreChannel is dominant
Loss Intolerant Applications
Strict In Order Delivery
Highly Available, Dual Path

Core-Edge Storage Networks

Core-Edge SAN Design

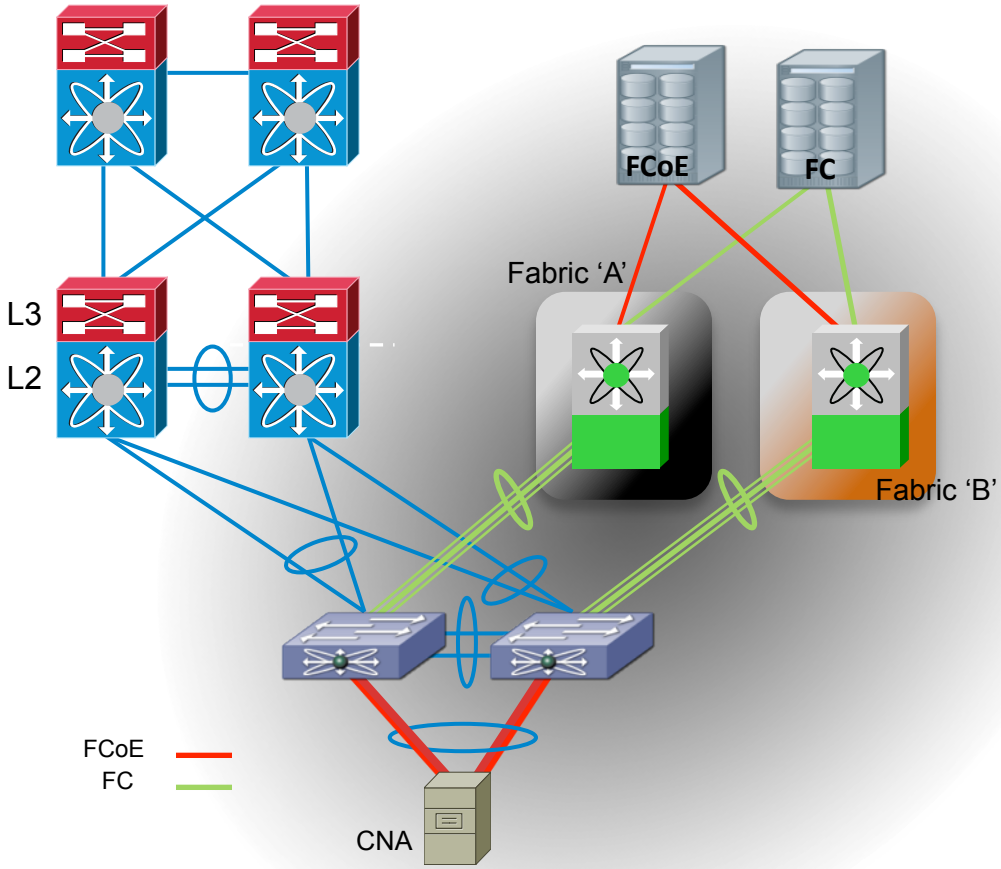
In a Traditional Fibre Channel SAN



- Distinct and Physically separate FC fabric
- Highly Available by Design
- Appropriate for situations where number of ports in the core are sufficient for storage ports available
- Core ISLs are port-channeled for greater aggregate bandwidth and link level availability
- Oversubscribed at the host edge

Core-Edge SAN Design

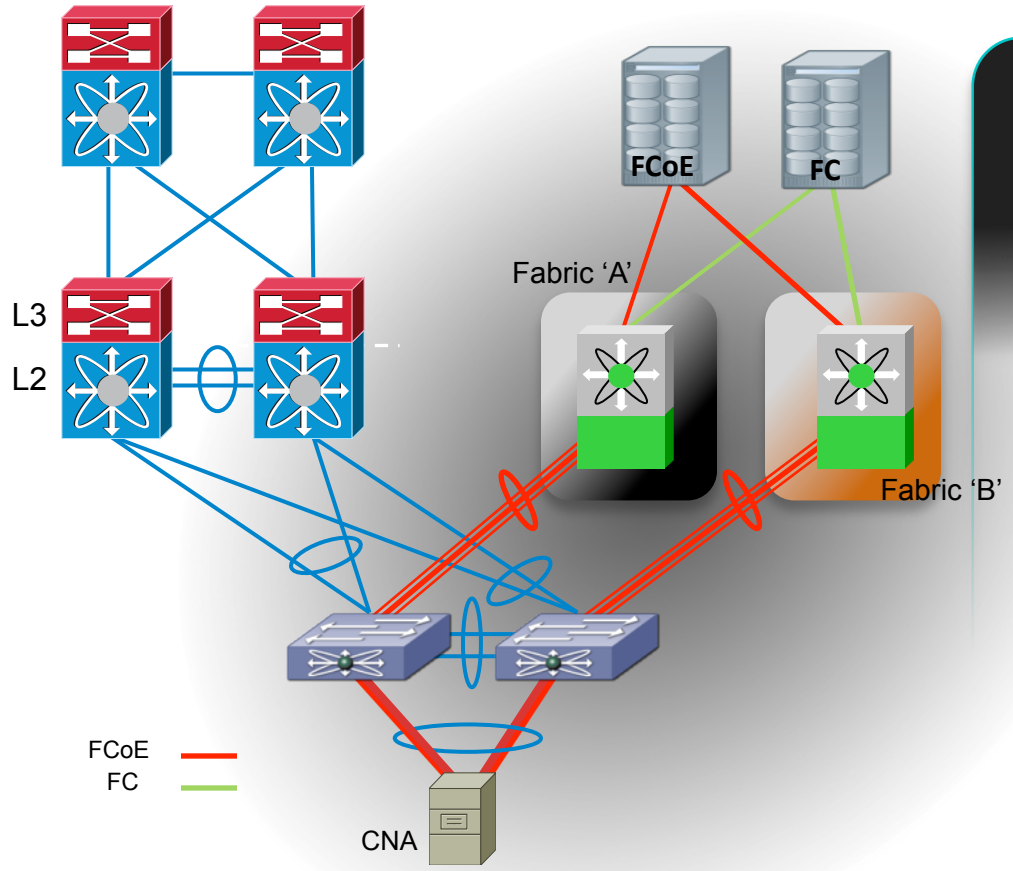
Consolidating SAN/LAN at Access layer



- Benefits of Consolidation at the Access Layer
- Common SAN to access FC and FCoE Storage
- Maintain Storage Fabric Separation
- ***Core-Edge*** Topology
- Facilitate End-to-End FCoE Deployment without a forklift SAN upgrade
- Extend port density by utilizing the Nexus 2000 Fabric Extender

Core-Edge SAN Design

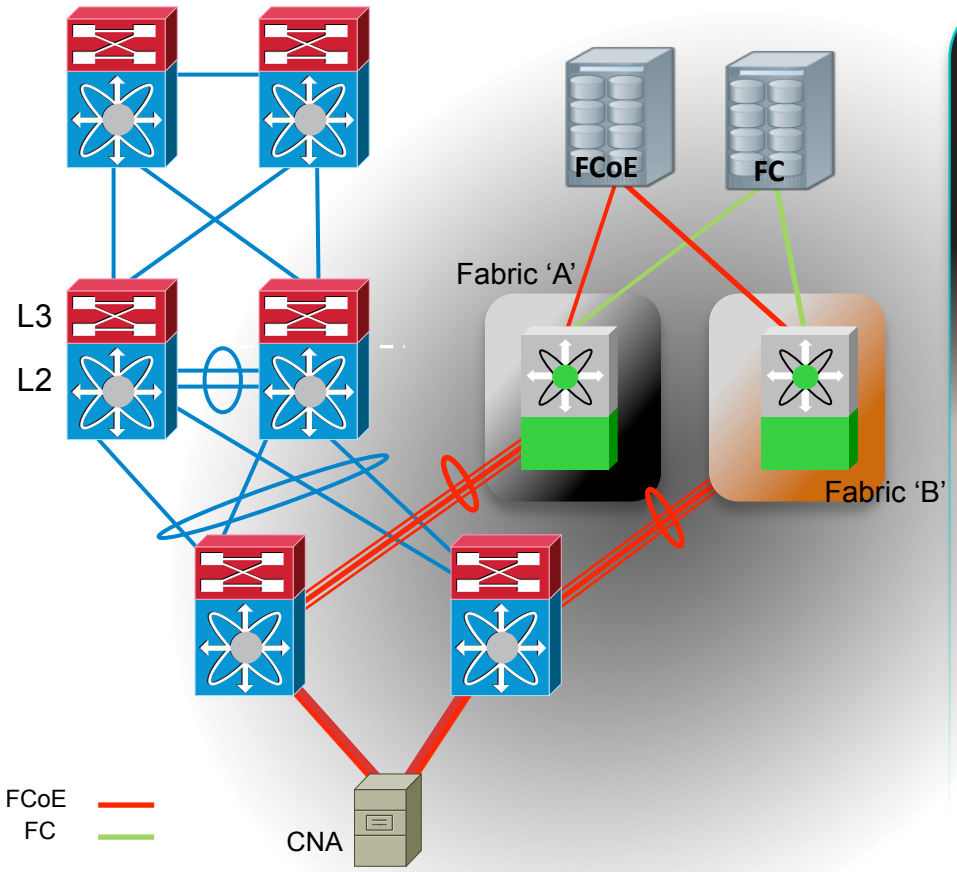
with FCoE Beyond the Access Layer



- Benefits of Consolidation at the Access Layer
- Take advantage of 10GE bandwidth for all ISLs (with a view to upgrade to 40/100GE)
- Maintain Storage Fabric Separation
- ***Edge-Core*** Topology
- Facilitate End-to-End FCoE Deployment without a forklift SAN upgrade
- Extend port density by utilizing the Nexus 2000

Edge-Core SAN Design

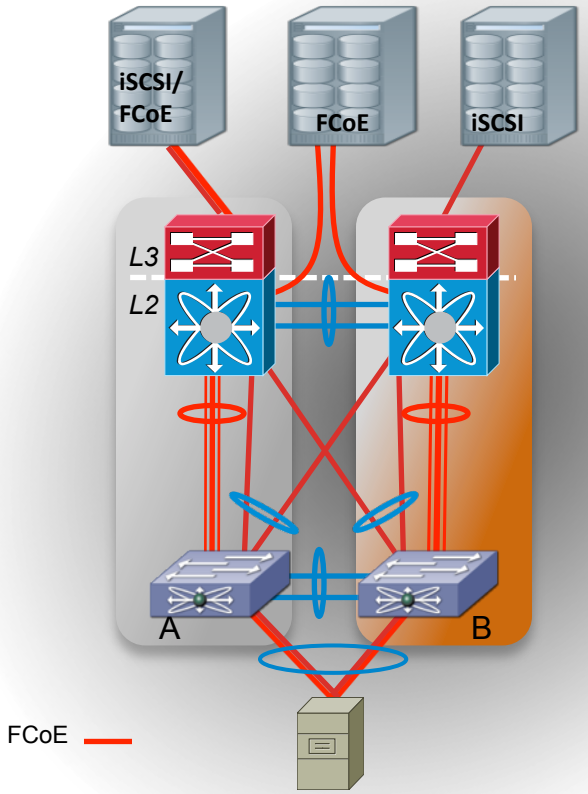
Director-Class at Access layer



- Convergence using a Director-Class platform – for cases where an MDS 9500 was deployed at the FC SAN edge
- Benefits of Consolidation at the Access Layer
- Take advantage of 10GE bandwidth for all ISLs (with a view to upgrade to 40/100GE)
- Maintain Storage Fabric Separation
- ***Edge-Core*** Topology
- **Roadmap:** Nexus 2000 support on Nexus 7000

Core-Edge SAN Design

A Unified Network

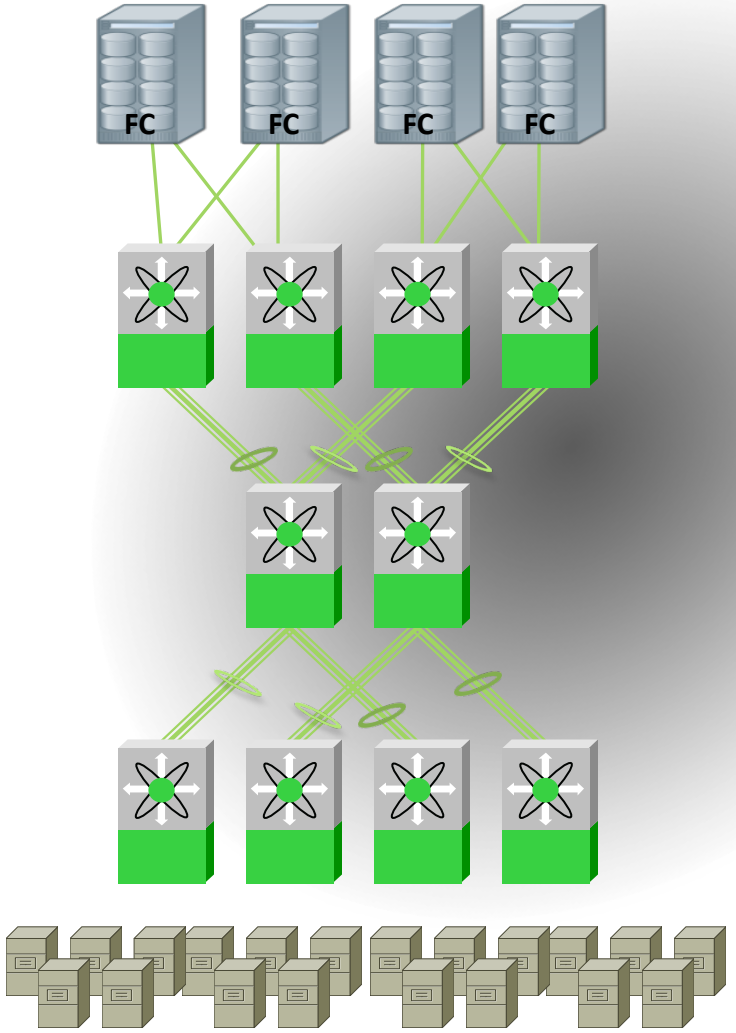


- **Nexus 7000**: Ethernet Storage Director
- Highly Available, Highly Scalable
- Single Network for Multiple Storage Protocols
- Maintain Storage Fabric Segregation
- Benefit from Ethernet's Economic Model
- Single Fabric, lowering points of management

Edge-Core-Edge Storage Networks

Edge-Core-Edge SAN Design

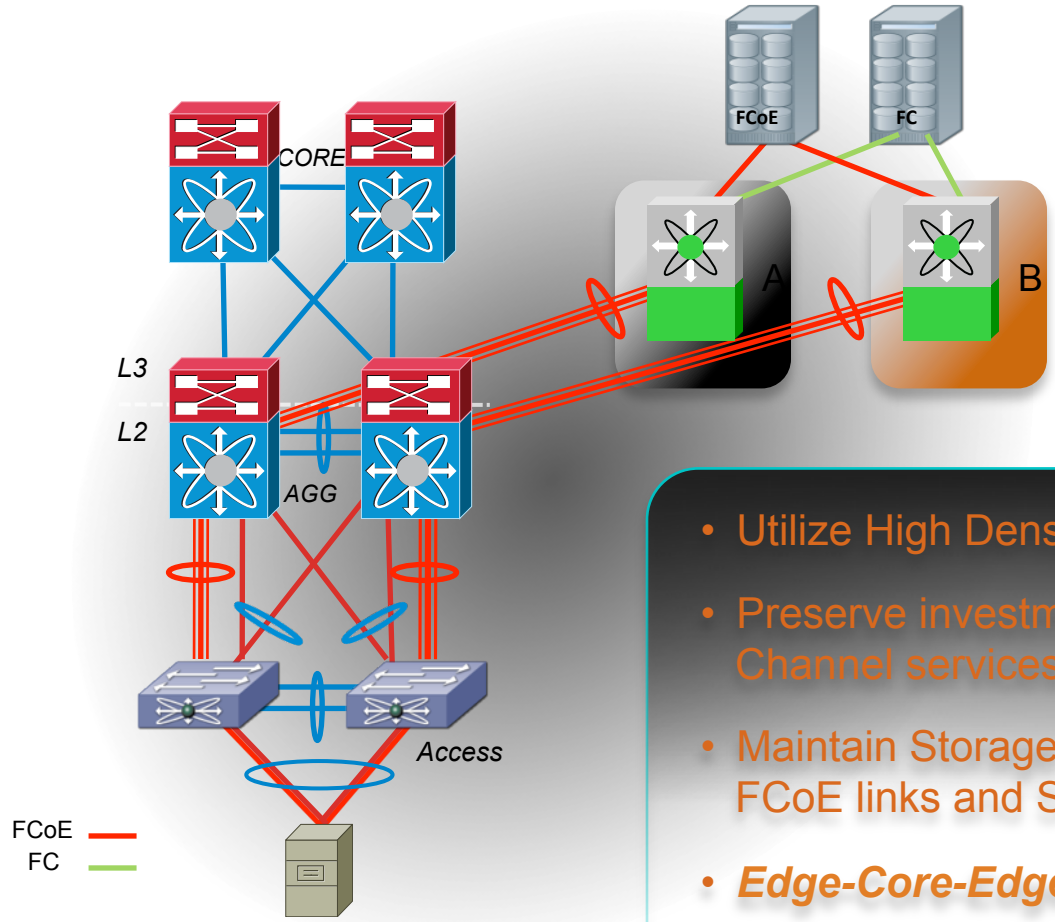
In a Traditional Fibre Channel SAN



- Distinct and Physically separate FC fabric
- Highly Available by Design
- Appropriate for situations where number of ports in the core are not sufficient for the foreseeable future growth of storage ports
- Core ISLs are port-channeled for greater aggregate bandwidth and link level availability
- Oversubscribed at the host edge

Converged Network Design

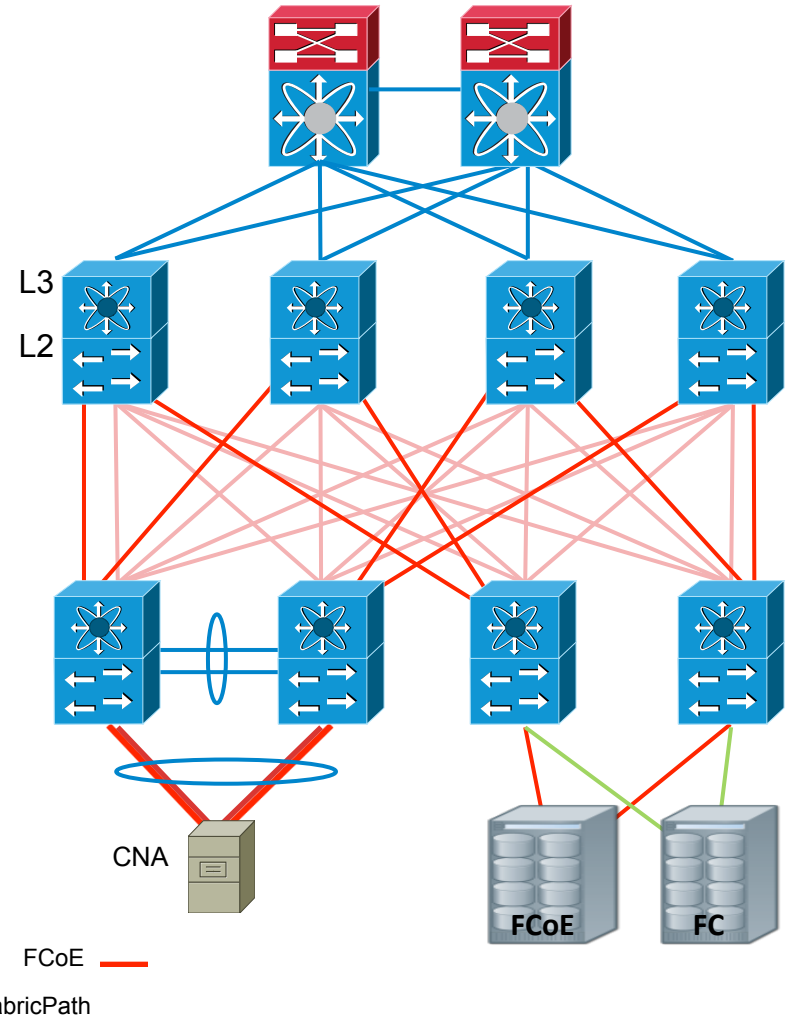
Integrating with Existing FibreChannel SAN



- Utilize High Density switches at Aggregation
- Preserve investments in Fibre Channel and Fibre Channel services
- Maintain Storage Fabric Segregation with Dedicated FCoE links and Storage VDC
- ***Edge-Core-Edge*** Topology
- Facilitate end-to-end FCoE Deployment for larger deployments

Converged Network Design

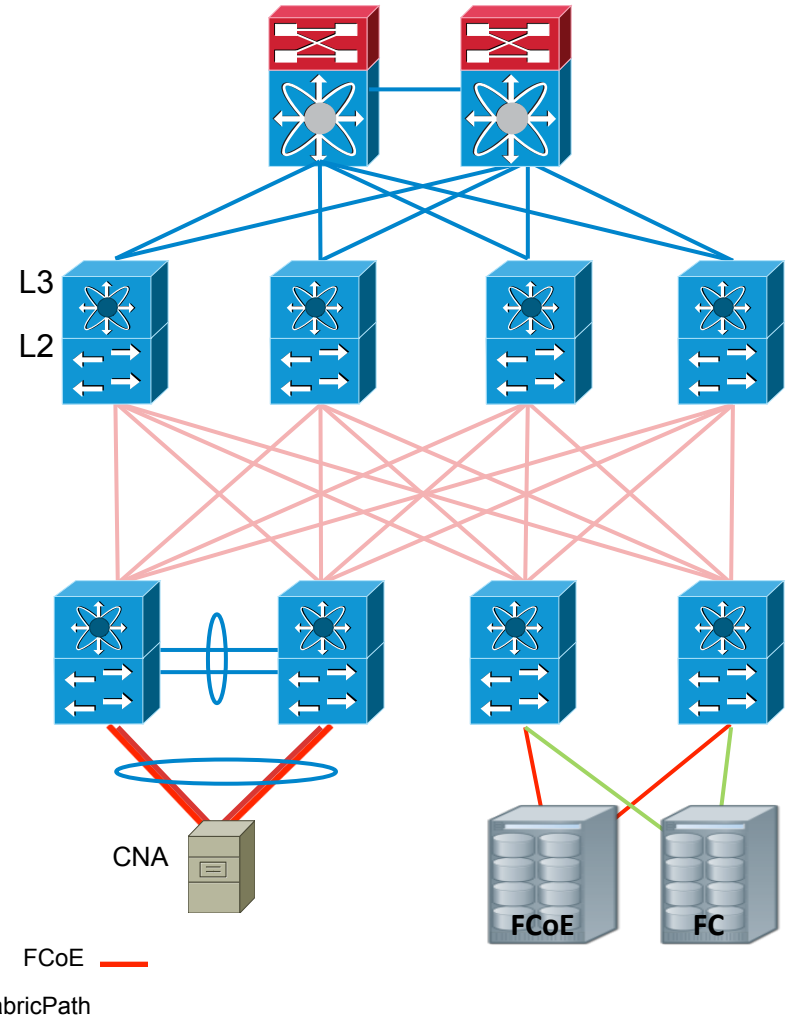
A Unified Network



- **TODAY:** NEXUS platforms can build a Highly Scalable, Multi-Purpose and Resilient Ethernet Fabric
- Converge on to this Fabric, all your LAN and SAN traffic
- Nexus 5000 and MDS used as Storage Edge
- Standardize on I/O, O/S and Platform

Converged Network Design

A Single Fabric



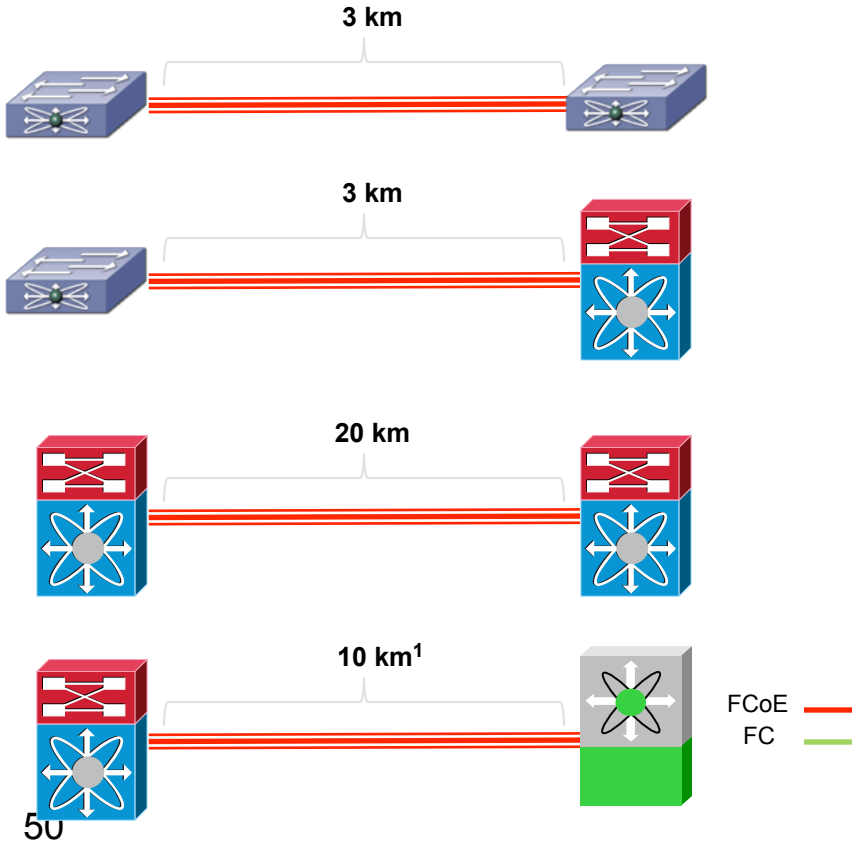
FORWARD LOOKING

- Multi-Purpose and Resilient Ethernet Fabric
- Link Level (ECMP) and Switch level Redundancy
- Converge on to this Fabric, all your LAN and SAN traffic
- Logical SAN A&B separation
- Currently **Not** supported:
  1. Majority of customers still require physical A&B separation
  2. Both Fabric Path and MultiHop FCoE are at their infancy. Perceived Risk is High
  3. As customer appetite for FCoE matures, in due course this model will be supported

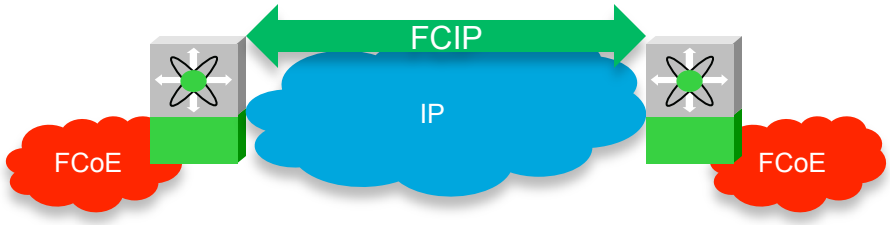
FCoE Extension Options

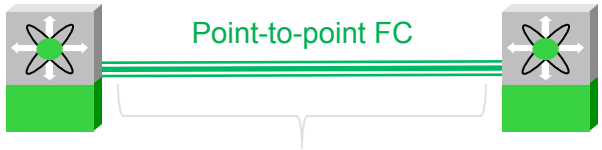
Short Distance Options

- **Requirement:** Maintain loss-less behavior across the point-to-point link
- Supported distance is governed by the ingress buffer size available on the switch



Longer Distance Options





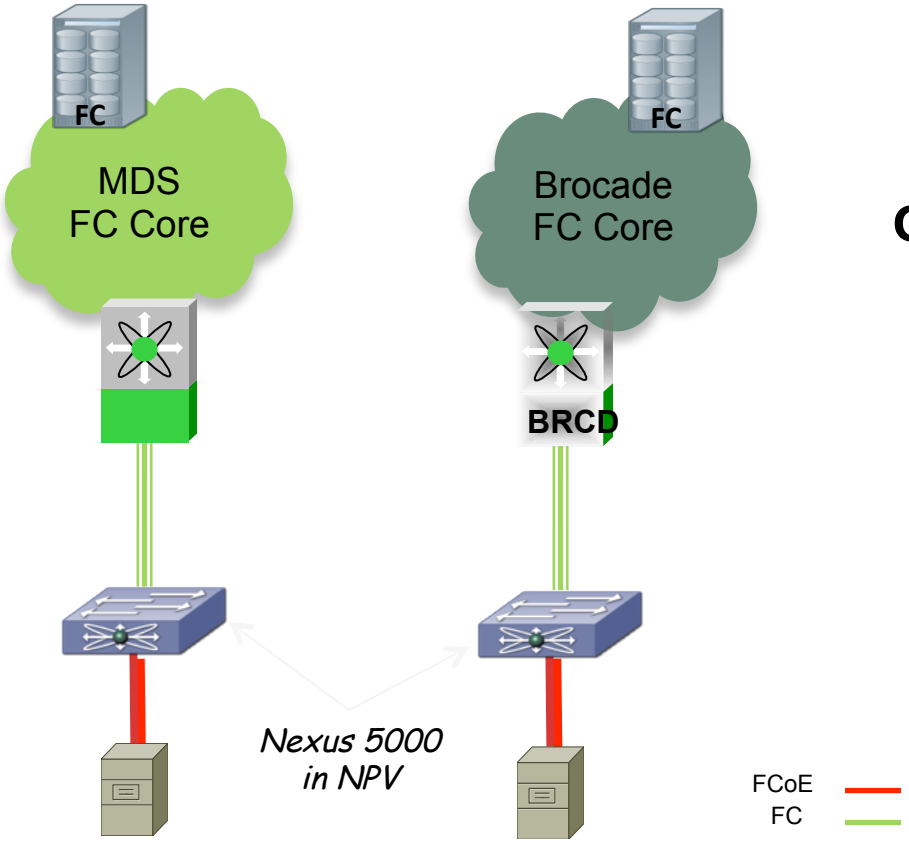
Speed (Gbps)	Max Distance (KM)
1	8000
2	4000
4	2000
8	1000
10	680

Common Deployments Today

Converged Access, FC Core

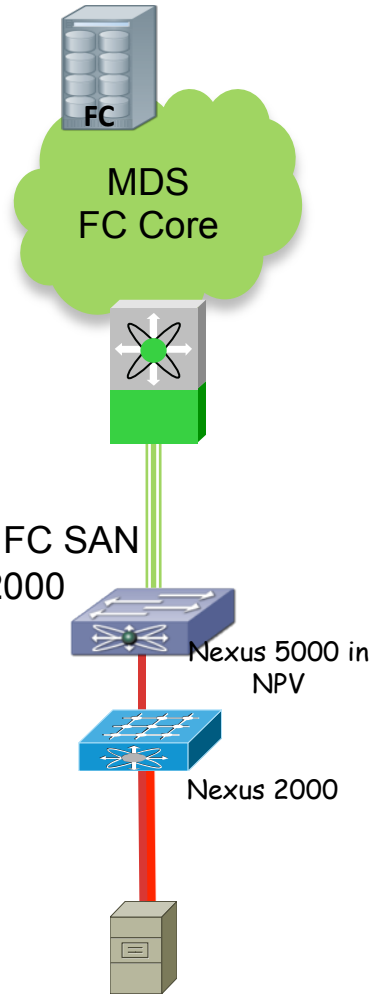
By Far the most popular deployment:

- Preserve Domain-ID on ToR switch
- Connectivity to Existing FC SAN
- Interoperability with Brocade installation



Gaining Traction:

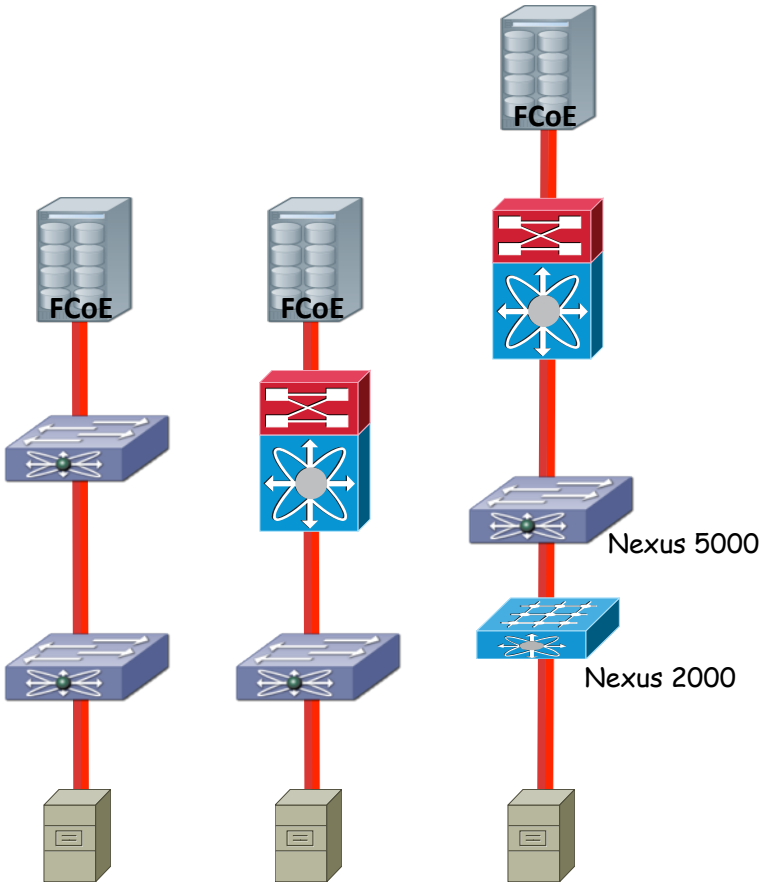
- Preserve Domain-ID
- Connectivity to Existing FC SAN
- Scale –Up with Nexus 2000

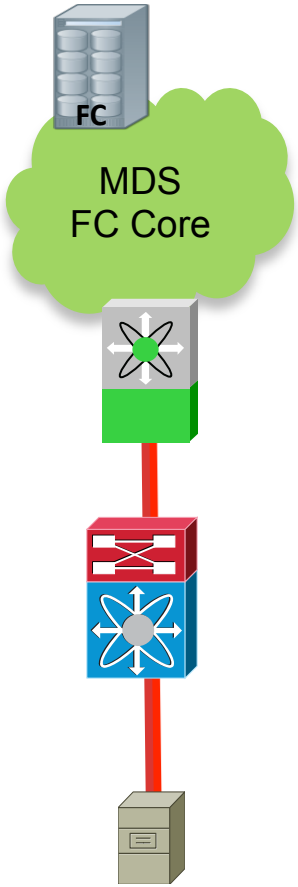


Multi-Hop FCoE

Slowly Gaining Momentum:

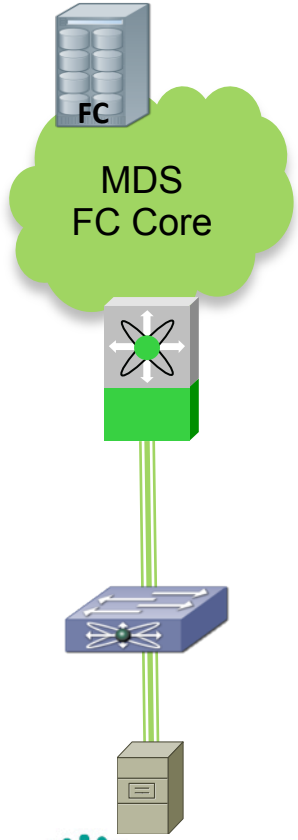
- A lot of interest for Nexus 2000 FCoE connectivity to Nexus 7000





Growing Interest:

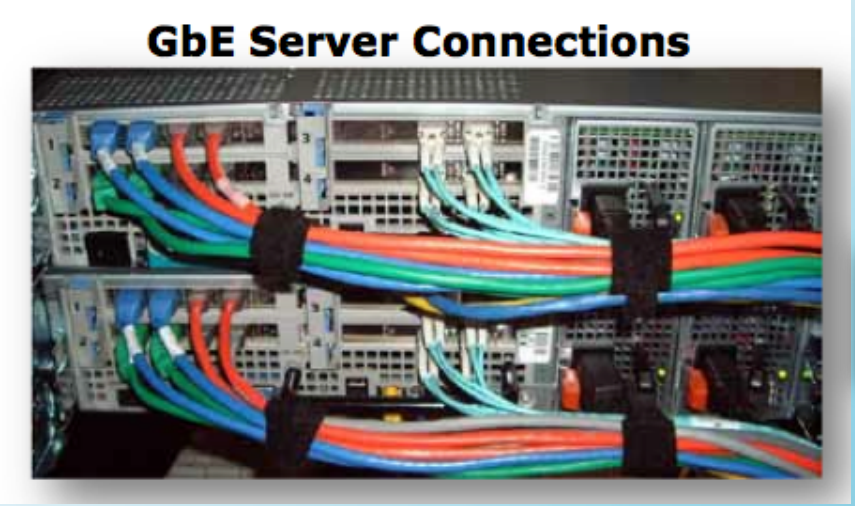
- Nexus 5000 as a Unified switch
- FC & Ethernet Connectivity
- Flexibility brought by Unified Port
- Especially attractive in cost sensitive and smaller markets

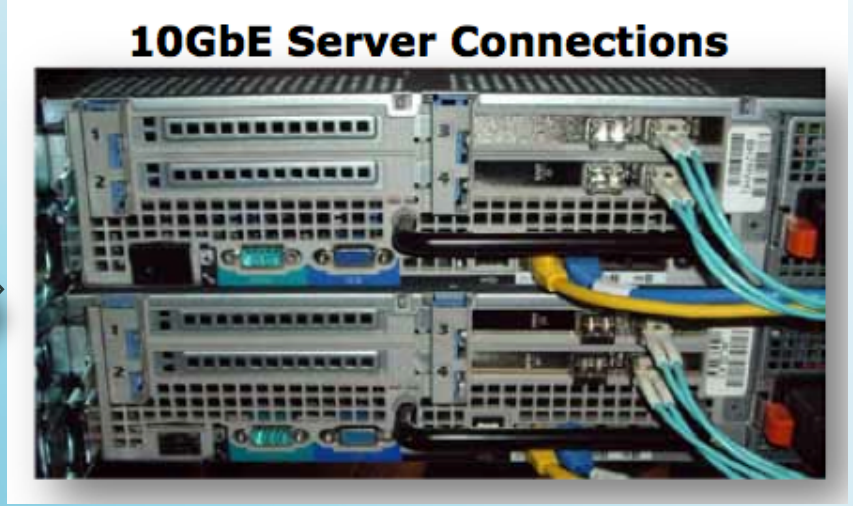


Some Deployments:

- Nexus 7000 as a storage director
- Analogous to MDS 9500 on the edge

10G Ethernet Simplifies Your Network





45% Reduction in Power per Rack

80% Reduction in Cables and Switch ports

15% Reduction in Infra-structure Costs

2x Improved Bandwidth per Server

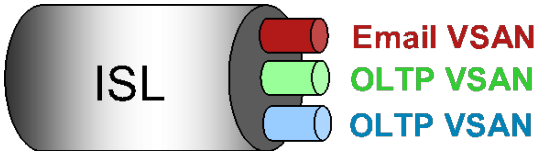
Source: Intel

MDS-Enabled Multi-layer SAN Virtualization

Virtual SAN (VSANs)

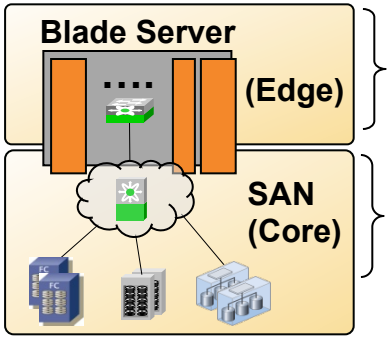
Partitioning of a SAN into multiple SANs (virtual SANS) for enabling fabric and storage consolidation

Link Virtualization



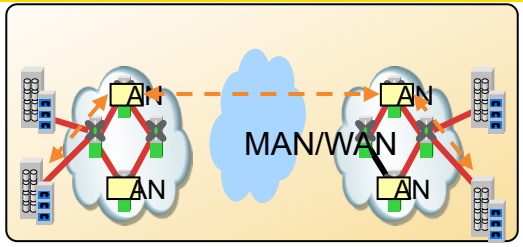
Enables multiple applications to share SANs with compromising performance (traffic management)

FlexAttach



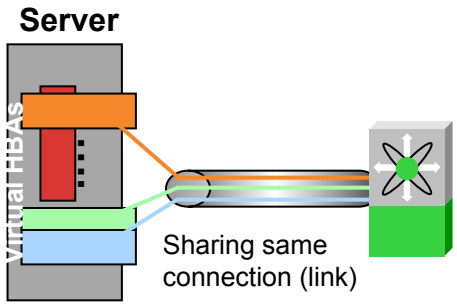
Enables server mobility, eliminating need for SAN and server teams to coordinate changes

Replication (FCIP)



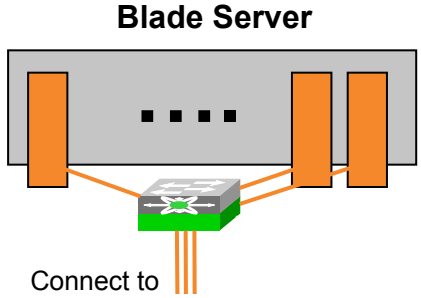
Enabling traditional FCIP replication for storage or tape using Cisco's vast expertise in TCP/IP. Also capable of advanced replication using IO/A or XRC.

Port Virtualization



Enables mobility, security, and QoS for SAN-attached Virtual Servers w/ N-Port ID Virtualization (NPIV)

Edge/Blade Switch Virtualization



Enables large-scale blade server deployments, simplifies management, and multi-vendor SAN connectivity