
TOMORROW starts here.
cisco
Cisco live!

Designing Storage Infrastructure in a Multi-protocol Data Centre

BRKDCT-2641

Adrian Quaife

Solution Consultant

Cisco live!

Agenda

- Introduction
- SAN Design Principles
- Technology Overview
- Storage Fabric Design Considerations
- Dedicated Data Centre SAN Topologies
- Converged Data Centre Topologies
- Distance Extension
- Q&A




Introduction

Trends and Key Requirements

Shaping Storage Networks

VIRTUALISATION / CLOUD

10X

growth in Physical and Virtual Servers by 2020*

MASSIVE DATA GROWTH

16X

growth in information created by 2020*

SOLID-STATE DRIVE

4X

growth in Solid-State Disks (SSD) by 2015**

MEGA DATA CENTRES

25%

growth in the Largest Data Centres by 2016***

High Bandwidth

Zero Downtime

Multi-protocol Storage Connectivity

Ease of Management, Fast Disaster Recovery, Low Latency

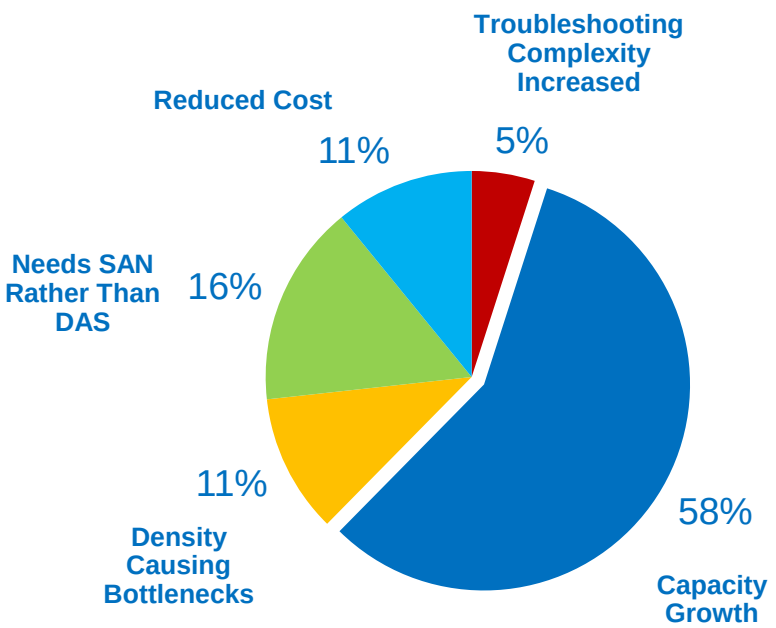
*2012 IDC Digital Universe Study

**: Gartner: "Marketing Essentials: Three Growth Opportunities in Storage Up to 2015", Aug. 2012

***: Gartner: "High-Tech Tuesday Webinar: The Shifting Face of the Data Centre", Feb. 2013

Server Virtualisation Impacts on Storage

What is the Predominant Impact of Server Virtualisation on Storage?



- Cost not predominant factor
  - Savings on servers offset with increased storage costs
- Server virtualisation places new demands on SAN
  - Zoning complexities
  - More diversity of storage access
- Managing capacity growth demands
- Need to deliver single wire access for all workloads in a DC

Source: The InfoPro. Used with Permission. Partial, interim results from a study begun in August 2011

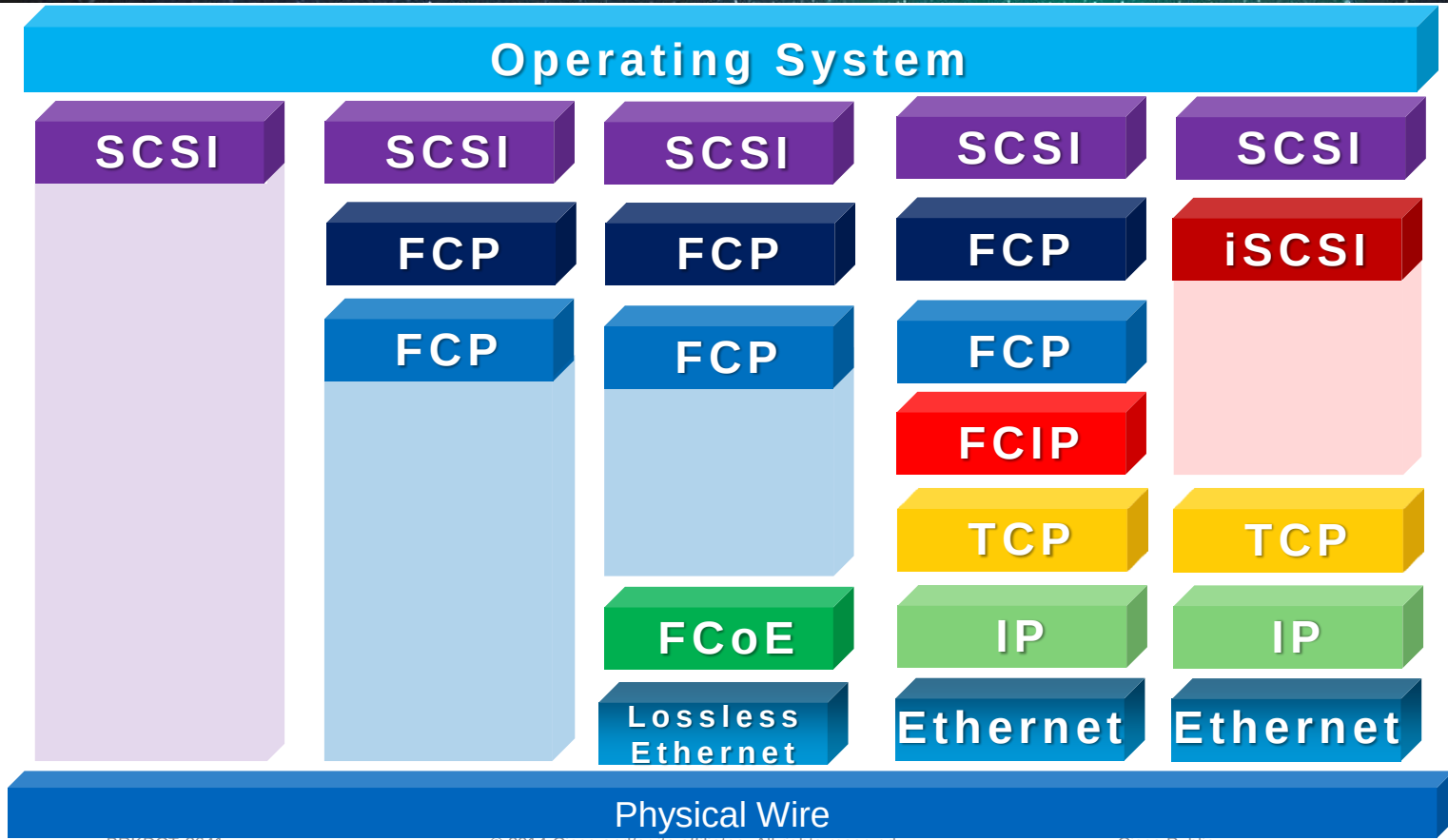



Design Principles

Protocols & Technologies

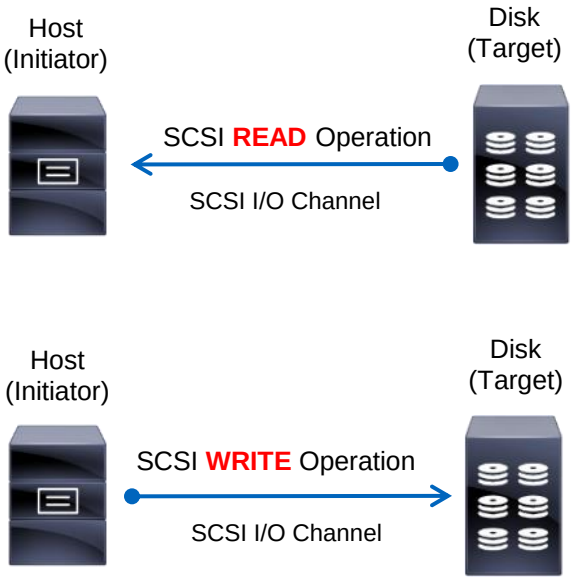
- SCSI foundations
- Addressing & Routing
- FCoE
- Connectivity Types
- VSAN
- NPV
- VSAN & IVR
- Smart Zoning
- Management
- SAN Security

SCSI is the Foundation



Fibre Channel – Foundations

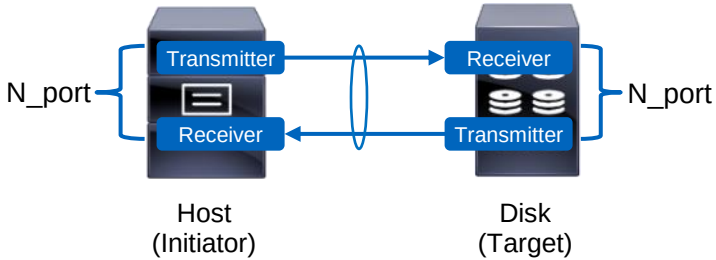
Based on SCSI

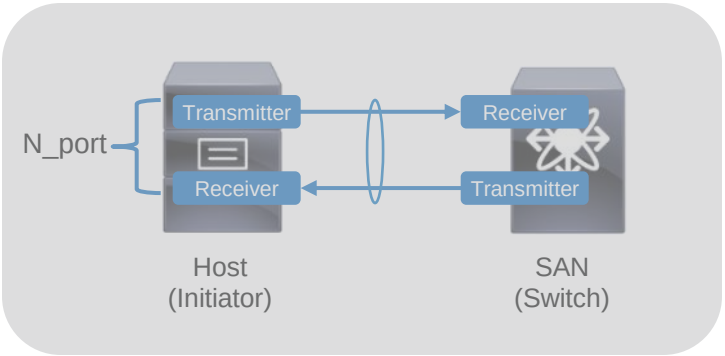


- Foundational protocol, forms the basis of an I/O transaction
- Communications are based upon Point to Point
- Storage is accessed at a block-level via SCSI
- High-performance interconnect providing high I/O throughput
- The foundation for all block-based storage connectivity
- Mature - SCSI-1 developed in 1986

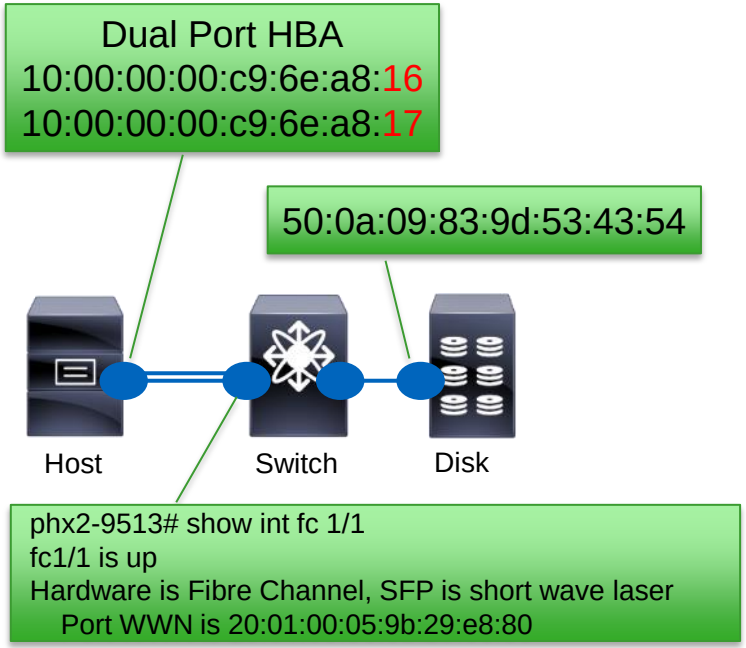
Fibre Channel - Communications

- 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 Addressing

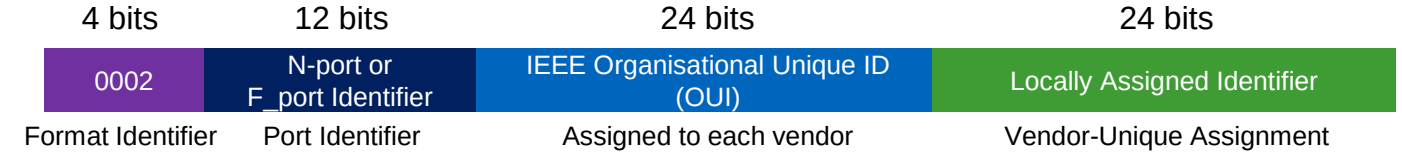


Every Fibre Channel port and node has a hard-coded address called World Wide Name (WWN)

- Allocated to manufacturer by IEEE
- Coded into each device when manufactured
- 64 or 128 bits

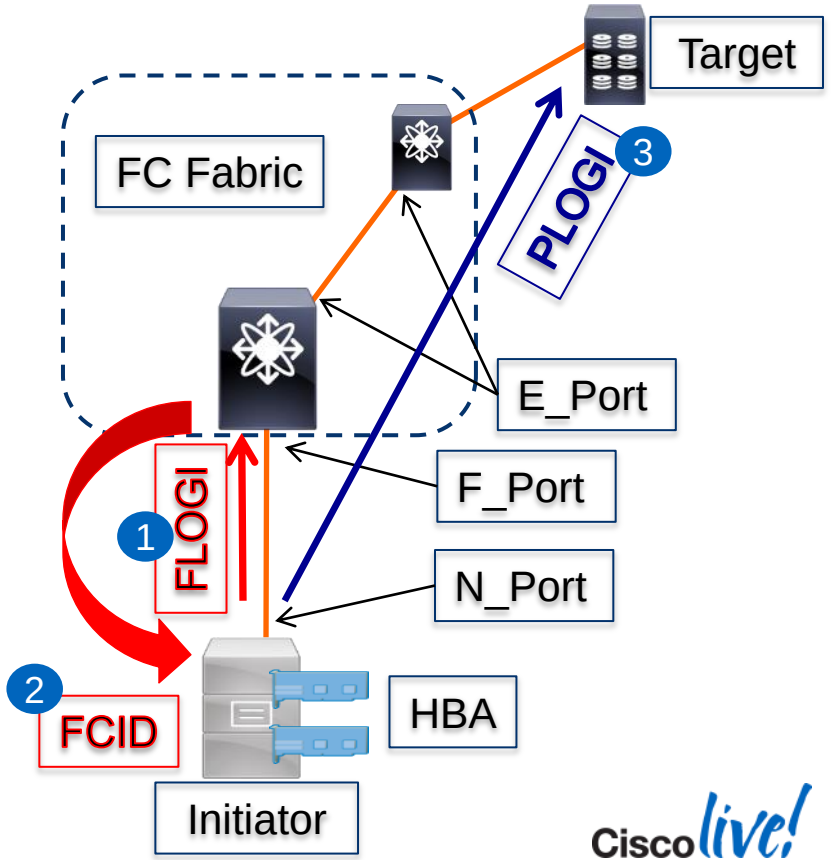
Switch Name Server maps WWNs to FCID

- WWNN uniquely identify devices
- WWPN uniquely identify each port in a device



My Port is up…Can I Talk Now?
FLOGIs/PLOGIs

- Step 1: Fabric Login (FLOGI)
  - Determines the presence or absence of a Fabric
  - Exchanges Service Parameters with the Fabric
  - Switch identifies the WWN in the service parameters of the accept frame and assigns a Fibre Channel ID (FCID)
  - Initialises the buffer-to-buffer credits
- Step 2: Port Login (PLOGI)
  - Required between nodes that want to communicate
  - Similar to FLOGI – Transports a PLOGI frame to the designation node port
  - In P2P topology (no fabric present), initialises buffer-to-buffer credits



FC_ID Address Model

- FC_ID address models help speed up FC 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)

	8 Bits	8 Bits	8 Bits
Switch Topology Model	Switch Domain	Area	Device
Private Loop Device Address Model	00	00	Arbitrated Loop Physical Address (AL_PA)
Public Loop Device Address Model	Switch Domain	Area	Arbitrated Loop Physical Address (AL_PA)

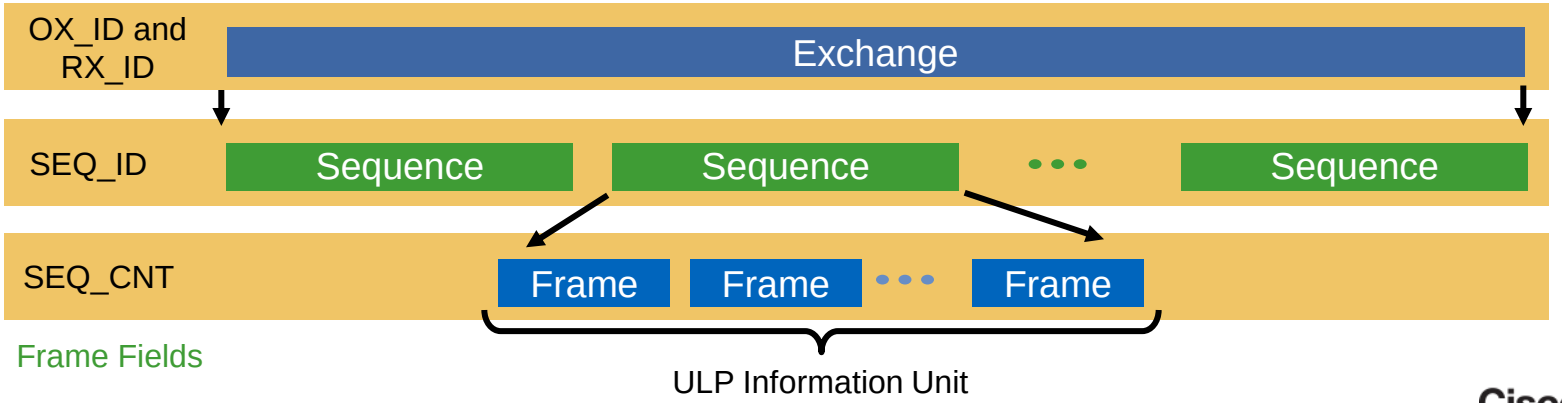
FSPF

Fabric Shortest Path First

- Provides routing services within any FC fabric
- Supports multipath routing
- Bases path status on a link state protocol similar to OSPF
- Routes hop by hop, based only on the domain ID
- Runs on E ports or TE ports and provides a loop free topology
- Runs on a per VSAN basis. Connectivity in a given VSAN in a fabric is guaranteed only for the switches configured in that VSAN.
- Uses a topology database to keep track of the state of the links on all switches in the fabric and associates a cost with each link
- Fibre Channel standard ANSI T11 FC-SW2

Fibre Channel FC-2 Hierarchy

- Multiple exchanges are initiated between initiators (hosts) and targets (disks)
- Each exchange consists of one or more bidirectional sequences
- Each sequence consists of one or more frames
- For the SCSI3 ULP, each exchange maps to a SCSI command



What Is FCoE?

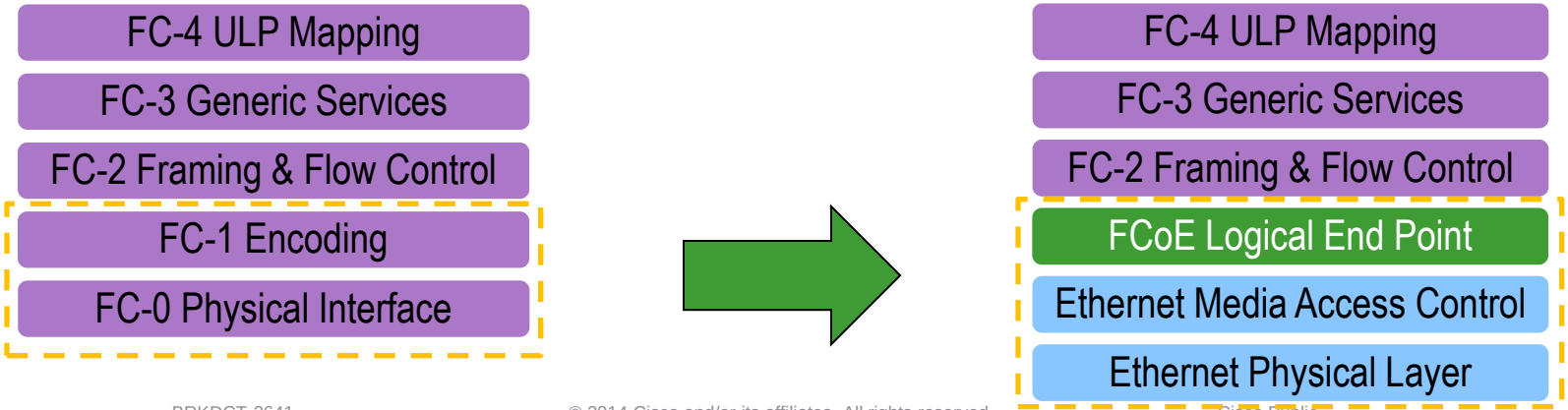
It's Fibre Channel

From a Fibre Channel standpoint it's

- FC connectivity over a new type of cable called… Ethernet

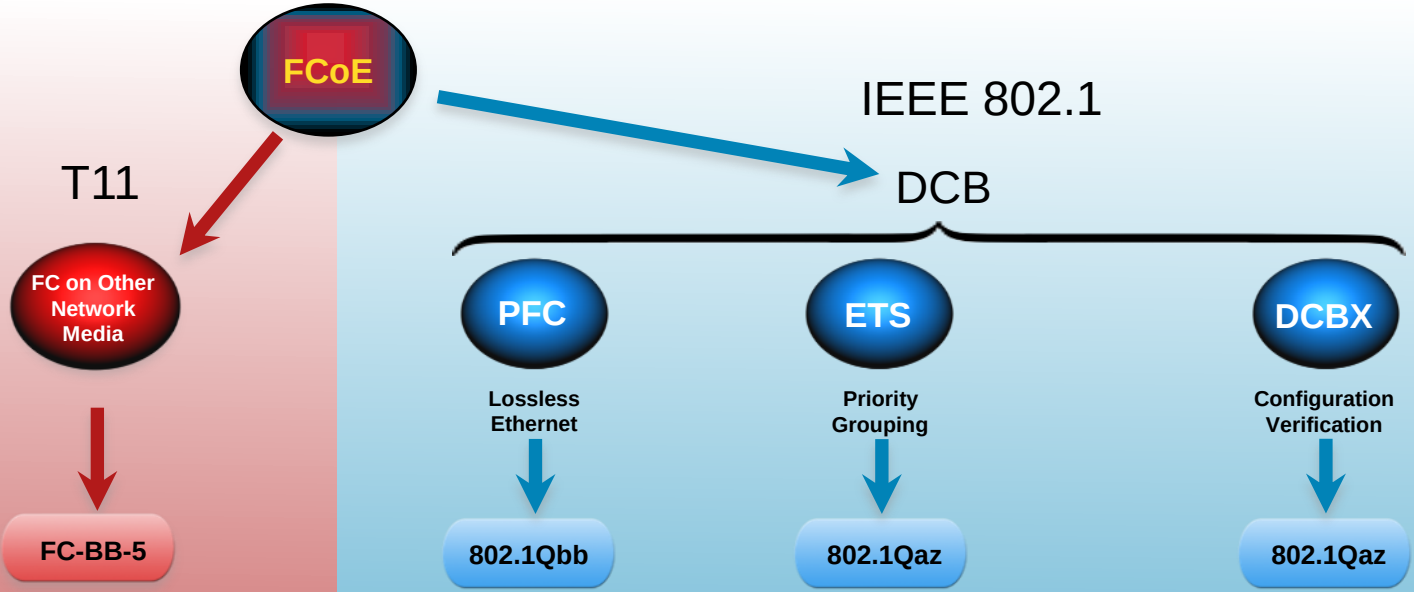
From an Ethernet standpoints it's

- Yet another ULP (Upper Layer Protocol) to be transported



Standards for FCoE

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



Standard Status	Technically stable October, 2008 Completed in June 2009 **Published in May, 2010**	Sponsor Ballot July 2010 **Published Fall 2011**	Sponsor Ballot October 2010 **Published Fall 2011**	Sponsor Ballot October 2010 **Published Fall 2011**

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 Initialisation Protocol)

- It is the control plane protocol
- It is used to discover the FC entities connected to an Ethernet cloud
- It is also used to login to and logout from the FC fabric
- Uses unique BIA on CNA for MAC

The Two Protocols Have

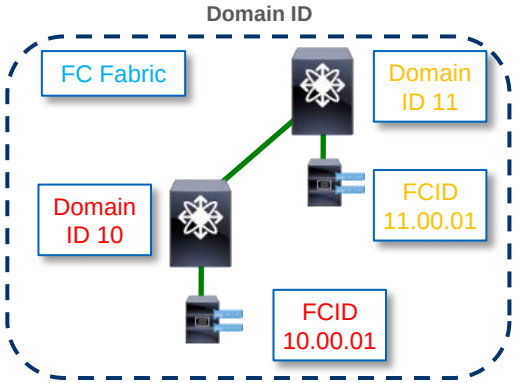
- Two different Ethertypes
- Two different frame formats
- Both are defined in FC-BB-5

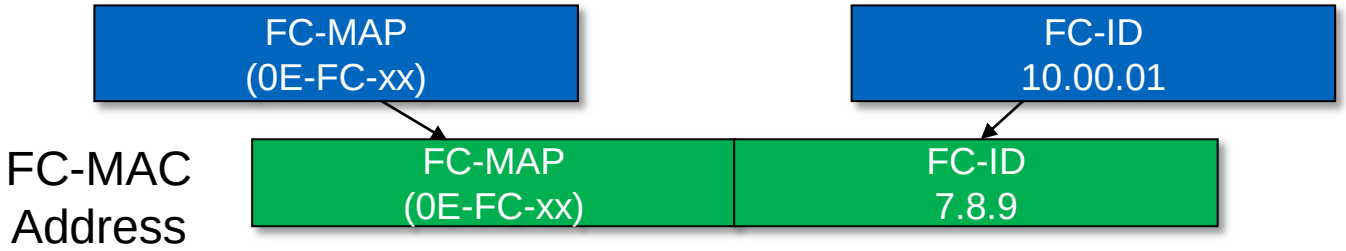
http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/white_paper_c11-560403.html

Enode MAC Address
Fibre Channel over Ethernet Addressing Scheme

Reference

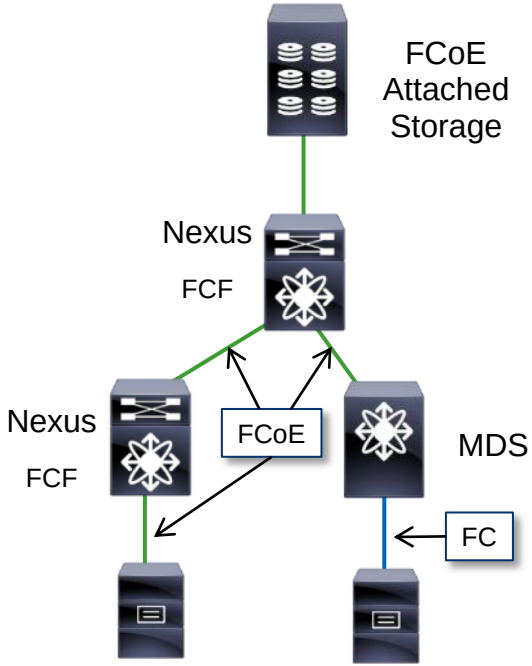
- Enode MAC assigned for each FCID
- Enode MAC composed of a FC-MAP and FCID
  - FC-MAP is the upper 24 bits of the Enode's MAC
  - FCID is the lower 24 bits of the Enode's MAC
- FCoE forwarding decisions still made based on FSPF and the FCID within the Enode MAC





What is an FCoE Switch?

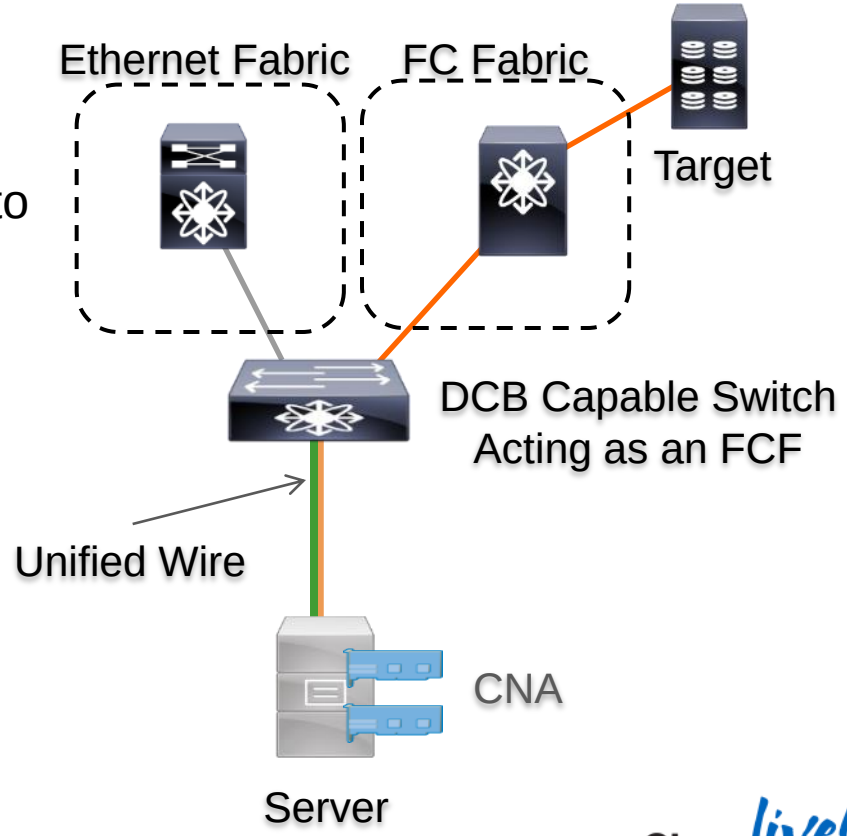
- FCF (Fibre Channel Forwarder) accepts the Fibre Channel frame encapsulated in an Ethernet packet and forwards that packet over a VLAN across an Ethernet network to a remote FCoE end device
- FCF is a logical FC switch inside an FCoE switch
  - Fibre Channel login happens at the FCF
  - Contains an FCF-MAC address
  - Consumes a Domain ID
- FCoE encapsulation/decapsulation happens within the FCF
- NPV devices are ***not*** FCF's and do ***not*** have domains
- Want more? BRKDCT-1044
  - FCoE for the IP Network Engineer



FCoE, Same Model as FC

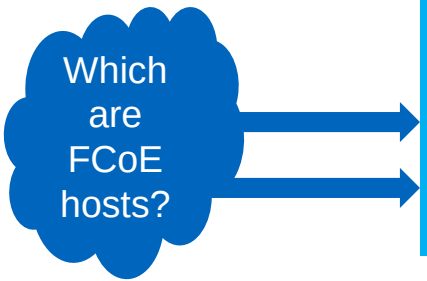
Reference

- 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



FCoE is Operationally Identical

- Supports both FC and FCoE
- FCoE is treated exactly the same as FC
- After zoning device perform registration and then performs discovery



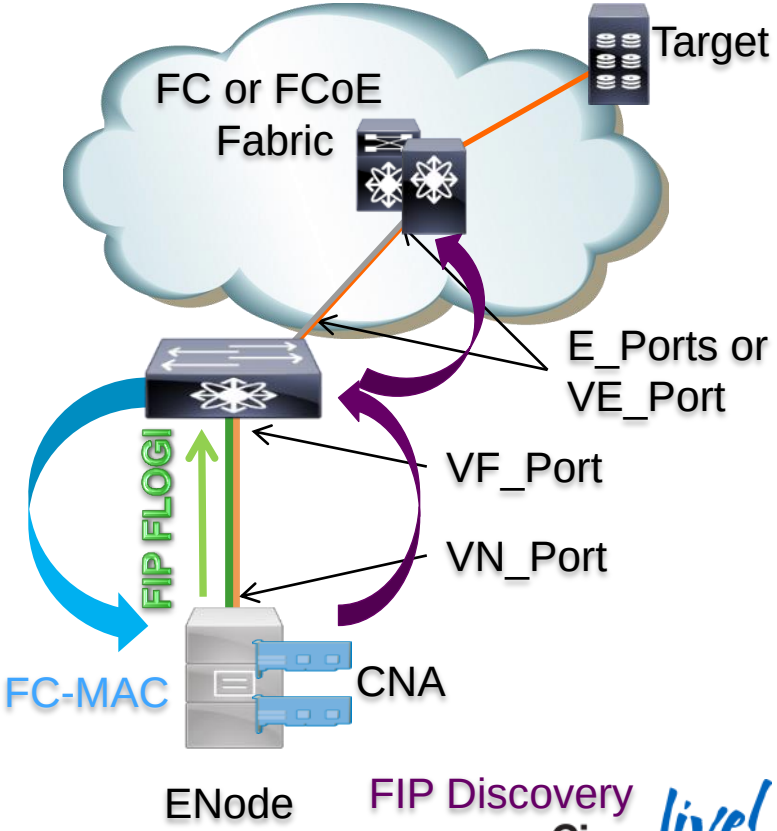

```
phx2-9513# show fcns database vsan 42

VSAN 42:
--------------------------------------------------------------------------
FCID        TYPE  PWWN                    (VENDOR)        FC4-TYPE:FEATURE
--------------------------------------------------------------------------
0xac0600    N     50:0a:09:83:8d:53:43:54 (NetApp)        scsi-fcp:target
0xac0700    N     50:0a:09:84:9d:53:43:54 (NetApp)        scsi-fcp:target
0xac0c00    N     20:41:54:7f:ee:07:9c:00 (Cisco)         npv
0xac1800    N     10:00:00:00:c9:6e:b7:f0                 scsi-fcp:init fc-gs
0xef0000    N     20:01:a0:36:9f:0d:eb:25                 scsi-fcp:init fc-gs
```

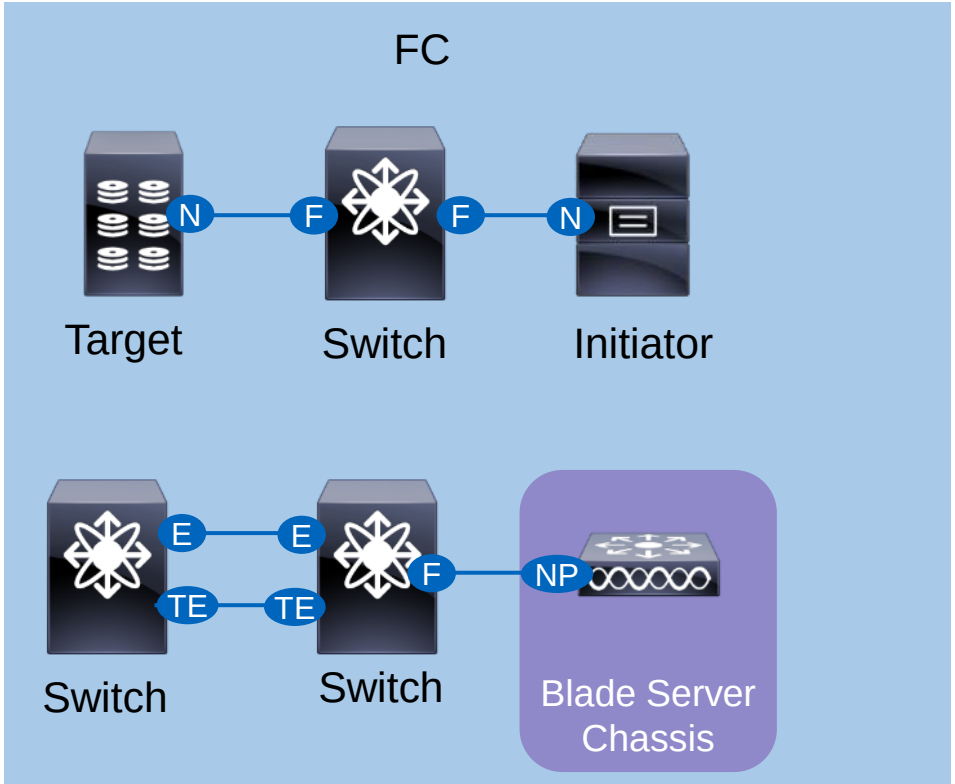
After Link Is Up, Accessing Storage
FIP and FCoE Login Process

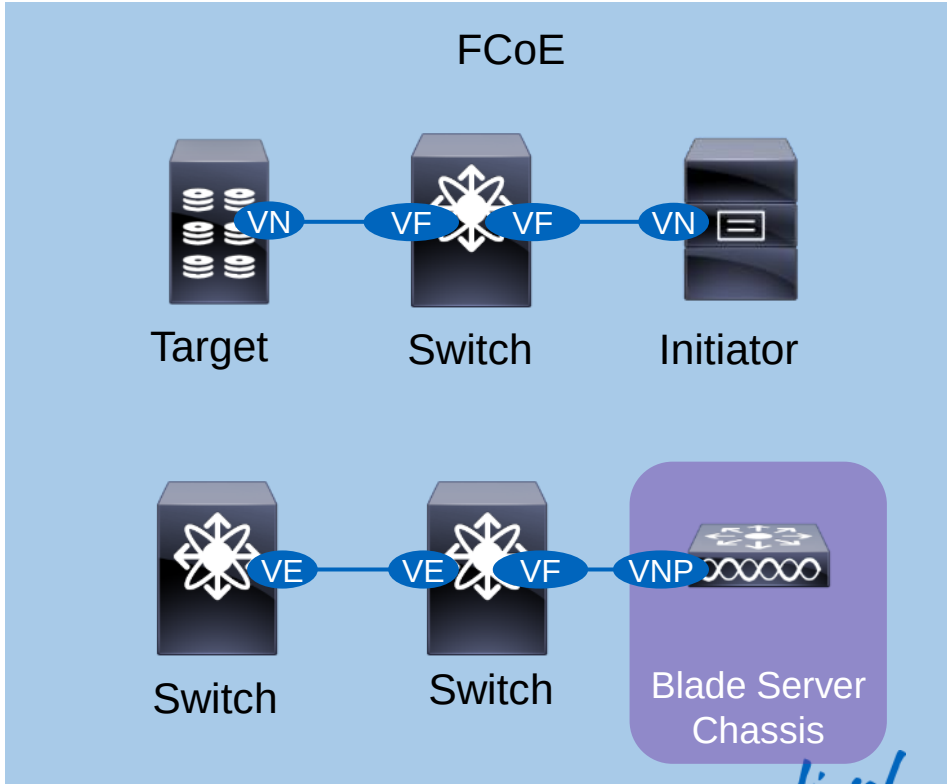
Reference

- Step 1: FIP Discovery Process
  - Enables FCoE adapters to discover which VLAN to transmit & receive FCoE frames
  - Enables FCoE adapters and FCoE switches to discover other FCoE capable devices
  - Verifies Lossless Ethernet is capable of FCoE transmission
- Step 2: FIP Login Process
  - Similar to existing Fibre Channel Login process - sends FLOGI to upstream FCF
  - Adds the negotiation of the MAC address to use Fabric Provided MAC Address (FPMA)
  - FCF assigns the host a Enode MAC address to be used for FCoE forwarding



Connectivity Types

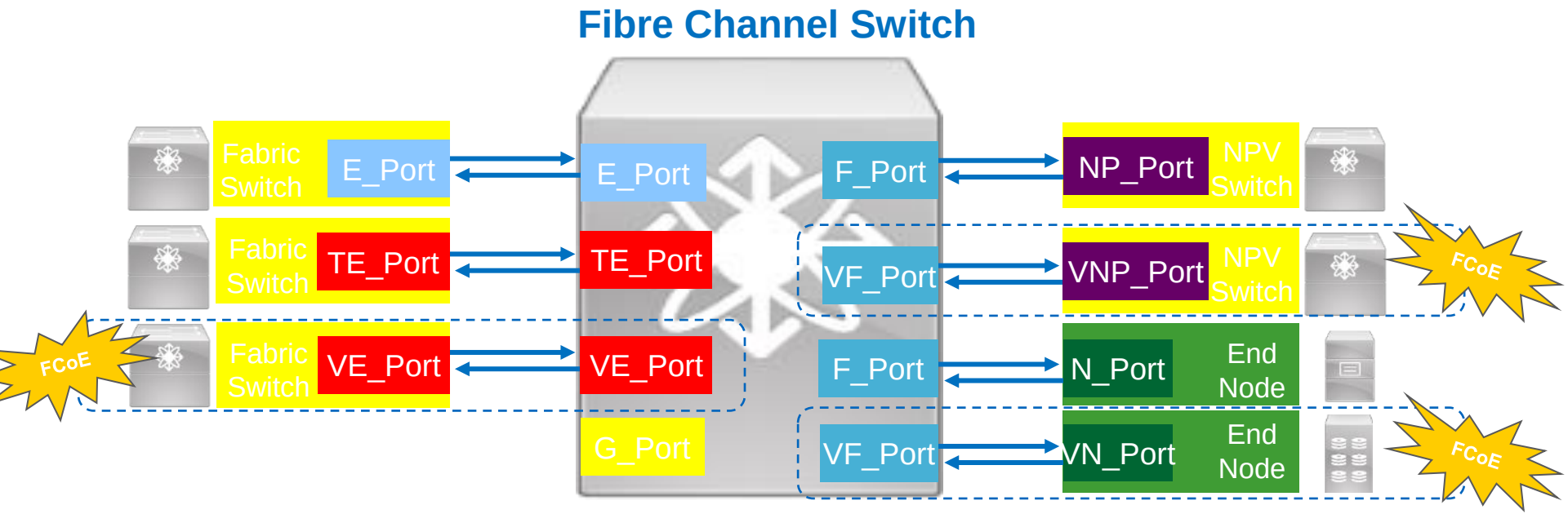




Fibre Channel Port Types

Summary

Reference



The Story of Interface Speeds

Protocol	Clocking Gbps	Encoding Data/Sent	Data Rate	
			Gbps	MB/s
8G FC	8.500	8b/10b	6.8	850
10G FC	10.51875	64b/66b	10.2	1,275
10G FCoE	10.3125	64b/66b	10.0	1,250
16G FC	14.025	64b/66b	13.6	1,700
32G FC	28.050	64b/66b	27.2	3,400
40G FCoE	41.250	64b/66b	40.0	5,000

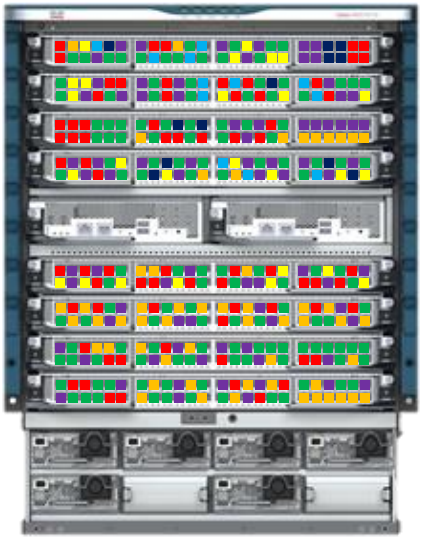
- Comparing speeds is more complex than just the “apparent” speed
- Data throughput is based on both the interface clocking (how fast the interface transmits) and how efficient the interface transmits (how much encoding overhead)

VSANs

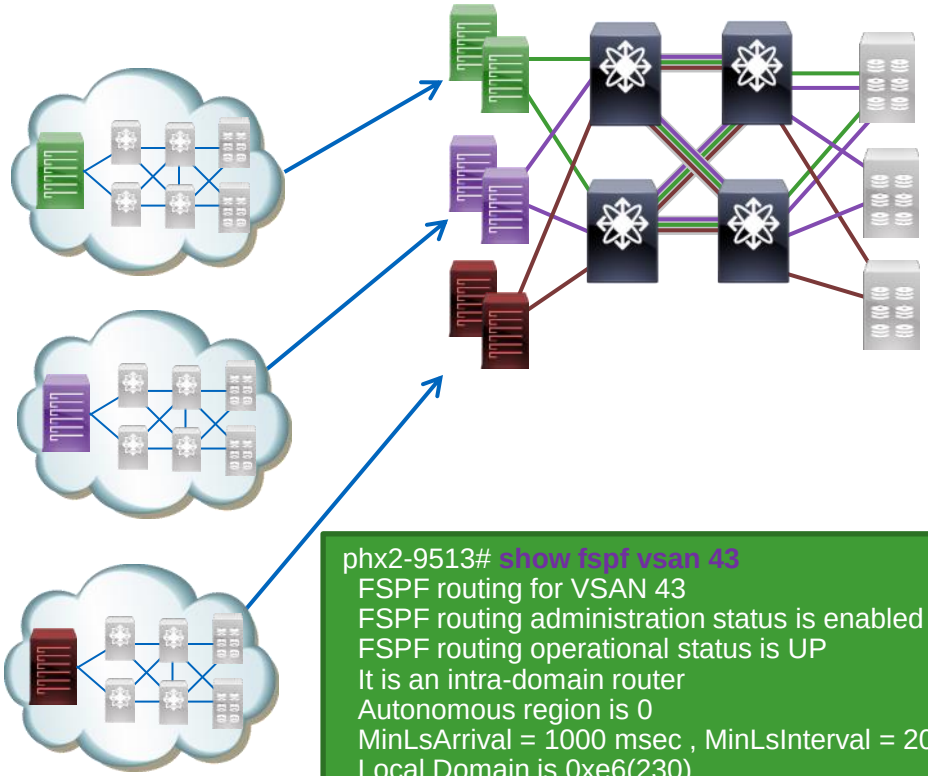
Introduced in 2002

- A Virtual SAN (VSAN) Provides a Method to Allocate Ports within a Physical Fabric and Create Virtual Fabrics
- 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.
- ANSI T.11 committee and is now part of Fibre Channel standards as Virtual Fabrics

Per Port Allocation



VSAN



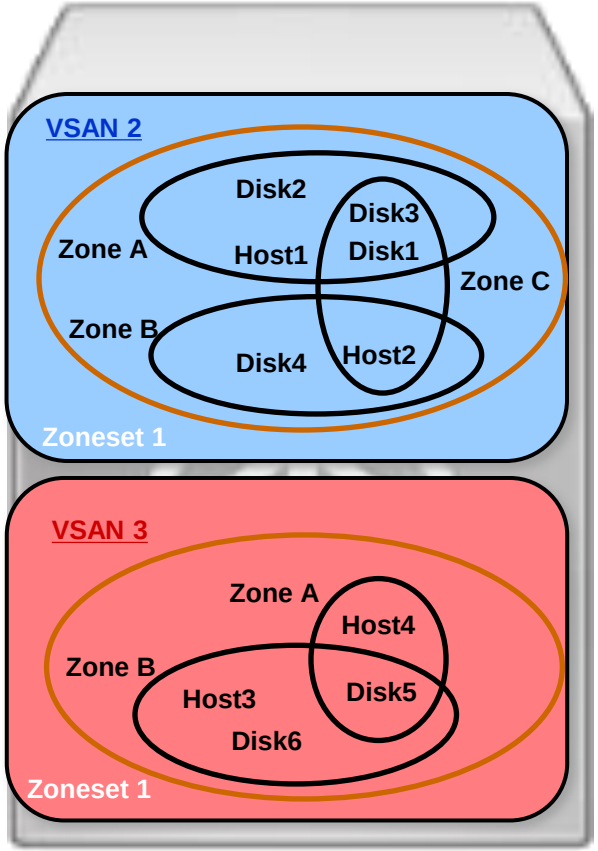
- Assign ports to VSANs
- Logically separate fabrics
- Hardware enforced
- Prevents fabric disruptions
  - RSCN sent within fabric only
- Each fabric service (zone server, name server, login server, etc.) operates independently in each VSAN
- Each VSAN is configured and managed independently

```
phx2-9513# show fspf vsan 43
  FSPF routing for VSAN 43
  FSPF routing administration status is enabled
  FSPF routing operational status is UP
  It is an intra-domain router
  Autonomous region is 0
  MinLsArrival = 1000 msec , MinLsInterval = 2000 msec
  Local Domain is 0xe6(230)
  Number of LSRs = 3, Total Checksum = 0x00012848
```

```
vsan database
  vsan 2 interface fc1/1
  vsan 2 interface fc1/2
  vsan 4 interface fc1/8
  vsan 4 interface fc1/9
```

```
phx2-9513# show zoneset active vsan 43
  zoneset name UCS-Fabric-B vsan 43
  zone name UCS-B-VMware-Netapp vsan 43
```

Zoning & VSANs



- First assign physical ports to VSANs
- Then configure zones within each VSAN
- Assign zones to active zoneset

- Each VSAN has its own zoneset
- A zone consists of multiple zone members
- Members in a zone can access each other; members in different zones cannot access each other
- Devices can belong to more than one zone

Zoning Examples

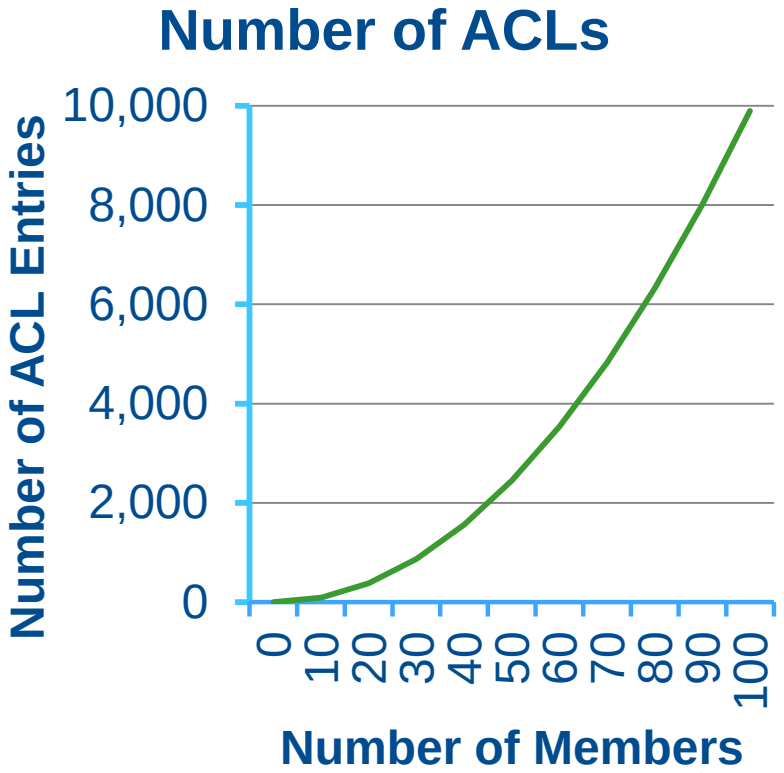
- Non-zoned devices are members of the default zone
- A physical fabric can have a maximum of 16,000 zones
- Attributes can include pWWN, FC alias, FCID, FWWN, Switch Interface fc x/y, Symbolic node name, Device alias

```
zone name AS01_NetApp vsan 42
  member pwwn 20:03:00:25:b5:0a:00:06
  member pwwn 50:0a:09:84:9d:53:43:54
```

```
device-alias name AS01
   pwwn 20:03:00:25:b5:0a:00:06
device-alias name NTAP
   member pwwn 50:0a:09:84:9d:53:43:54
zone name AS01_NetApp vsan 42
   member device-alias AS01
   member device-alias NTAP
```

The Trouble with Sizable Zoning
All Zone Members are Created Equal

- Standard zoning model just has “members”
- Any member can talk to any other member
- Recommendation: 1-1 zoning
- Each pair consumes an ACL entry in TCAM
- Result: n*(n-1) entries
- Admin pays price for internal inefficiency



Smart Zoning

Operation	Today – 1:1 Zoning			Today – Many - Many			Smart Zoning		
	Zones	Cmds	ACLs	Zones	Cmds	ACLs	Zones	Cmds	ACLs
Create zones(s)	32	96	64	1	13	132	1	13	64
Add an initiator	+4	+12	+8		+1	+24		+1	+8
Add a target	+8	+24	+16		+1	+24		+1	+16

8 x I
4 x T

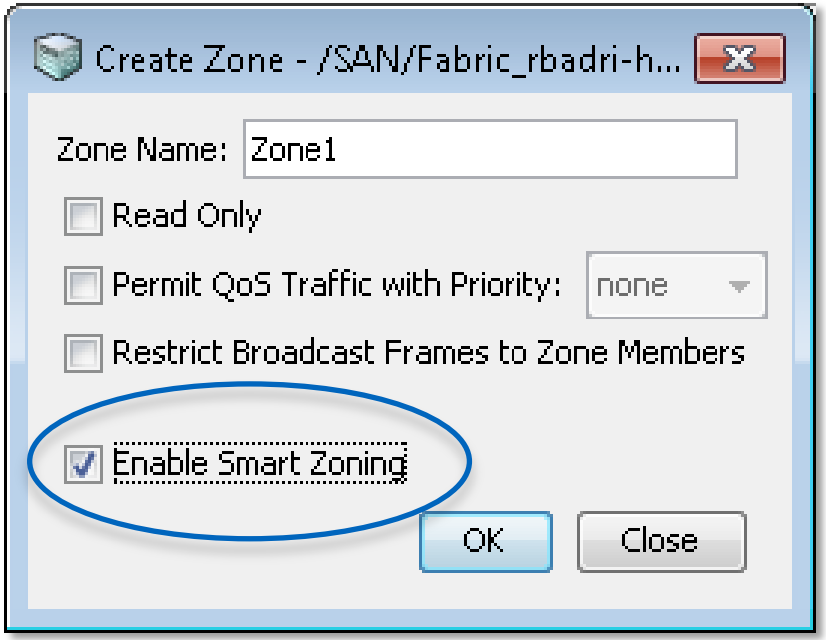
- New in NX-OS 5.2(6)
- Allows storage admins to create larger zones while still keeping premise of single initiator & single target
- Dramatic reduction SAN administrative time for zoning
- Utility to convert existing zone or zoneset to Smart Zoning

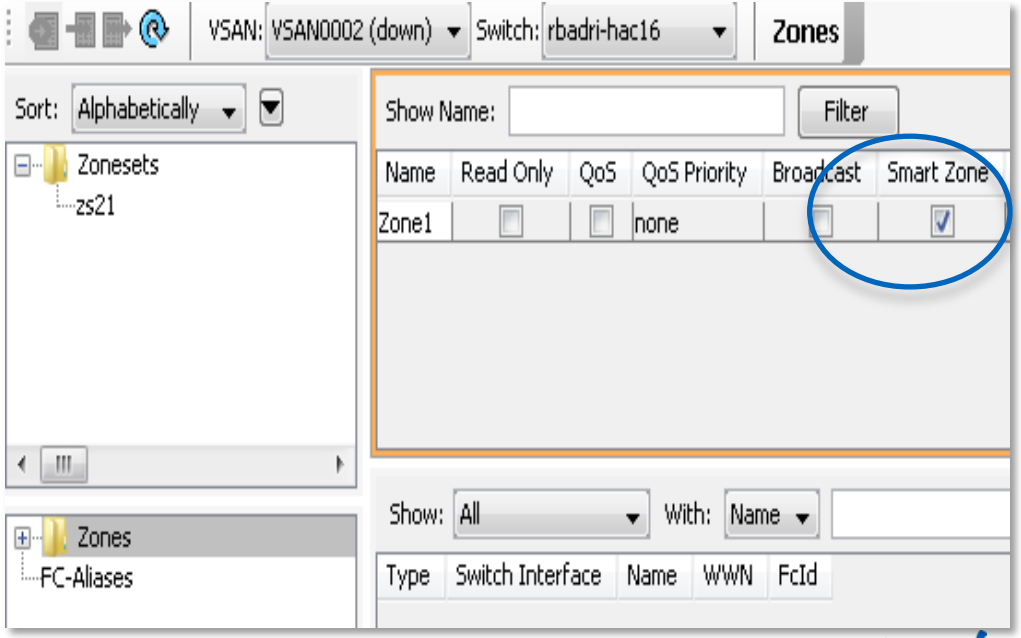
How to Enable Smart Zoning

Reference

New Zone

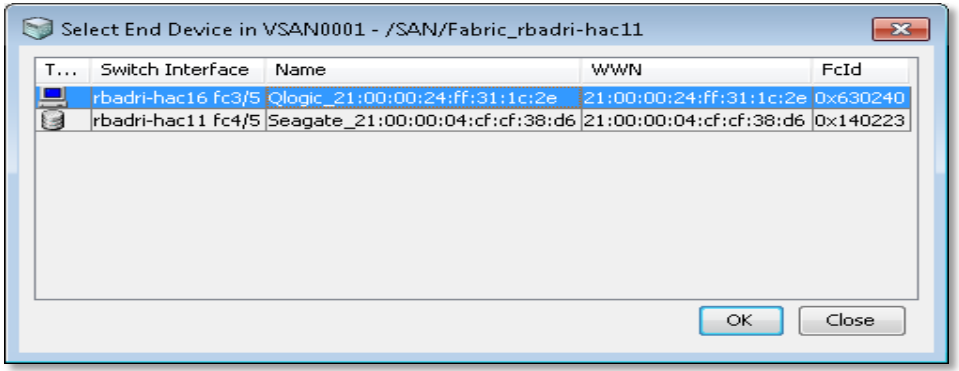
Existing Zone

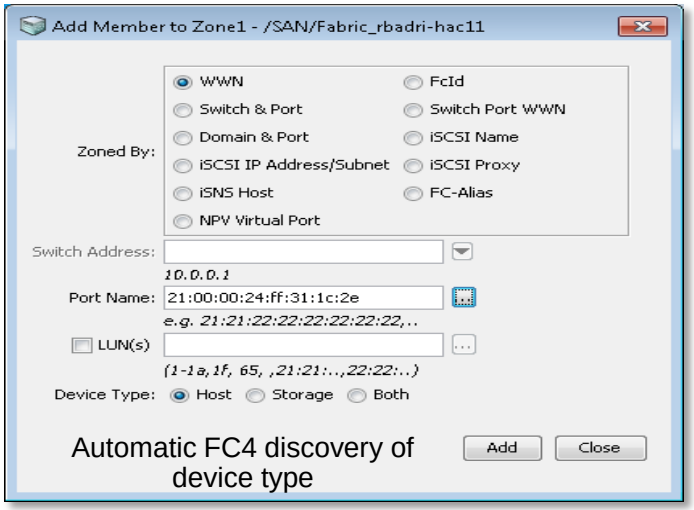




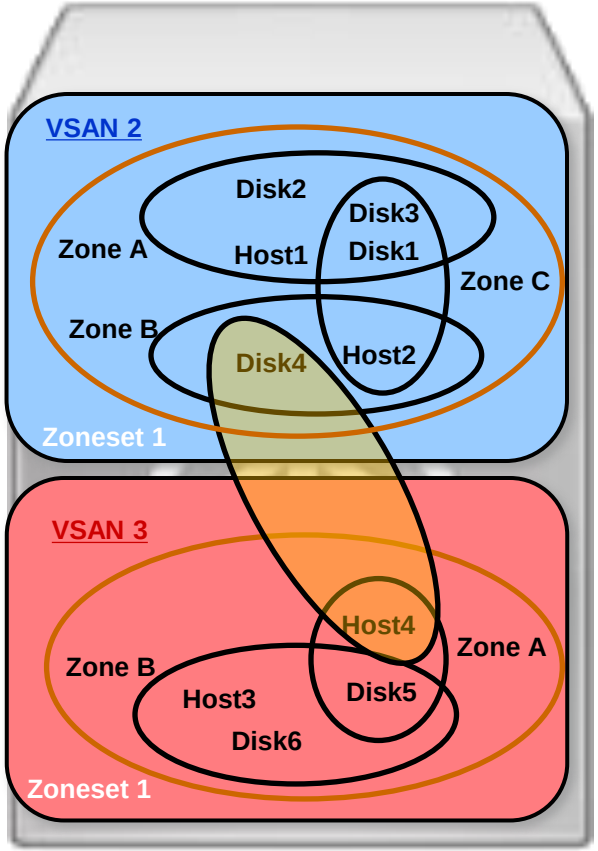
Adding Pre-Discovered End Devices as Zone Members

Reference





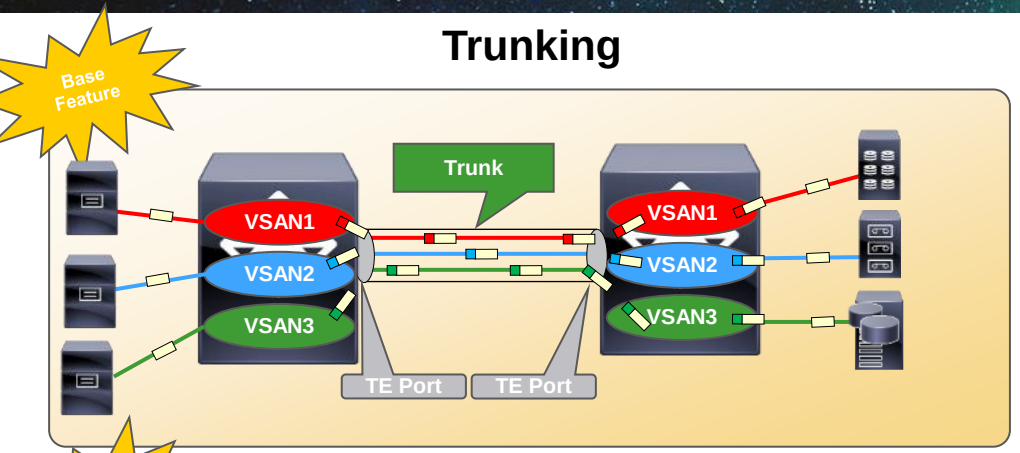
Inter-VSAN Routing



- Enables devices in different VSANs to communicate
- Allows selective routing between specific members of two or more VSANs
  - Traffic flow between selective devices
- Resource sharing, i.e., tape libraries and disks
- IVR Zoneset
  - A collection of IVR zones that must be activated to be operational

Trunking & Port Channels

Trunking

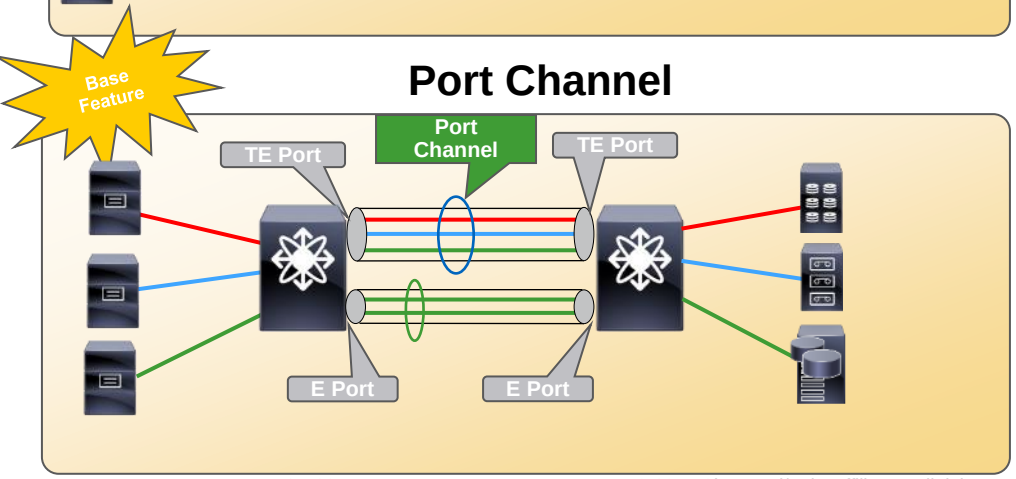


Single-link ISL or PortChannel ISL can be configured to become EISL – (TE_Port)

Traffic engineering with pruning VSANs on/off the trunk

Efficient use of ISL bandwidth

Port Channel



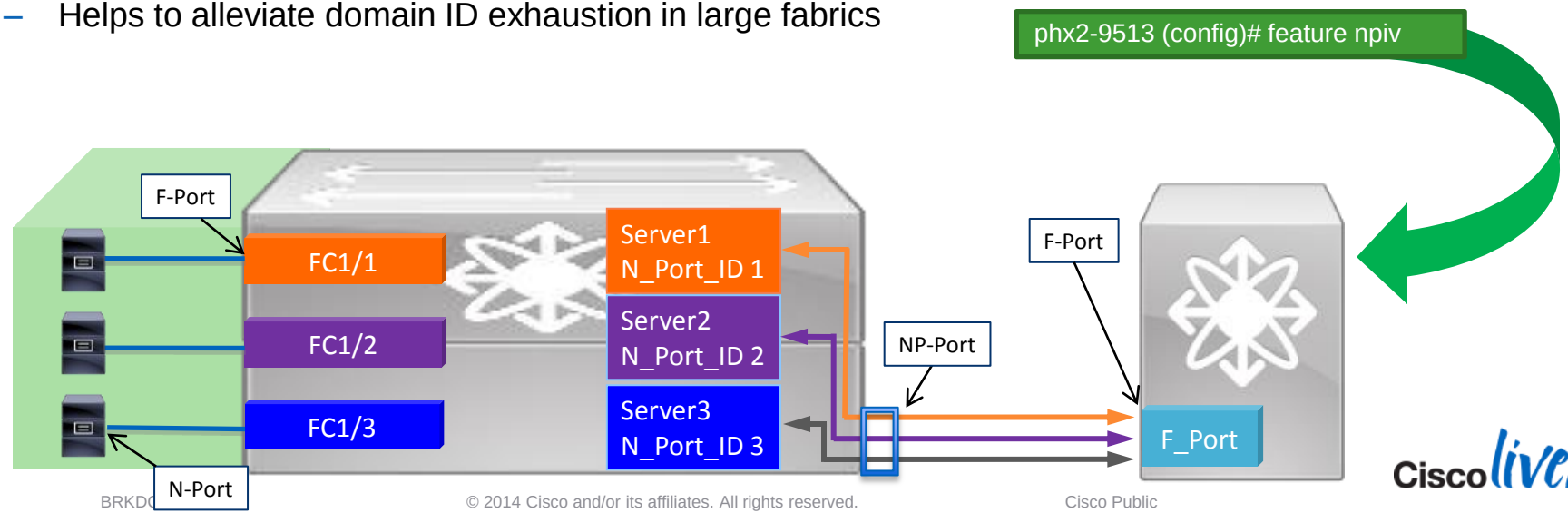
Up to 16 links can be combined into a PortChannel increasing the aggregate bandwidth by distributing traffic granularly among all functional links in the channel

Load balances across multiple links and maintains optimum bandwidth utilisation. Load balancing is based on the source ID, destination ID, and exchange ID

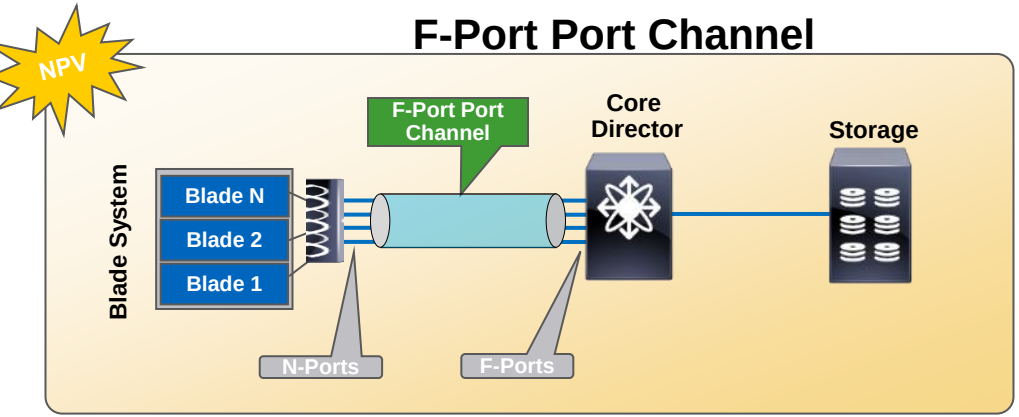
If one link fails, traffic previously carried on this link is switched to the remaining links. To the upper protocol, the link is still there, although the bandwidth is diminished. The routing tables are not affected by link failure

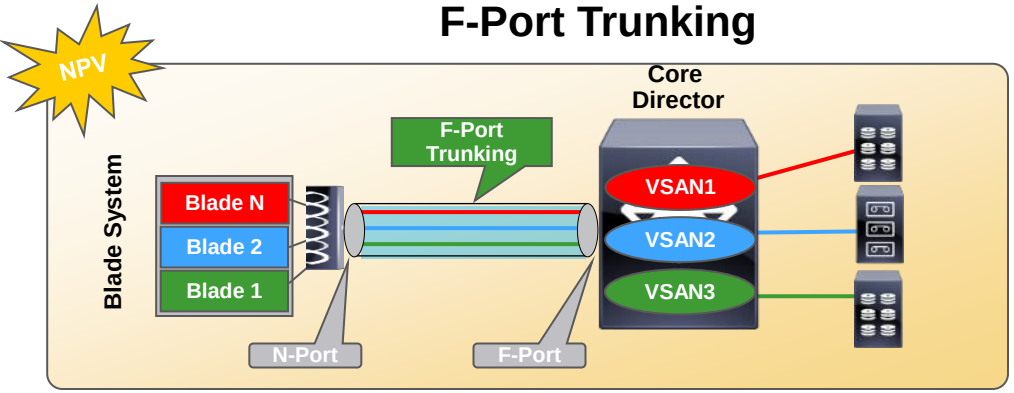
N-Port Virtualisation
Scaling Fabrics with Stability

- N-Port Virtualiser (NPV) utilises NPIV functionality to allow a “switch” to act like a server/HBA performing multiple fabric logins through a single physical link
- Physical servers connect to the NPV switch and login to the upstream NPIV core switch
- No local switching is done on an FC switch in NPV mode
- FC edge switch in NPV mode does not take up a domain ID
  - Helps to alleviate domain ID exhaustion in large fabrics



F-Port Port Channel and F-Port Trunking
Enhanced Blade Switch Resiliency





F-Port Port Channel w/ NPV

- Bundle multiple ports in to 1 logical link
  - Any port, any module
- High-Availability (HA)
  - Blade Servers are transparent if a cable, port, or line cards fails
- Traffic Management
  - Higher aggregate bandwidth
  - Hardware-based load balancing

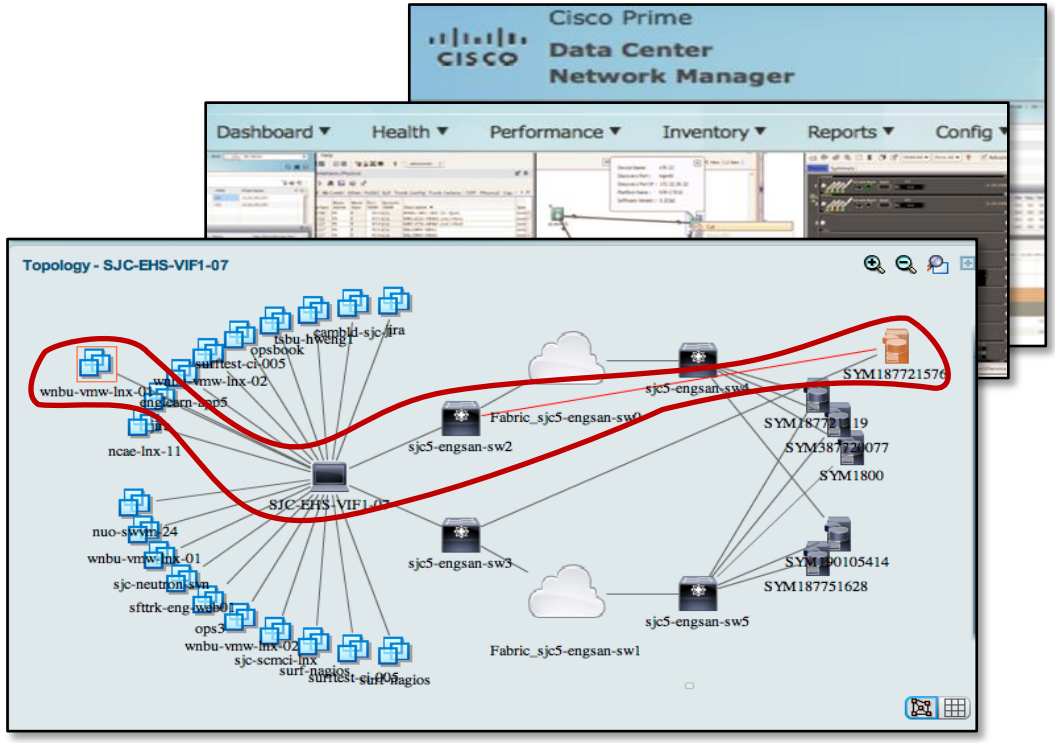
F-Port Trunking w/ NPV

- Partition F-Port to carry traffic for multiple VSANs
- Extend VSAN benefits to Blade Servers
  - Separate management domains
  - Separate fault isolation domains
  - Differentiated services: QoS, Security

Cisco Prime Data Centre Network Manager Feature Support and User Interface

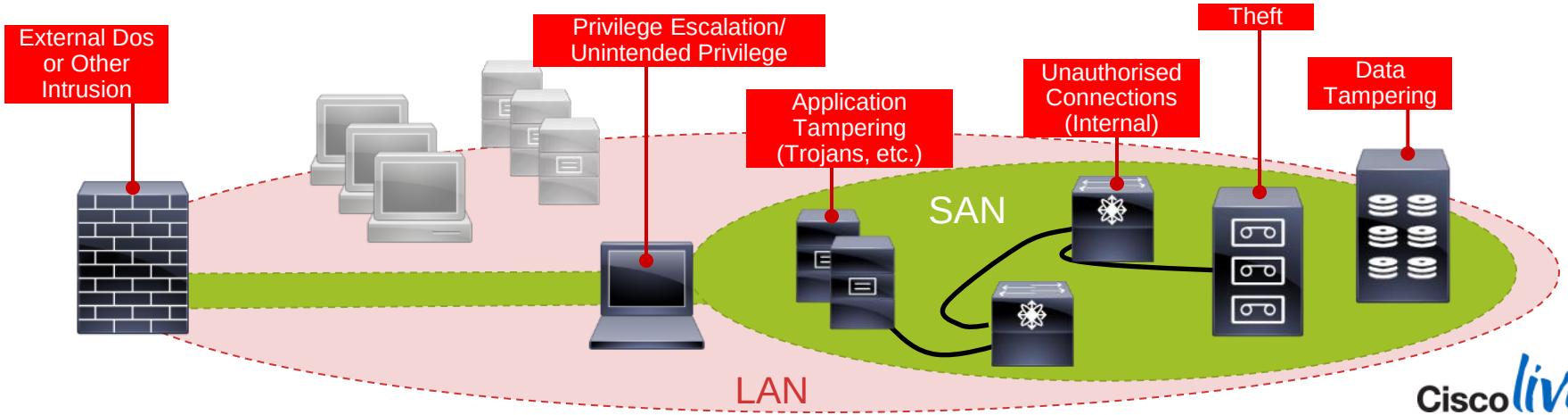
VMpath Analysis provides VM connectivity to network and storage across Unified Compute and Unified Fabric

- Visibility past physical access (switch) layer
- Standard & Custom Reports
- On Nexus and MDS platforms
- Dynamic Topology Views
- Rule-based event filtering and forwarding
- Threshold Alerting
- Integration via vCenter API



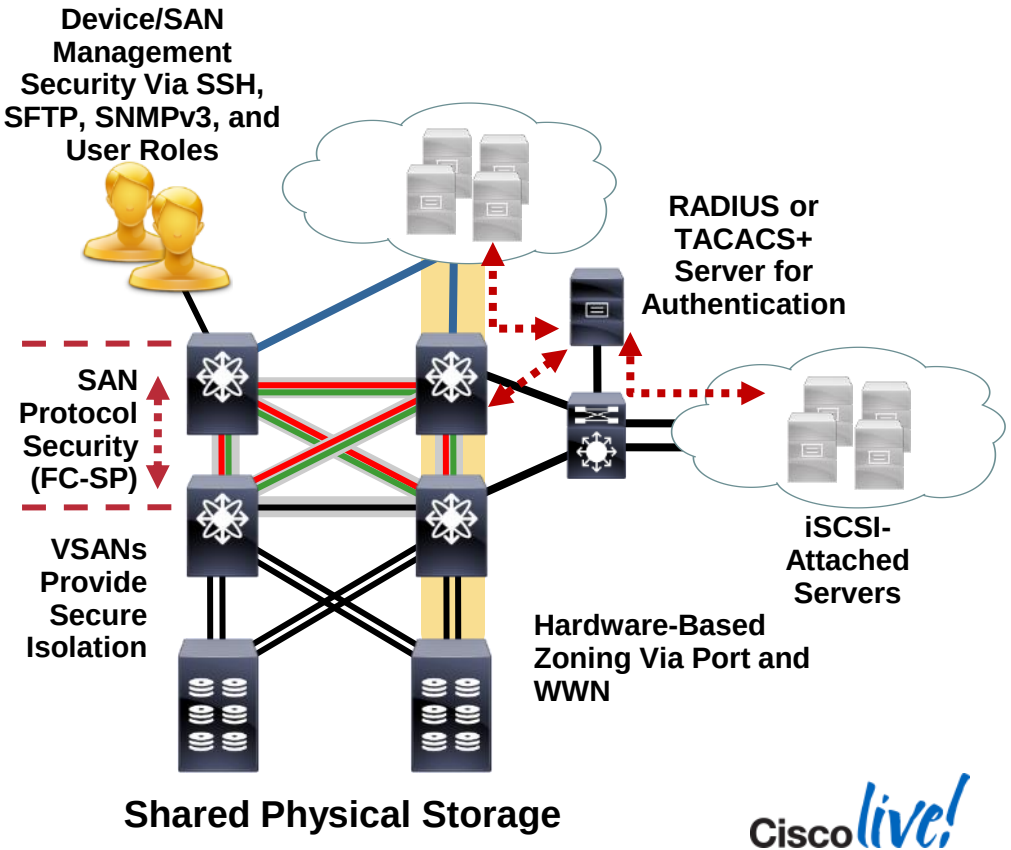
SAN Design Security Challenges

- SAN design security is often overlooked as an area of concern
  - Application integrity and security is addressed, but not back-end storage network carrying actual data
  - SAN extension solutions now push SANs outside datacentre boundaries
- Not all compromises are intentional
  - Accidental breaches can still have the same consequences
- SAN design security is only one part of complete data centre solution
  - Host access security—one-time passwords, auditing, VPNs
  - Storage security—data-at-rest encryption, LUN security



SAN Security

- Secure management access
  - Role-based access control
  - CLI, SNMP, and web access
- Secure management protocols
  - SSH, SFTP, and SNMPv3
- Secure switch control protocols
  - TrustSec
  - FC-SP (DH-CHAP)
- RADIUS AAA and TACACS+
  - User, switch and iSCSI host authentication
- Fabric Binding
  - Prevent unauthorised switches from joining the fabric



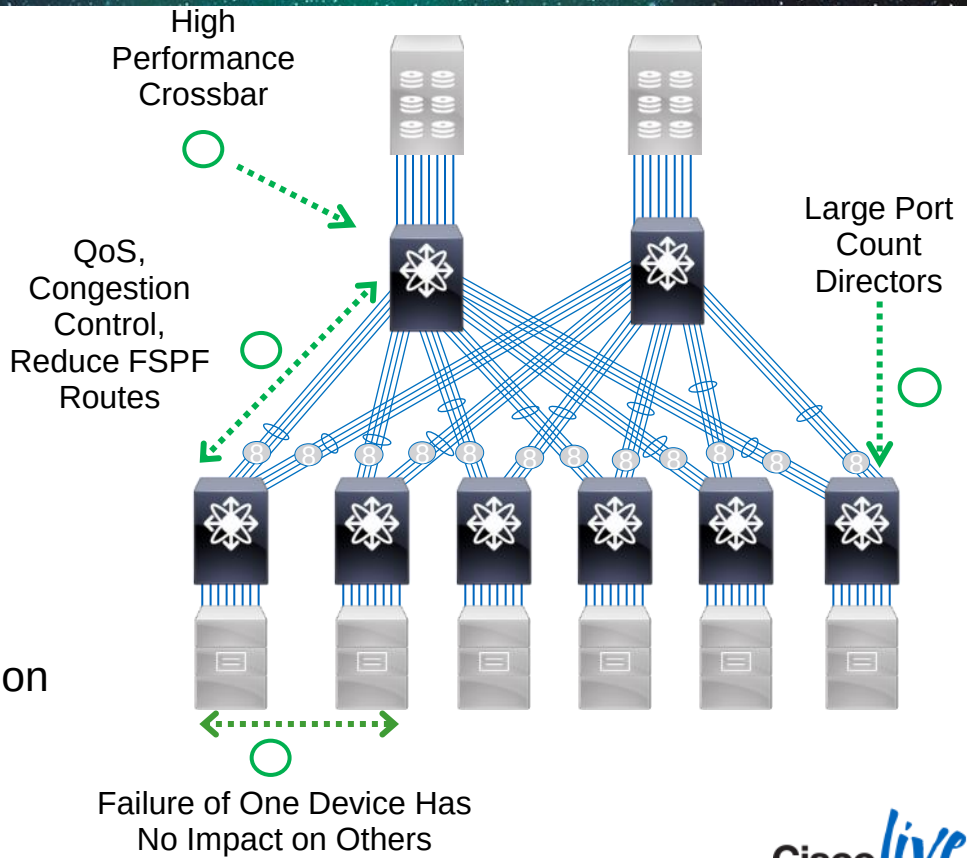
Storage Fabric Design Considerations

The Importance of “Architecture”

- SAN designs traditionally robust: dual fabrics, data loss is not tolerated
- Must manage ratios
  - Fan in/out
  - ISL oversubscription
  - Virtualised storage IO streams (NPIV attached devices, server RDM, LPARs, etc.)
  - Queue depth
- Latency
  - Initiator to target
  - Slow drain
  - Performance under load: does my fabric perform the same
- Application independence
  - Consistent fabric performance regardless of changes to SCSI profile
    - Number of frames
    - Frame size
    - Speed or throughput

SAN Major Design Factors

- Port density
  - How many now, how many later?
  - Topology to accommodate port requirements
- Network performance
  - What is acceptable? Unavoidable?
- Traffic management
  - Preferential routing or resource allocation
- Fault isolation
  - Consolidation while maintaining isolation
- Management
  - Secure, simplified management



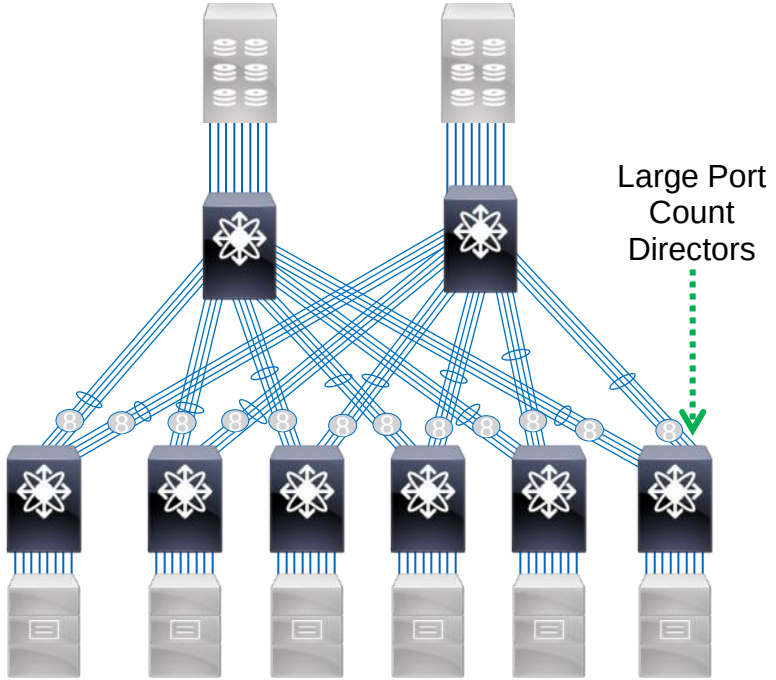
Scalability—Port Density
Topology Requirements

Considerations

- Number of ports for end devices
- How many ports are needed now?
- What is the expected life of the SAN?
- How many will be needed in the future?
- Hierarchical SAN design

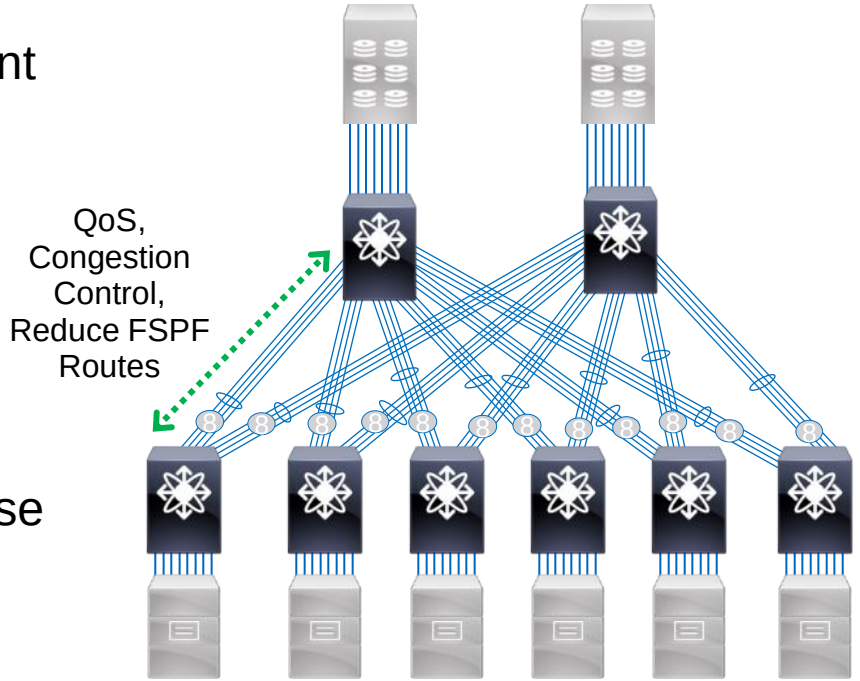
Best Practice

- Design to cater for future requirements
- Doesn't imply "build it all now," but means "cater for it" and avoids costly retrofits tomorrow



Traffic Management

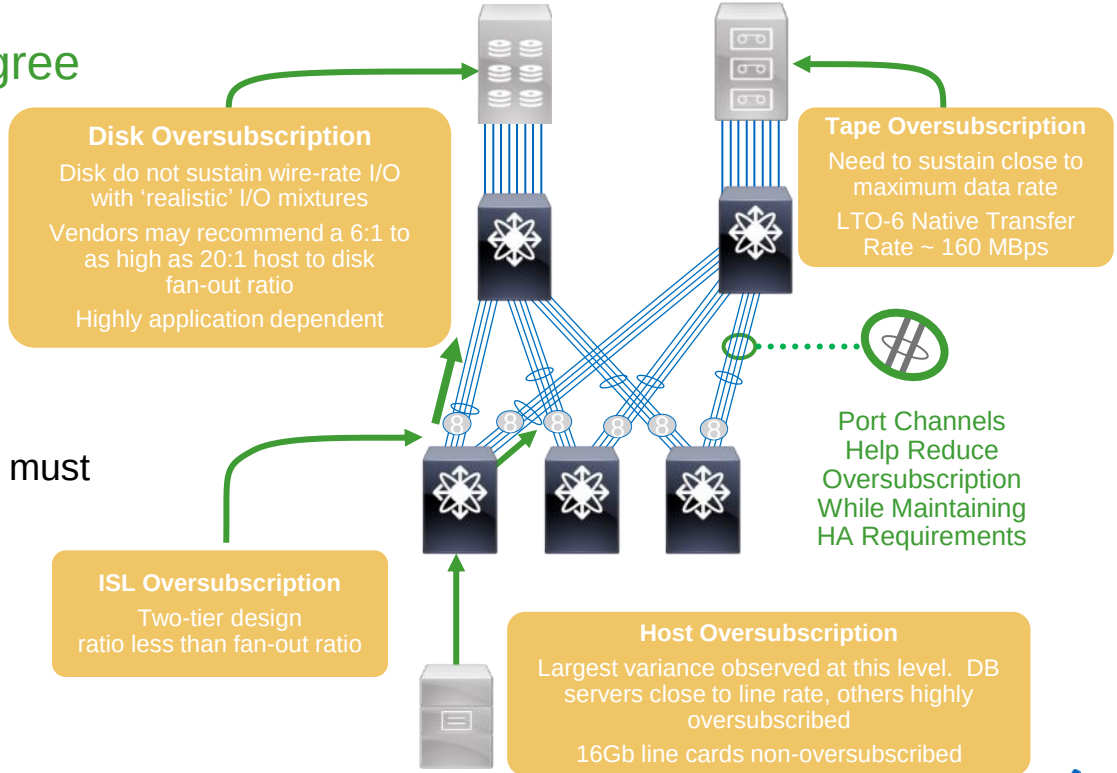
- Do different apps/servers have different performance requirements?
  - Should bandwidth be reserved for specific applications?
  - Is preferential treatment/ QoS necessary?
- Given two alternate paths for traffic between data centres, should traffic use one path in preference to the other?
  - Preferential routes



Network Performance
Oversubscription Design Considerations

All SAN Designs Have Some Degree of Oversubscription

- Without oversubscription, SANs would be too costly
- Oversubscription is introduced at multiple points
- Switches are rarely the bottleneck in SAN implementations
- Device capabilities (peak and sustained) must be considered along with network oversubscription
- Must consider oversubscription during a network failure event
- Remember, all traffic flows towards targets—main bottlenecks



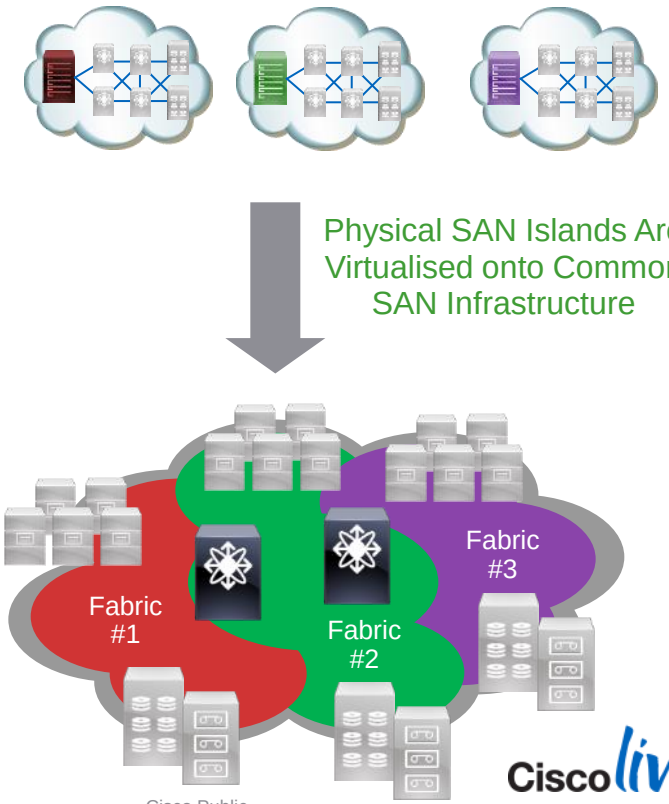
Fault Isolation

Consolidation of Storage

- Single Fabric = Increased Storage Utilisation + Reduced Administration Overhead

Major Drawback

- Faults Are No Longer Isolated
- Technologies such as VSANs enable consolidation and scalability while maintaining security and stability
- VSANs constrain fault impacts
- Faults in one virtual fabric (VSAN) are contained and do not impact other virtual fabrics



Dedicated Data Centre SAN Topologies

Data Centre Trends

The transition to 10G & Traffic Impact from Hosts

What is spurring the upgrade?

- Server Adoption
  - Intel's Romley shipped ~2X the volume of its predecessor within the same timeframe
- Efficiency and Cost Control
  - Scale VMs, LAN/SAN Convergence
  - Fewer NICs, cables, switch ports
  - 10G LAN on Motherboards (LOM)

What is the impact to SAN?

- Higher throughput from host
- Opportunity for consolidation with FCoE

" 10GE server attach-rates growing from 30% in 2012 to 90% in 2014" - Oppenheimer

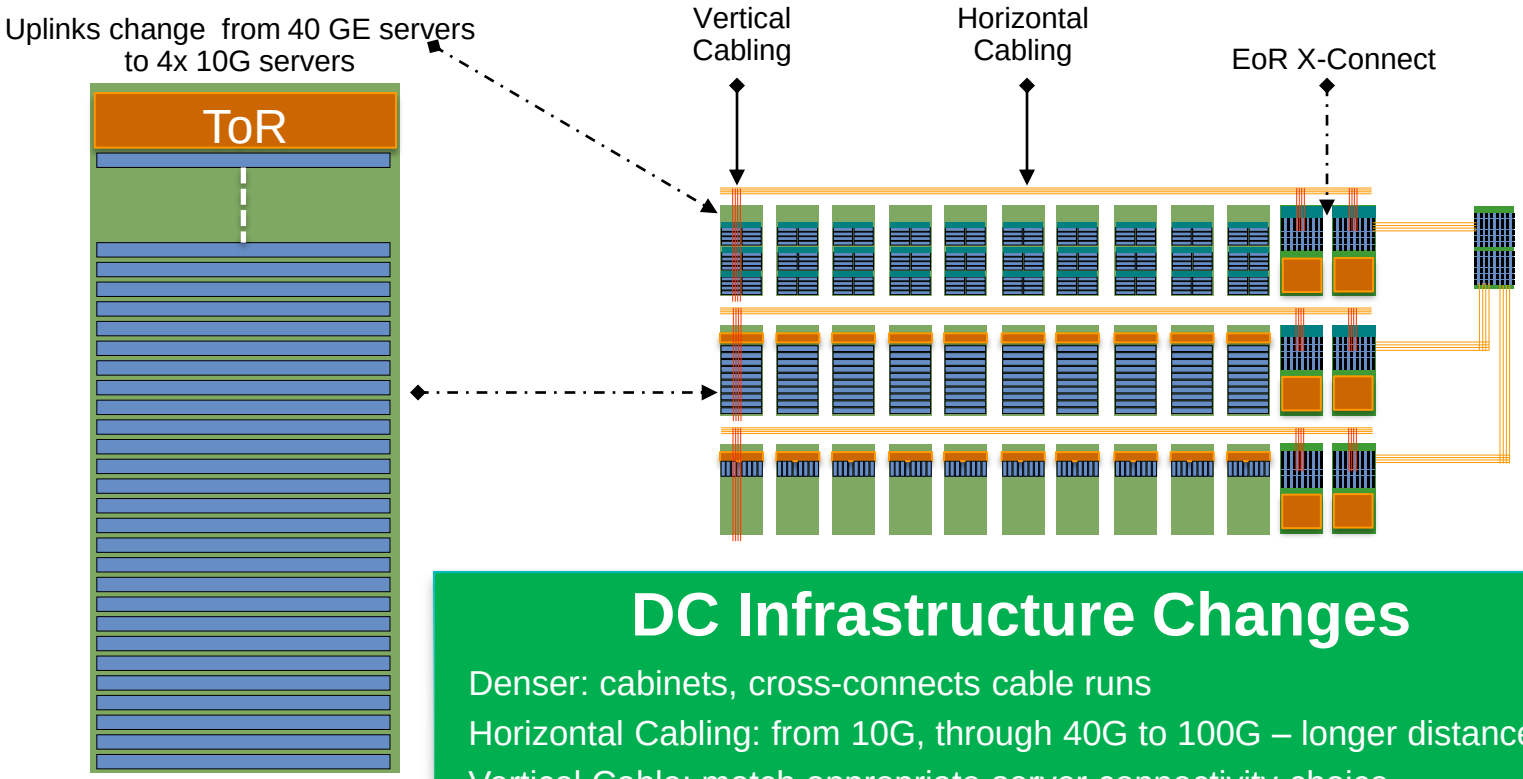
"...expect the Ethernet switch market to see a significant boost, doubling its (year-over-year) growth rates in 2013/14"

"10 GE port shipments to increase from 8.5M in 2011 to 62.2 M in 2016. Revenues increase from $6.1 B in 2011 to $13.2 B in 2016" – Dell'Oro

"40G/100G Ethernet to be a $3 billion market by 2016." – Dell'Oro

Denser Server Cabinets

What are the Implications?



DC Infrastructure Changes

Denser: cabinets, cross-connects cable runs

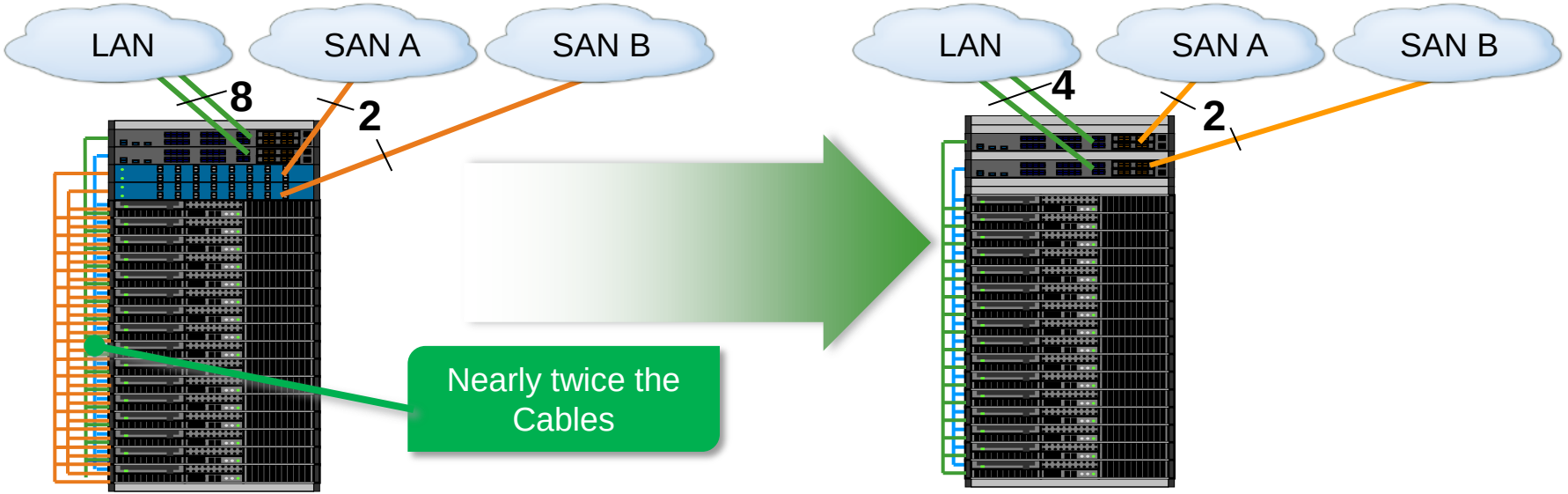
Horizontal Cabling: from 10G, through 40G to 100G – longer distances

Vertical Cable: match appropriate server connectivity choice

Is SAN EoR economical now?

Top-of-Rack Consolidated I/O

FCoE with I/O Consolidation at Access

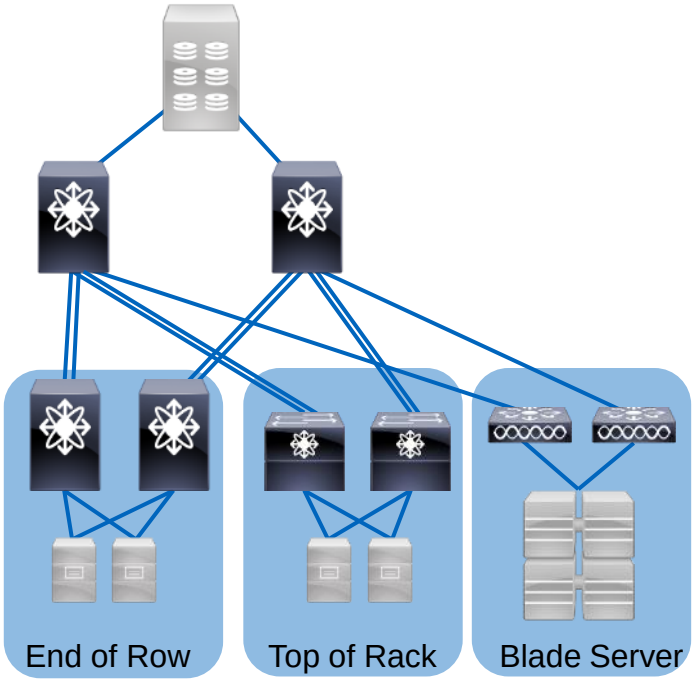


10 Servers	Enet	FC	Total
Adapters	20	20	40
Switches	2	2	4
Cables	40	40	80
Mgmt Pts	2	2	4

10 Servers	Enet	FC	Total
Adapters	20	0	20
Switches	2	0	2
Cables	40	0	40
Mgmt Pts	2	0	2

Core-Edge

Highly Scalable Network Design



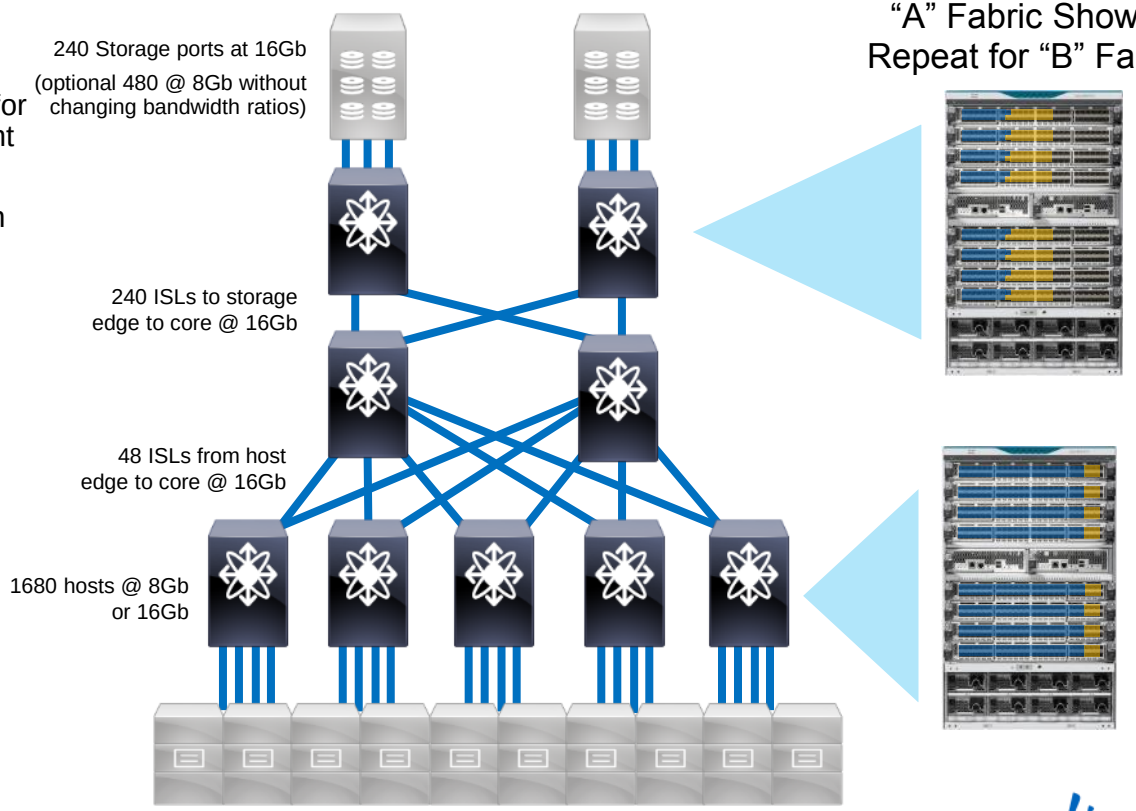
- Traditional SAN design for growing SANs
- High density directors in core and fabric switches, directors or blade switches on edge
- Predictable performance
- Scalable growth up to core and ISL capacity
- Evolves to support EoR & ToR

Large Edge-Core-Edge/End-of-Row Design

Large Edge/Core/Edge
(3456 Usable Ports per Fabric)

- Traditional Edge-Core-Edge design Is ideal for very large centralised services and consistent host-disk performance regardless of location
- Full line rate ports, no fabric oversubscription
- 8Gb or 16Gb hosts and targets
- Services consolidated in the core
- Easy expansion
- Massive cabling if used for EoR designs

Ports Deployed	**6,912**
Used Ports	**5,760 @ 16Gb 6,240 @ 8Gb**
Storage Ports	**480 @ 16Gb, or 960 @ 8Gb**
Host Ports	**3360**
Host ISL Oversubscription	**7:1 @ 16Gb**
End to End Oversubscription	**7:1 @ 16Gb storage 7:1 @ 8Gb storage**

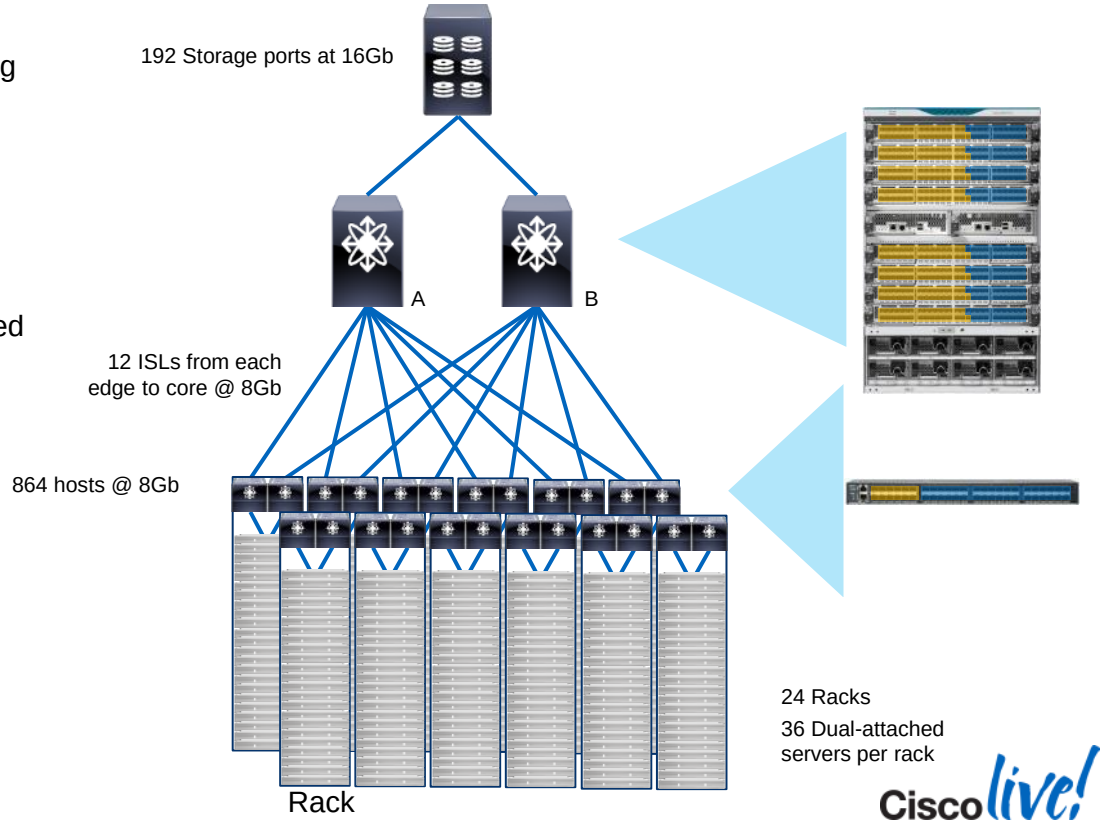


SAN Top of Rack – MDS 9148

SAN Top of Rack
(3072 Usable Ports per Fabric)

- Ideal for centralised services while reducing cabling requirements
- Consistent host/target performance regardless of location in rack
- 8Gb hosts & 16Gb targets (if they exist)
- Easy edge expansion
- Massive cabling infrastructure avoided as compared to EoR designs
- Additional efficiencies with in rack IO convergence

Ports Deployed	**3,072**
Used Ports	**192 @ 16Gb 2,880 @ 8Gb**
Storage Ports	**192 @ 16Gb, or 192 @ 8Gb**
Host Ports	**1,920**
Host ISL Oversubscription	**3:1 @ 8Gb**
End to End Oversubscription	**4.5:1 @ 16G storage 9:1 @ 8G storage**



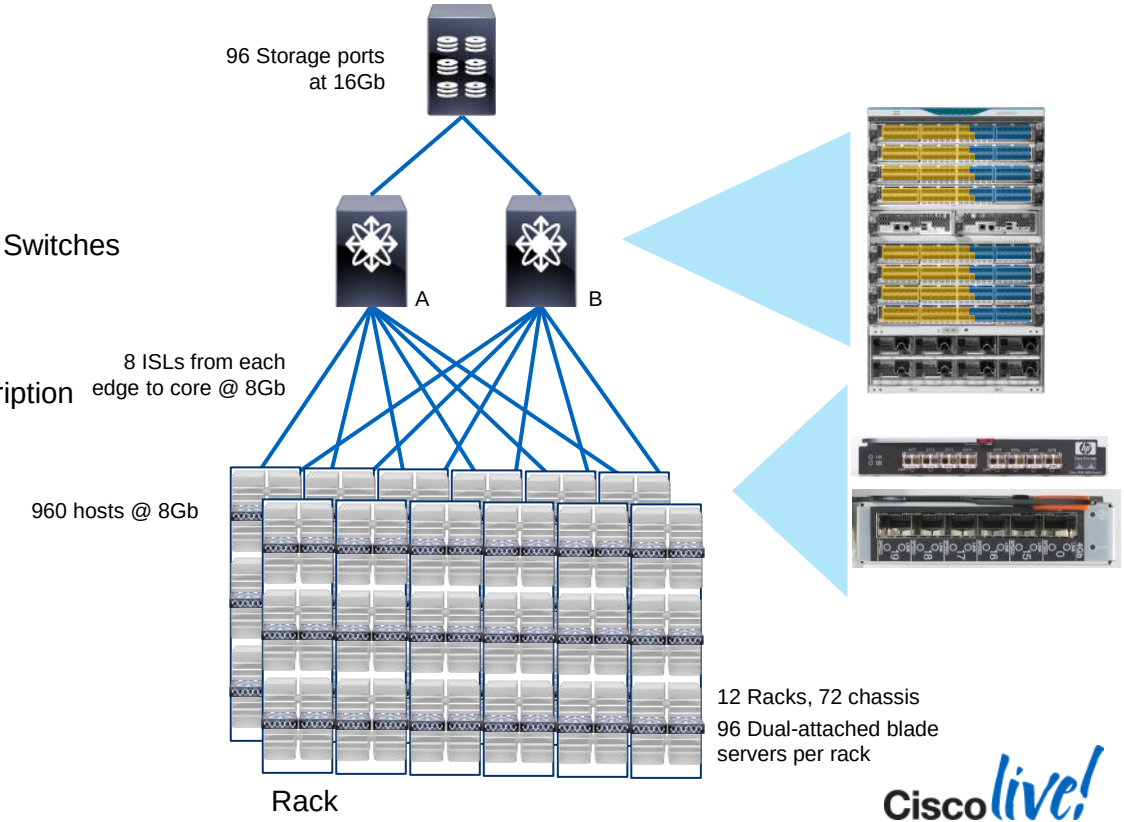
Top-of-Rack Design - Blade Centres

SAN Top of Rack – Blade Centres

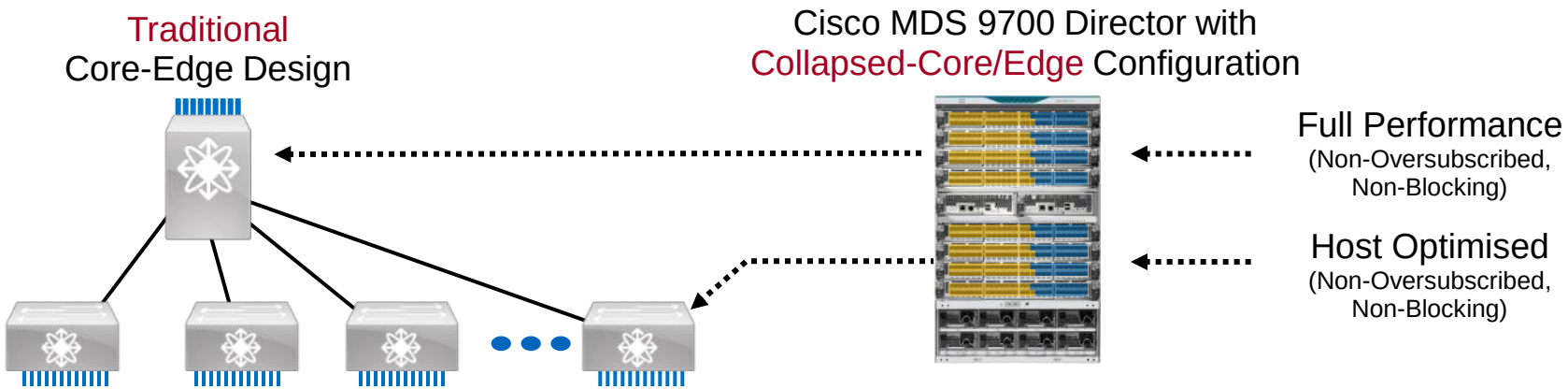
(1,920 Usable Ports per Fabric)

- Ideal for centralised services
- Consistent host/target performance regardless of location in blade enclosure or rack
- 8Gb hosts & 16Gb targets (if they exist)
- Need to manage more SAN Edge switches/Blade Switches
- NPV attachment reduces fabric complexity
- Assumes little east-west SAN traffic
- Add blade server ISLs to reduce fabric oversubscription

Ports Deployed	**1,920**
Used Ports	**192 @ 16Gb** **1056 @ 8Gb**
Storage Ports	**192 @ 16Gb, or** **192 @ 8Gb**
Host Ports	**2304**
Host ISL Oversubscription	**4:1 @ 8G**
End to End Oversubscription	**6:1 @ 16Gb Storage** **12:1 @ 8Gb Storage**



Collapsed Core/Edge Design



Collapsed Core/Edge

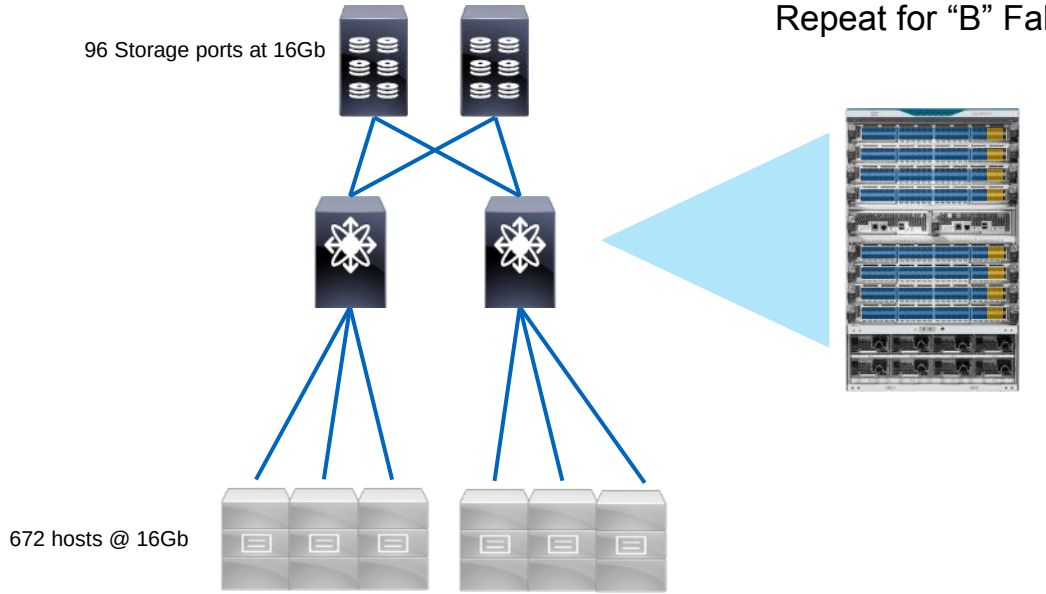
- Typically a lower oversubscription ratio
- Room to grow—empty slots = future port count growth
- While director ports are more expensive than fabric switch ports, collapsed core/edge design has no wasted ports for ISLs—similar cost/usable port
- More important in previous MDS generations with oversubscribed linecards

Medium Scale Dual Fabric
Collapsed Core/Edge Design

Medium Scale Dual Fabric
(768 Usable Ports per Fabric)

- Ideal for centralised services
- Consistent host/target performance regardless of location
- 8Gb or 16Gb hosts & targets (if they exist)
- Relatively easy edge expansion to Core/Edge
- EoR design
- Supports blade centres connectivity

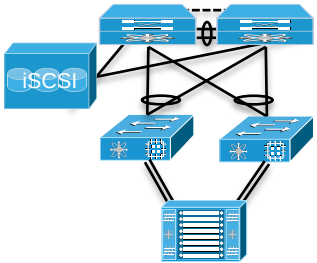
Ports Deployed	**768**
Used Ports	**768@ 16Gb**
Storage Ports	**96 @ 16Gb**
Host Ports	**672 @ 16Gb**
Host ISL Oversubscription	**N/A**
End to End Oversubscription	**7:1 @ 16Gb**

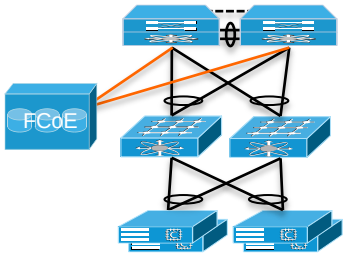


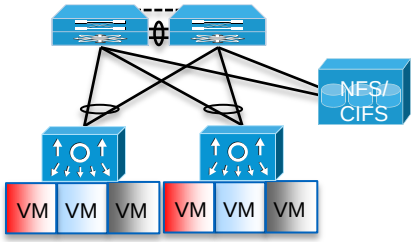
Converged Data Centre SAN Topologies

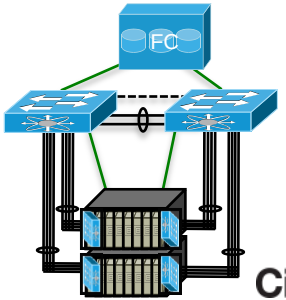
How do I Support Everything !

- Form Factor
  - Unified Computing Fabric
  - 3rd Party Blade Servers
  - Rack Servers (Non-UCSM)
- Storage Protocols
  - Fibre Channel
  - FCoE
  - iSCSI, NFS
- Virtualisation Requirements
  - vSwitch/DVS
  - Nexus 1000v
  - VM-FEX HW Switching
- NIC Connectivity Model
  - 10 or 1-GigE Server ports
  - Physical Interfaces per server
  - NIC Teaming models



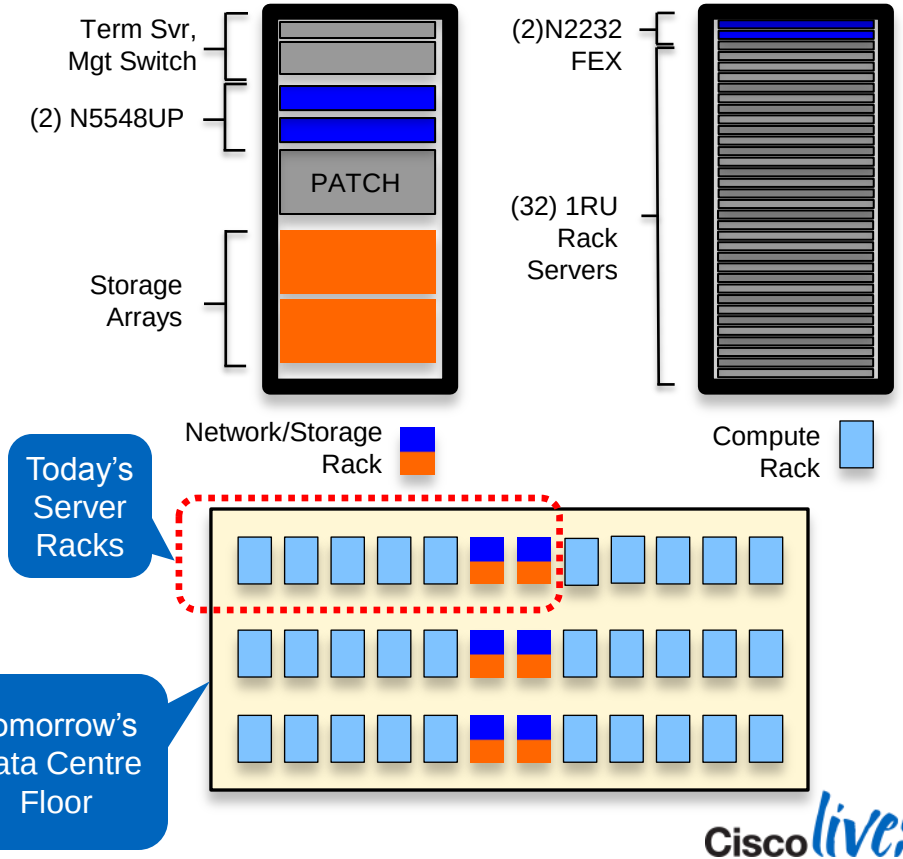






Converged Data Centre Access Requirements

- Plan for growth in a modular, pod-based repeatable fashion.
- Your own “pod” will be based on compute, network, or storage requirements.
- Consider future topology needs eg. FabricPath
- Map physical Data Centre needs to a flexible communication topology.
- Use MDS Storage Services to interconnect with legacy environments
- Manage service delivery with QoS and CoS enforcement



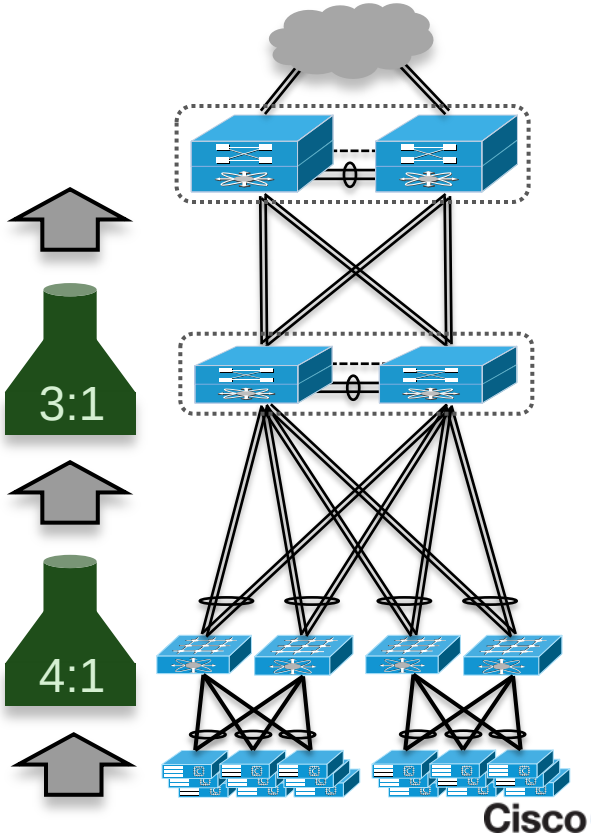
Oversubscription & Resource Utilisation

Oversubscription:

- Most servers will not be consistently filling a 10 GigE interface.
- A switch may be line-rate non-blocking, but still introduce oversubscription into an overall topology by design.
- Consider Ethernet-based storage traffic when planning ratios, don't be overly aggressive.

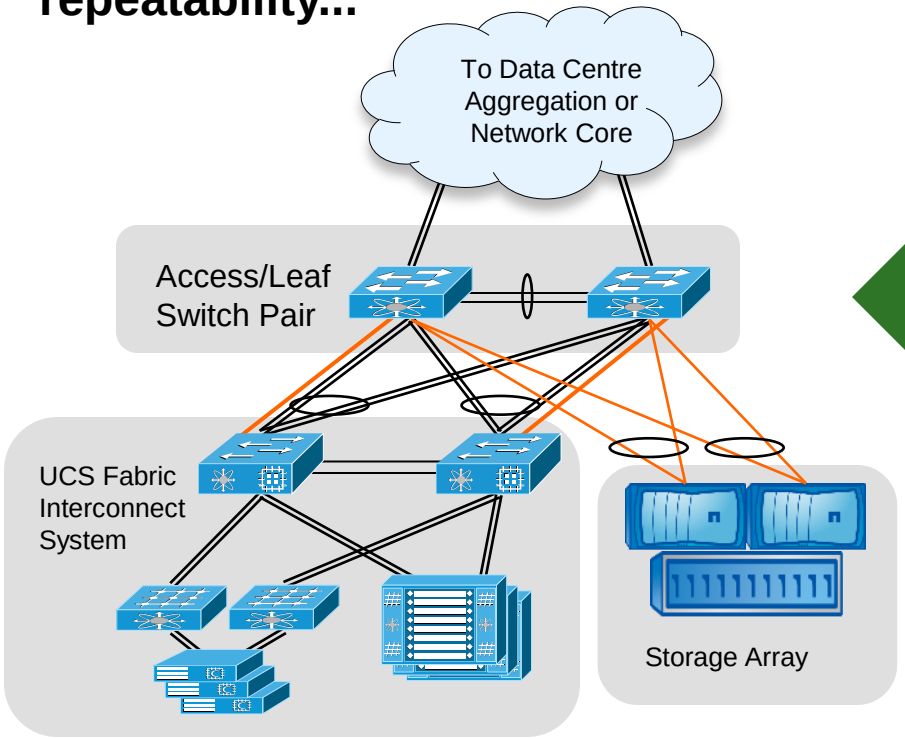
Resource utilisation:

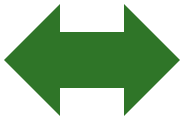
- Consider in Converged Pod architectures logical limits
- These include MAC address counts, Zone Member counts
- Remember every feature you turn on consumes resources
- In heavily virtualised environments expect logical resource contention to be exceeded **first**

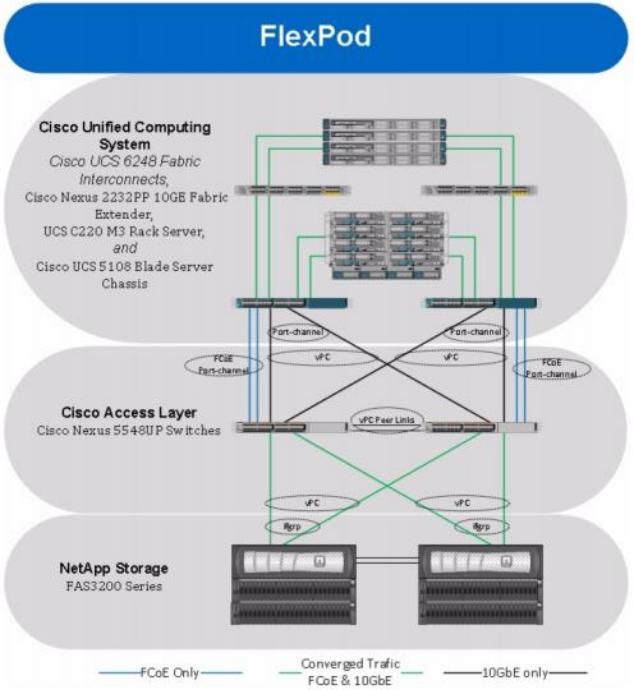


Pod or Block Style Architectures

Block architectures increase repeatability...



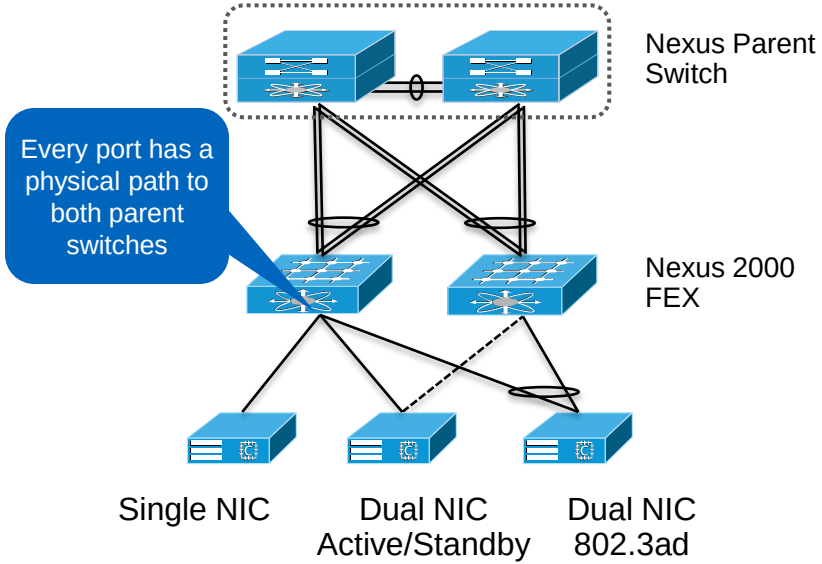




Converged Access Pod : Enhanced vPC (EvPC)

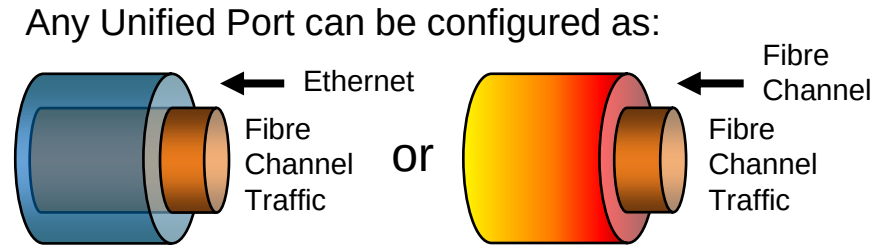
- In an Enhanced vPC configuration any server NIC teaming configuration will be supported on any port. No ‘orphan ports’ in the design.
- All components in the network path are fully redundant.
- Supported dual-homed FEX parent switch is Nexus 6000 or 5500.
- Provides flexibility to mix all three server NIC configurations (single NIC, Active/Standby and NIC Port Channel).

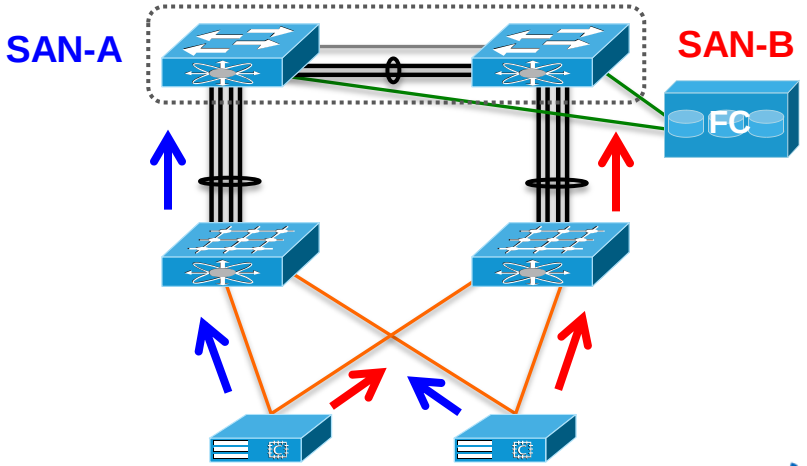
Note, Port Channel to active/active server is standard port channel, not configured as “vPC”.



Access Pod Features: Unified Ports and FCoE

- Unified Port allows a physical port to be configured to support either native Fibre Channel or Ethernet.
- SFP+ optic needs to be chosen to support the setting of the port
- Fibre Channel over Ethernet (FCoE) allows encapsulation and transport of Fibre Channel traffic over a shared Ethernet network
- Traffic may be extended over Multi-Hop FCoE, or directed to an FC SAN
- SAN “A” / “B” isolation is maintained across the network

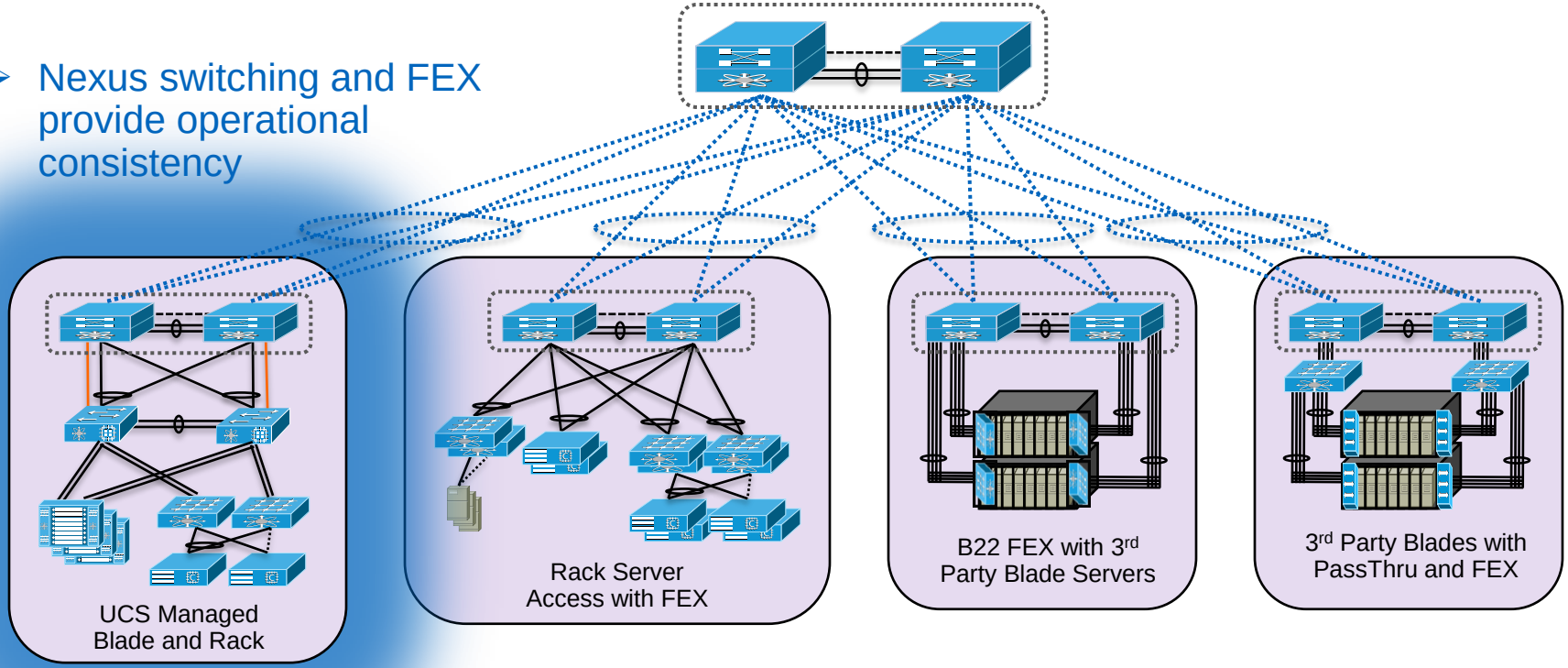




Pod Architecture Expansion Options

Mix and Match Layer-2 Compute Connectivity for Migration or Scale Requirements

- Nexus switching and FEX provide operational consistency



More features, highest value and physical consolidation

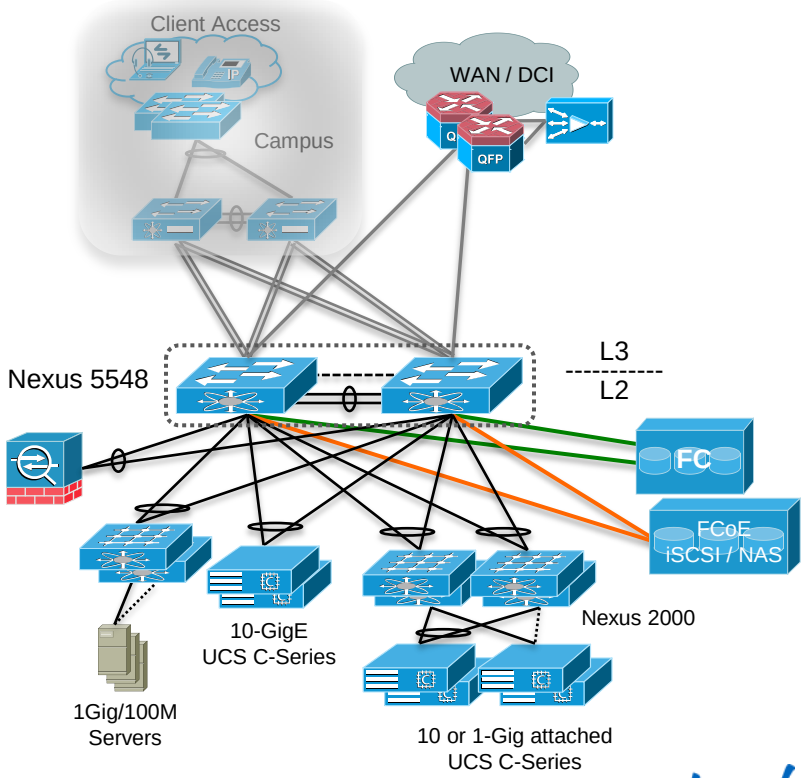
Small DC Converged Infrastructure Example

Dedicated Nexus 5548-based DC Switch Pair (or 5596 for higher port count)

- Unified Ports cater for native Fibre Channel, FCoE, iSCSI or NAS Storage.
- Non-blocking, line-rate 10Gpbs Layer-2 switching with low latency.
- CoS / QoS used to ensure storage priority
- Nexus 5500 supports physical FEX, Adapter-FEX, VM-FEX capabilities.

Notes:

- ISSU not supported with Layer-3 module in 5500 switches
- Tape/legacy support through use of FC



Single Layer Data Centre, Nexus 7004

Highly Available Virtualised Chassis Access/Aggregation Model

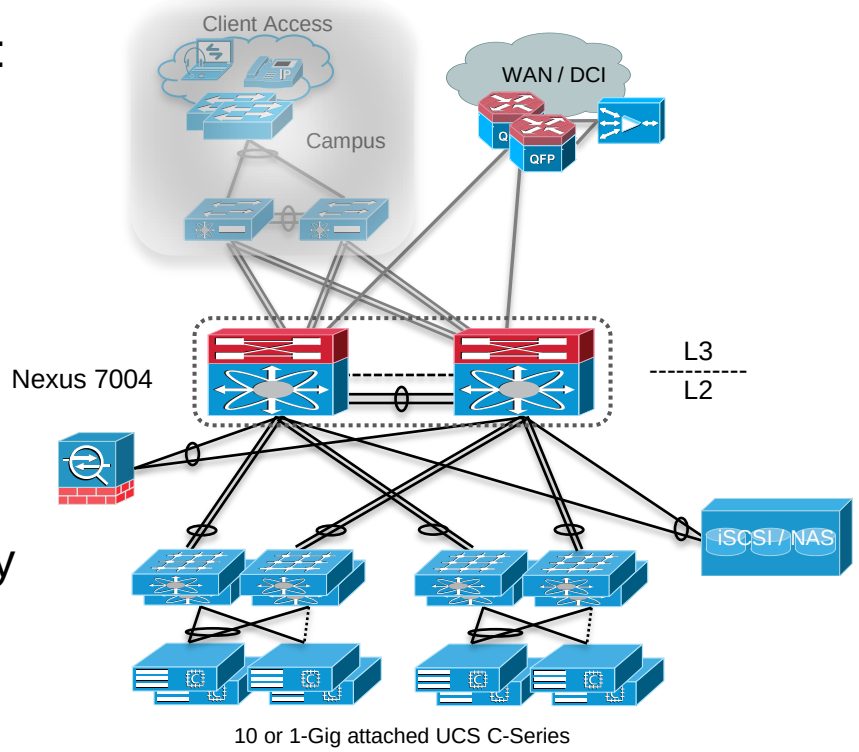
Benefits of Nexus 7004 with F2 I/O Modules:

- Supervisor High Availability
- Layer-2/3 ISSU
- Virtual Device Contexts (VDC)
- FabricPath support for future expansion

Benefits:

- Supports all key DC features with higher availability
- Allows for scaling of L3 and L2 separately
- Can Support FC via attached MDS or Unified port device

Note: For native Fibre Channel or FCoE add Nexus 5500 access layer or MDS SAN.



Dual Layer Nexus 7700/6000 Fabric with VDCs

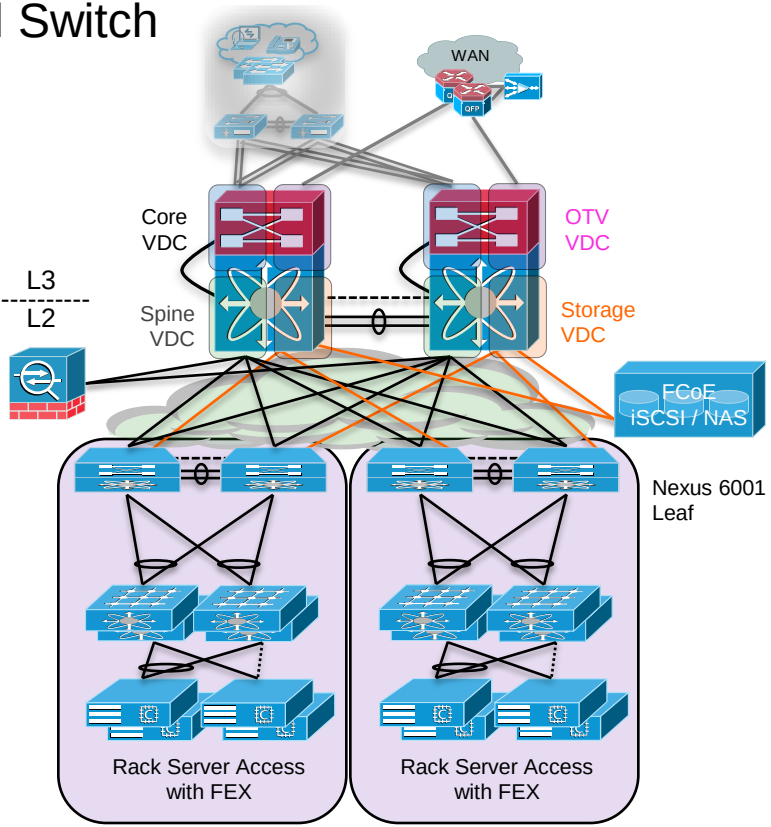
Virtual Device Contexts Partitioning the Physical Switch

Nexus 7700 series FabricPath Spine

- Highly Available design with dual-supervisor
- Add leaf pairs for greater end node connectivity
- Add spine nodes for greater fabric scale and HA
- FCoE support over dedicated links and VDC

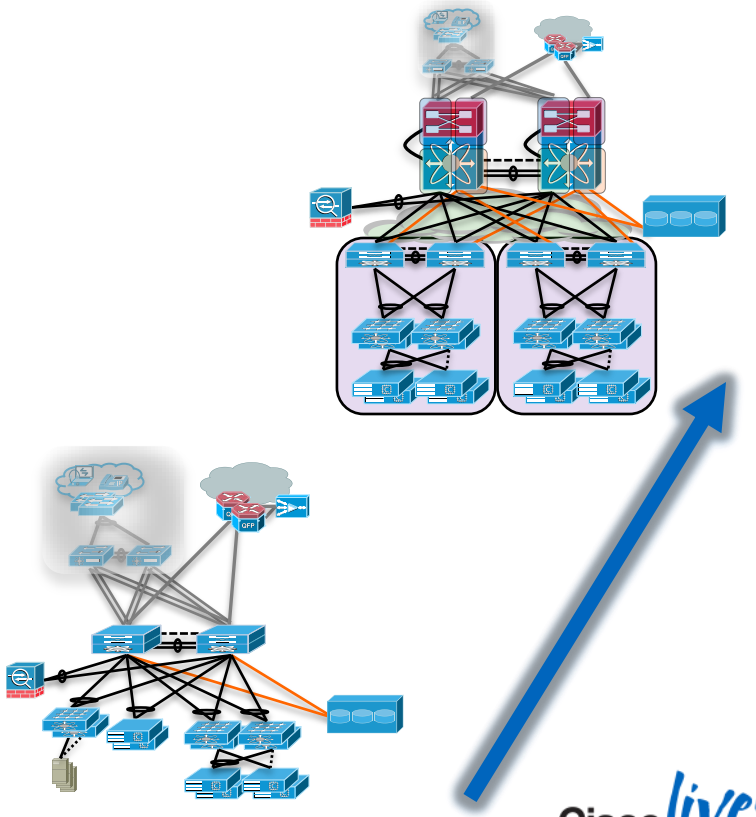
Benefits:

- Integrated DCI support with OTV, LISP, and MPLS
- Feature-rich switching fabric with VDC's, FEX, vPC, FabricPath, FCoE
- Investment protection of a chassis-based solution
- All director class topology allows you to reach 99.999% reliability



Summary: Converged Access Data Centre Designs

- Start with a structured approach that allows modular design as requirements grow.
- Evaluate Nexus switching options based on feature support, scale, and performance.
- Use Unified Port devices to support legacy workloads
- Consider using MDS storage services to interconnect between new converged pods and legacy DC components
- Plan ahead for re-use of components in new roles as needs change.

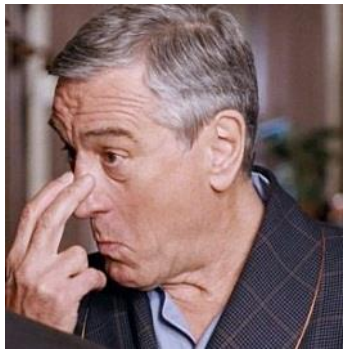


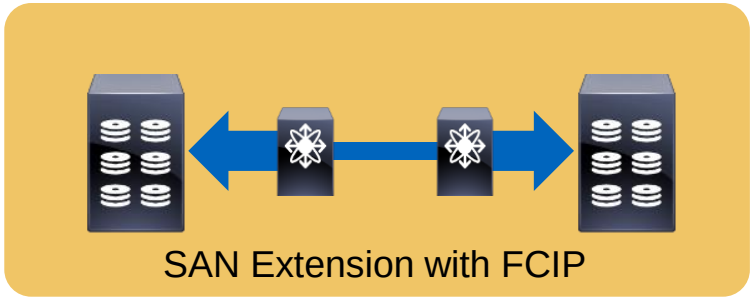


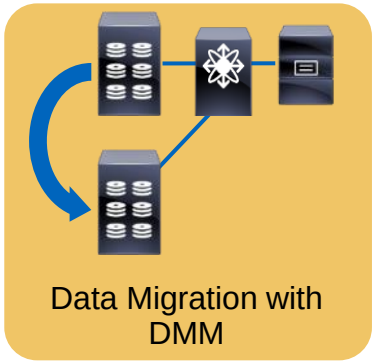

Distance Extension

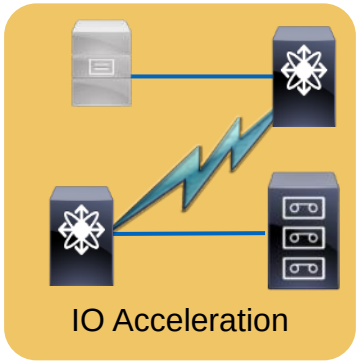
Enhancing SAN Design with Services

- Extend Fabrics
  - FCIP
  - Extended Buffer to Buffer credits
  - Encrypt the pipe
  - Efficiently use optical networks
- SAN Services extend the effective distance for remote applications
  - SAN IO acceleration
  - Write acceleration
  - Tape acceleration
- Enhance array replication requirements
- Reduces WAN-induced latency
- Improves application performance over distance
- Data Migration







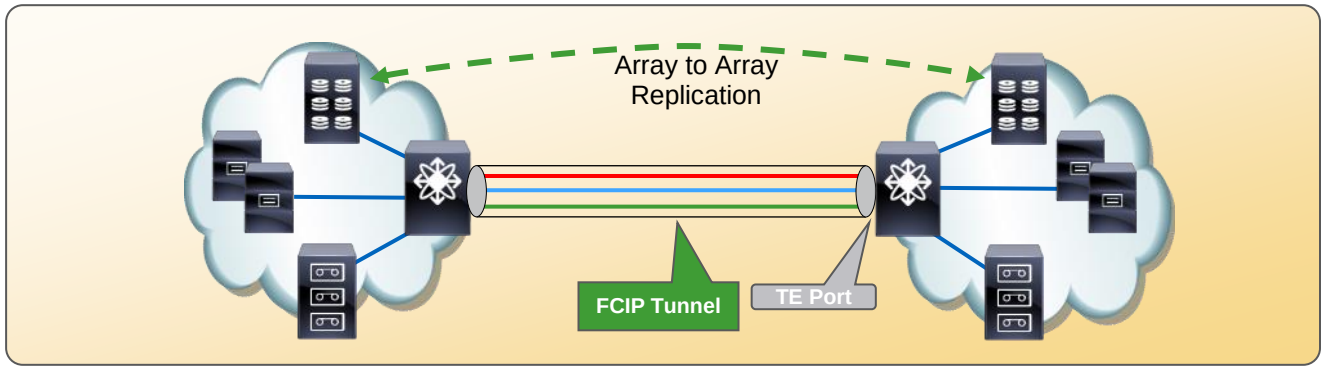


Fabric is aware of all data frames from initiator to target

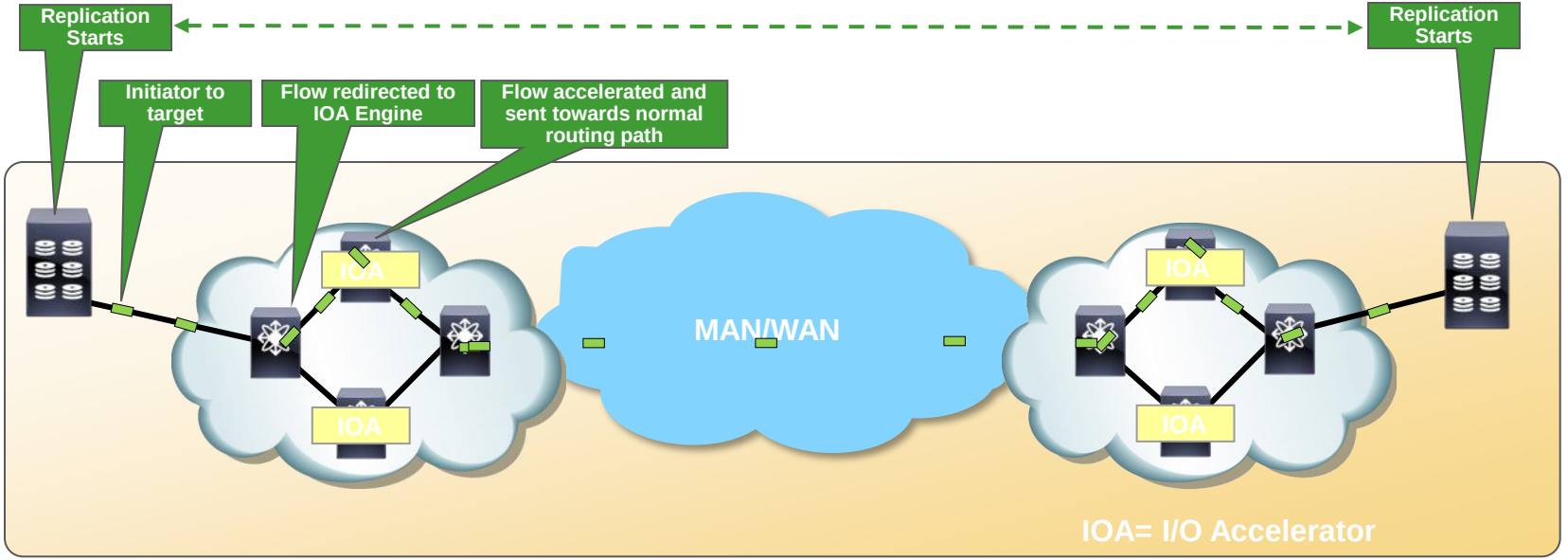
SAN Extension with FCIP

Fibre Channel over IP

- Encapsulation of Fibre Channel frames into IP packets and tunnelling through an existing TCP/IP network infrastructure, in order to connect geographically distant islands
- Write Acceleration to improve throughput and latency
- Hardware-based compression
- Hardware-based IPSec encryption



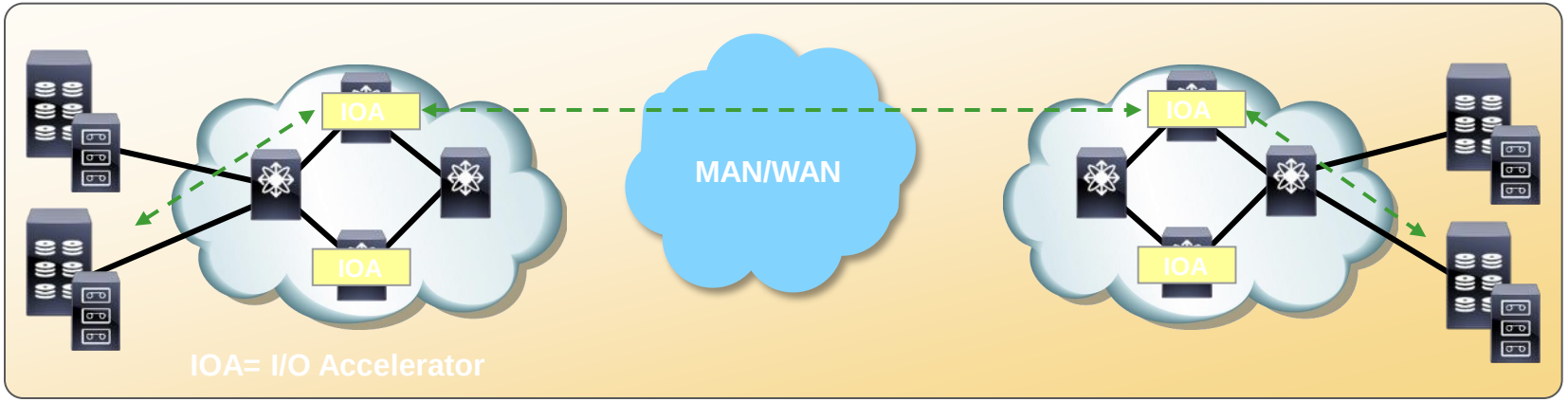
FC Redirect - How IOA Works



Data Acceleration

A Fabric Service to Accelerate I/O Between SAN Devices

- Accelerate SCSI I/O
  - Over both Fibre Channel (FC) and Fibre Channel over IP (FCIP) links
  - For both Write Acceleration (WA) and Tape Acceleration (TA)
- I/O Acceleration Node platforms: MSM-18/4, SSN-16, MDS-9222i, MDS-9250i
- Uses FC Redirect



IOA FCIP Tape Backup

Large Health Insurance Firm

- Highly resilient– Clustering of IOA engines allows for load balancing and failover
- Improved Scalability- Scale without increasing management overhead
- Significant re-utilisation of existing infrastructure- All chassis and common equipment re-utilised
- Flat VSAN topology- Simple capacity and availability planning

Summary

- Drivers in DC are forcing change
  - 10G convergence & server virtualisation
  - It's not just about FCP anymore. FCoE, NFS, iSCSI are being adopted
- Proper SAN design is holistic in the approach
  - Performance, Scale, Management attributes all play critical roles
  - Not all security issues are external
  - Fault isolation goes beyond SAN A/B separation
  - Consider performance under load
  - Design for SAN services
- Many design options
  - Optimised for performance
  - Optimised for data centre space




Q & A

Complete Your Online Session Evaluation

Give us your feedback and receive a Cisco Live 2014 Polo Shirt!

Complete your Overall Event Survey and 5 Session Evaluations.

- Directly from your mobile device on the Cisco Live Mobile App
- By visiting the Cisco Live Mobile Site www.ciscoliveaustralia.com/mobile
- Visit any Cisco Live Internet Station located throughout the venue

Polo Shirts can be collected in the World of Solutions on Friday 21 March 12:00pm - 2:00pm

Learn online with Cisco Live!

Visit us online after the conference for full access to session videos and presentations.
www.CiscoLiveAPAC.com


CISCO
TM