



Cisco Expo
2008

Data Center:

DC Trends towards the
Next Generation
Ethernet



Wolfgang Riedel
Consulting System Engineer
Technical Marketing - Data Center
CCIE #13804
wr@cisco.com



- Environmental Challenges
- What's up in the Datacenter
- DC Architecture Trends
- Fibre Channel compared to Ethernet
- I/O Consolidation
- Data Center Ethernet (DCE)
- Fibre Channel over Ethernet (FCoE)
- Unified I/O Use Case
- Summary

The Journey from Coal to the Server....

Cost = €1M per MW Year
Energy Efficiency ~ 5%

**33.2 MW
Coal Energy
Input**

**5-10%
distribution loss**

**10 MW
to the
Data Centre**

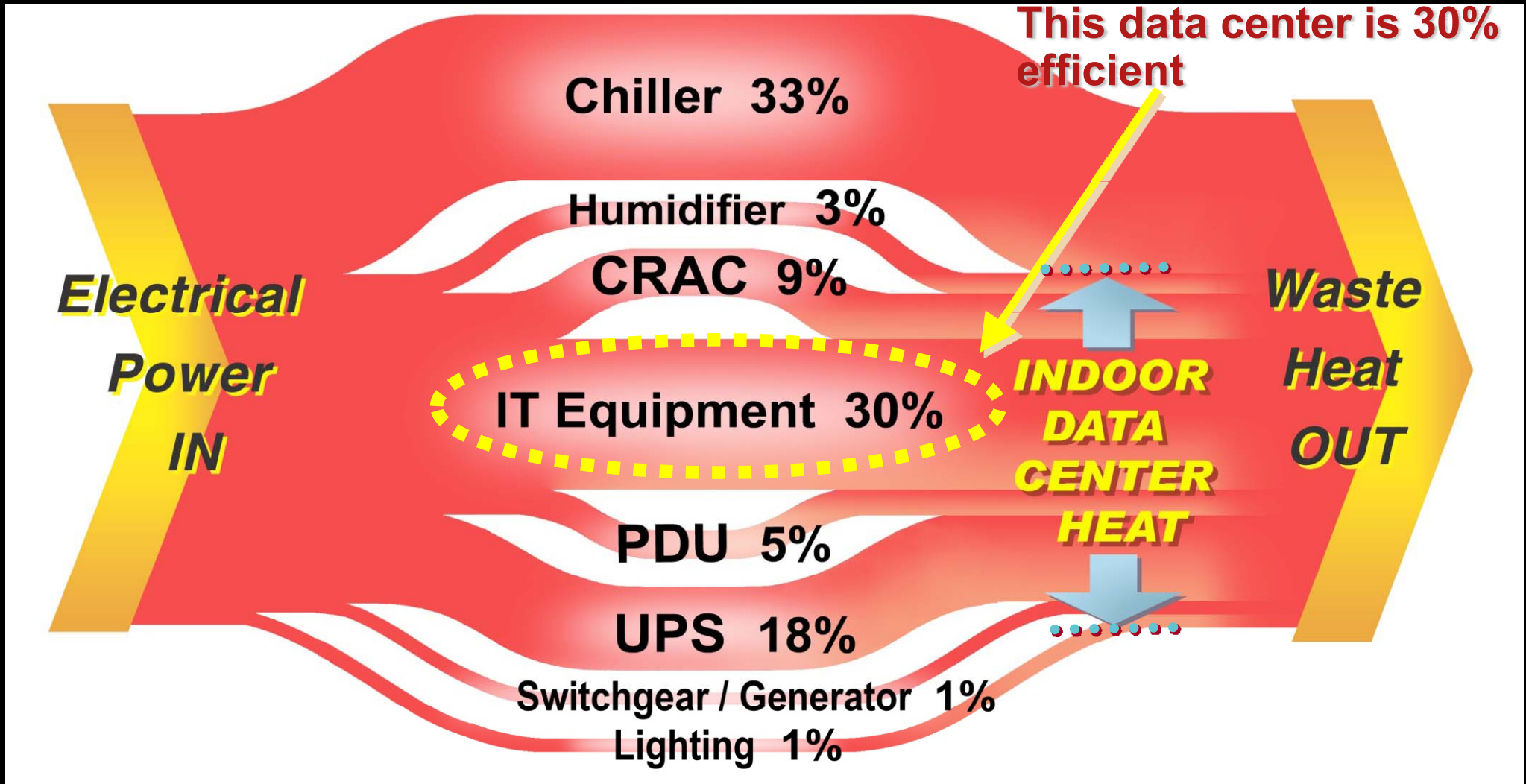
**65% DC power
Conversion and cooling
loss**

**0.4-0.8 MW
“Useful Server
Cycles”**

**60% Lost
through chimney**

**15-30% Server and
Storage
utilisation**

Where does the power go?

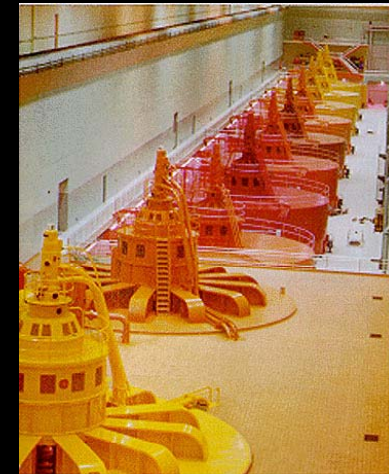
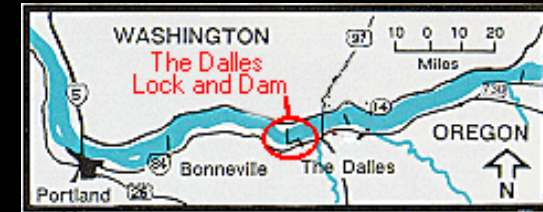


Load: 30% design capacity

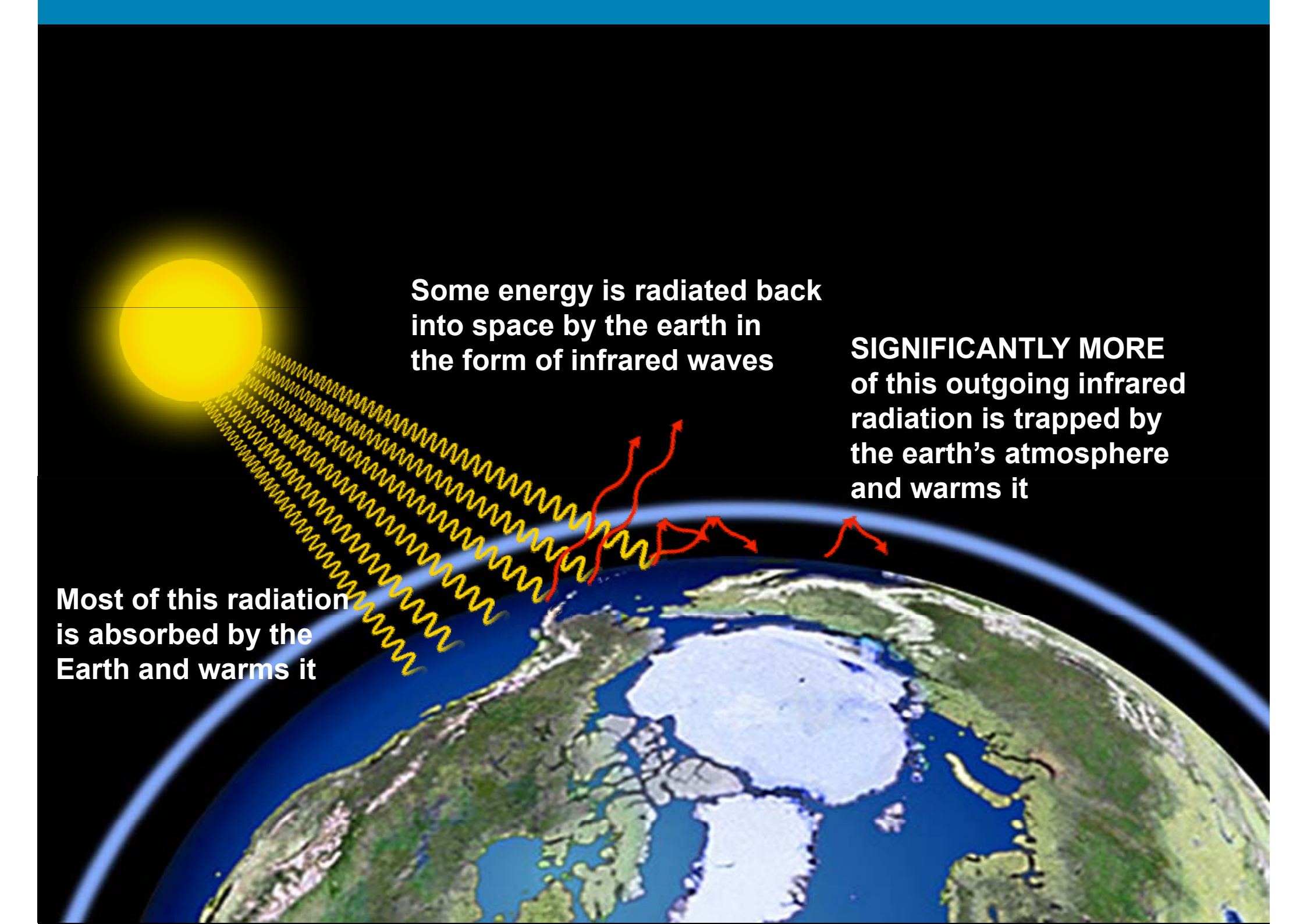
CRACS: N+1 redundancy



The Green Data Centre Going Hydro???



Google has two computing centers, top and left, each the size of a football field, @ Columbia River, The Dalles, Oregon



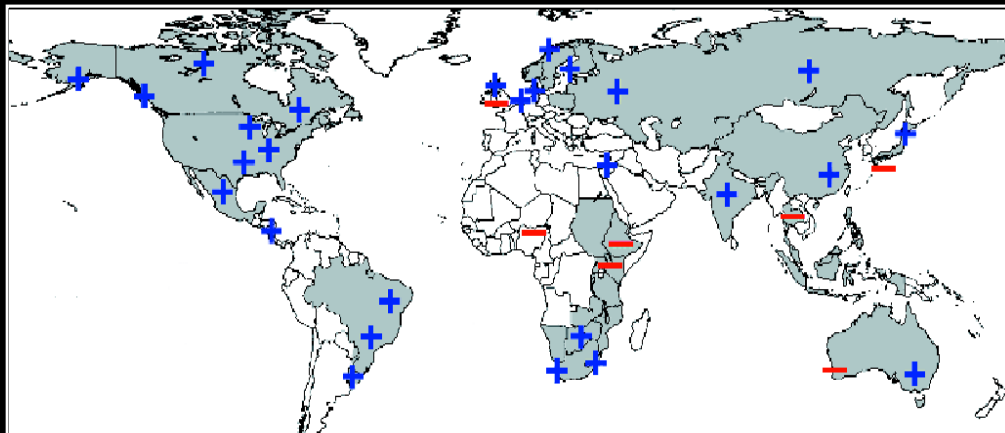
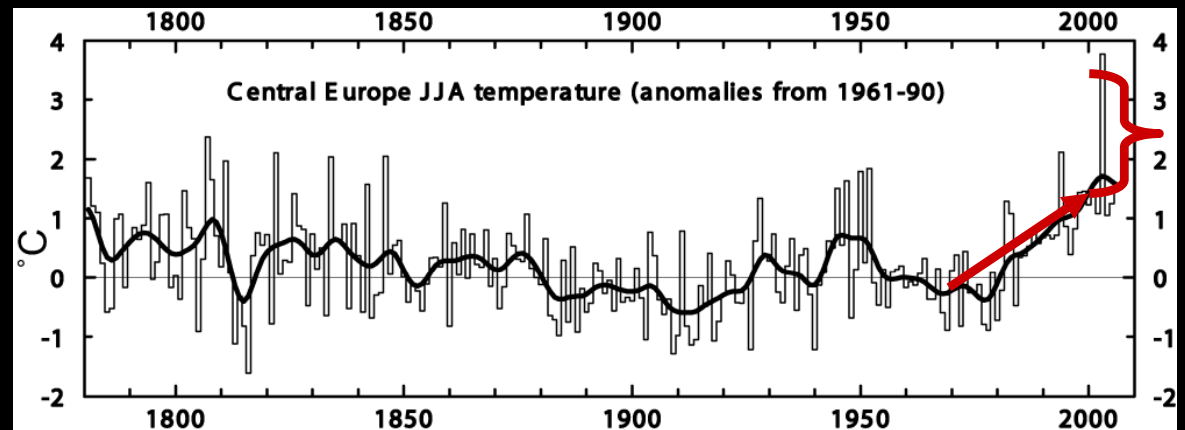
Some energy is radiated back into space by the earth in the form of infrared waves

SIGNIFICANTLY MORE of this outgoing infrared radiation is trapped by the earth's atmosphere and warms it

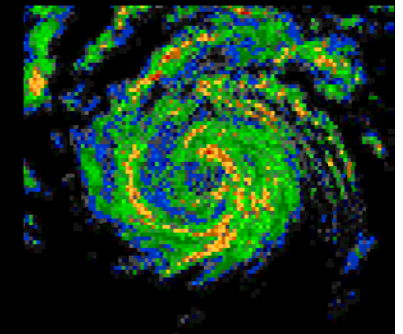
Most of this radiation is absorbed by the Earth and warms it

As well as averages, climate change causes more extreme events

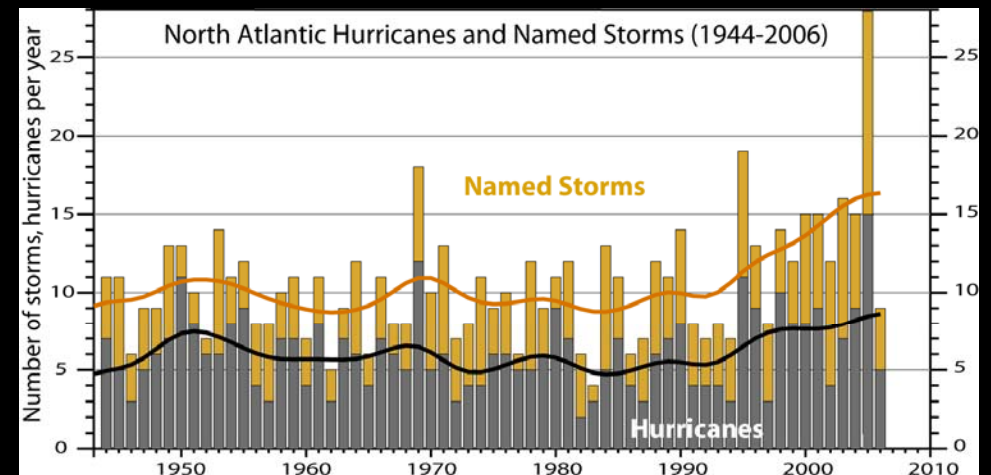
Heatwaves



Extreme rainfall

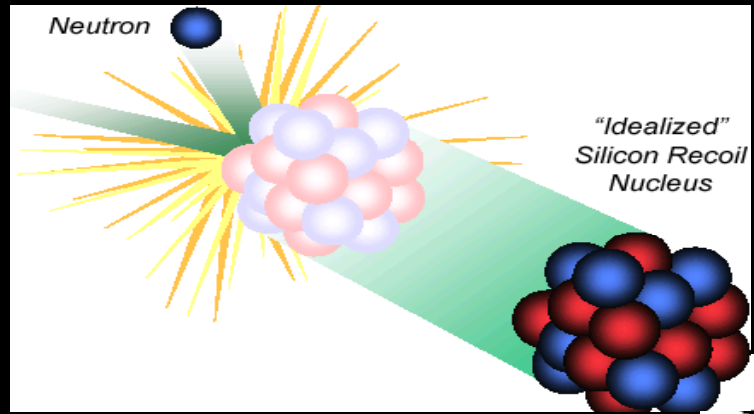


Hurricanes



Sources: Intergovernmental Government Panel on Climate Change (2007)

Cosmic Radiation or living with soft errors....



α particles α particles emitted from impurities in mold resin, solder bump, aluminum layers

High energy neutron High energy neutron in cosmic ray reacts with Si nuclei and generates secondary particles.

Thermal neutron Thermal neutron reacts with ¹⁰B which can be found in BPSG and produces alpha particles and Lithium ion.

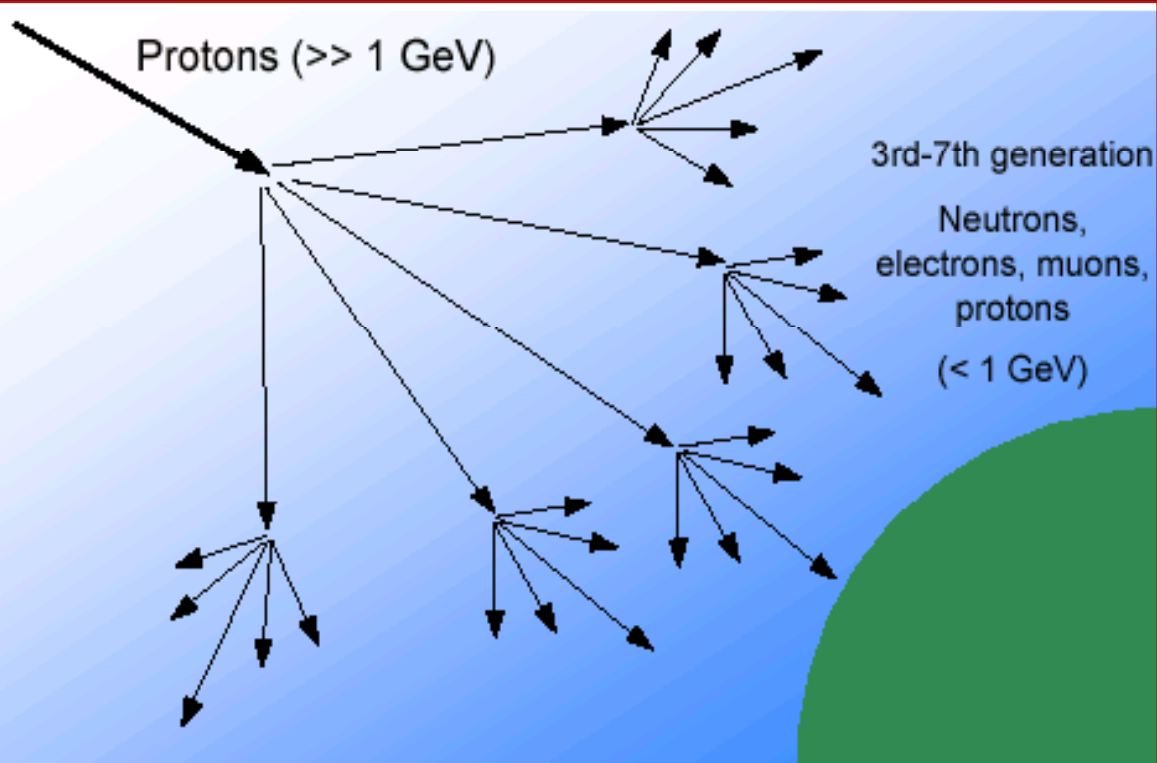
High energy neutron Thermal neutron α particle

resin polyimide Metal BPSG

Spallation Fission Ionization

memory cell

Si NWELL Li PWell NBL P-EPI P-SUB



²⁸Si

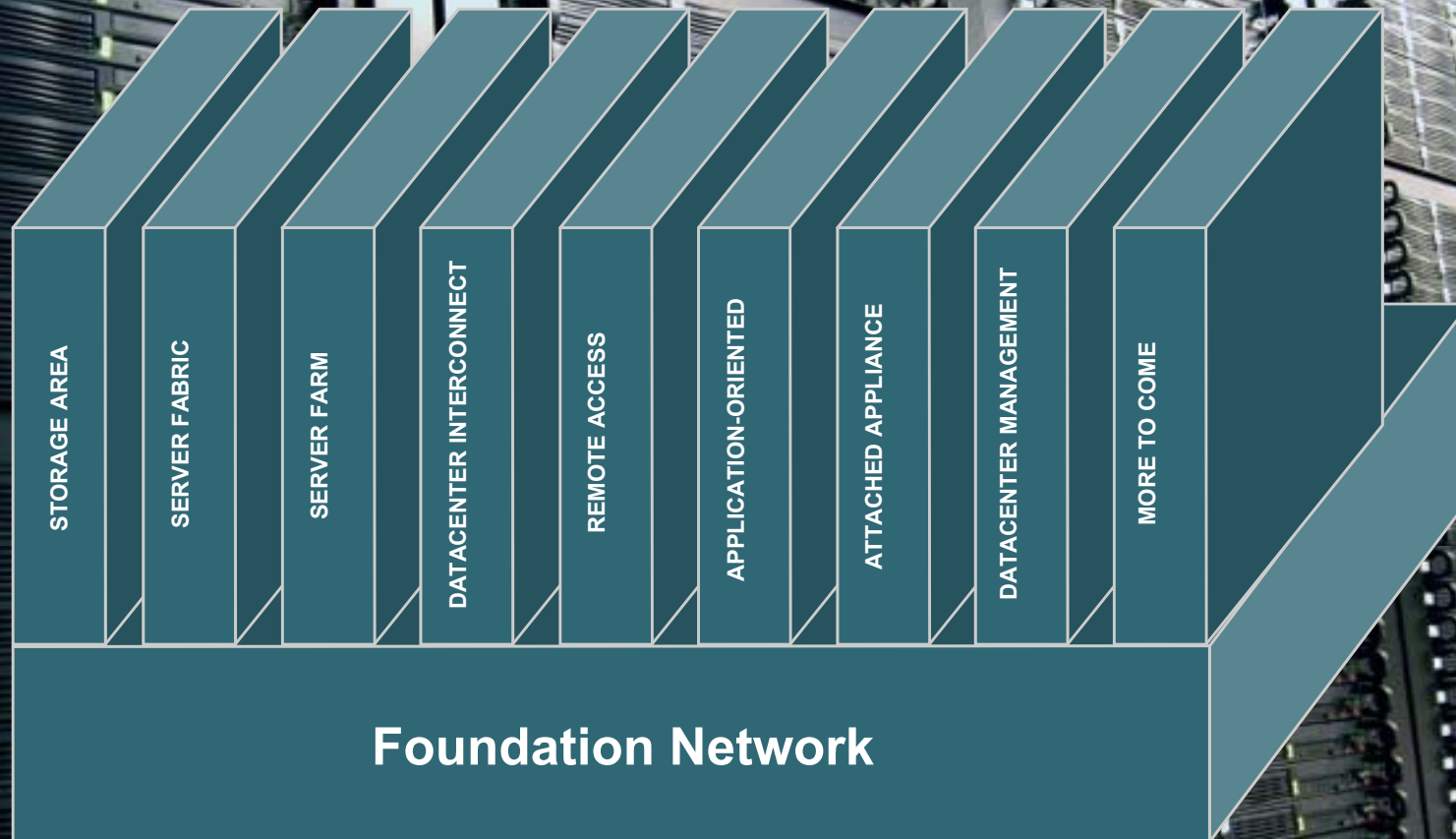
²⁵ Mg + α	2.75 MeV
²⁸ Al + p	4.00 MeV
²⁷ Al + d	9.70 MeV
²⁴ Mg + n + α	10.34 MeV
²⁷ Al + n + p	12.00 MeV
²⁶ Mg + ³ He	12.58 MeV
²¹ Ne + 2α	12.99 MeV

- Alpha Particles
- High Energy Neutrons:
- Thermal (low energy) Neutrons



- Environmental Challenges
- What's up in the Datacenter
- DC Architecture Trends
- Fibre Channel compared to Ethernet
- I/O Consolidation
- Data Center Ethernet (DCE)
- Fibre Channel over Ethernet (FCoE)
- Unified I/O Use Case
- Summary

What's Up in the Datacenter?



DATACENTER FOUNDATION NETWORK

All this are applications running on top of our Data Network

Prediction

1983	802.3	Ethernet (10 Mbps) Thick Coax
1986	802.3a	Ethernet Thin Coax
1991	802.3i	Ethernet Twisted Pair
1995	802.3u	Fast Ethernet (100 Mbps) Fiber, TP
1998	802.3z	Gigabit Ethernet (1 Gbps) Fiber
1999	802.3ab	Gigabit Ethernet Twisted Pair
2003	802.3ae	10G Ethernet (10 Gbps) Fiber
2003	802.3af	PoE, Power over Ethernet 4 / 7 / 15,4 watts
2006	802.3an	10G Ethernet Twisted Pair
2006	802.3ap	Backplane Ethernet (Blade servers and communications equipment)
2007	TRILL WG	Solve L2 STP forwarding limitations, with a link state protocol
2007	IEEE 802.1aq	Shortest Path Bridging (Optimal Bridging) in L2 Ethernet topologies
2007	IEEE 802.1Qau	Congestion Management (per virtual link)
2007	IEEE 802.3ar	Enhancements for rate limiting (per physical link)
2008	ePoE	ePoE, Enhanced Power over Ethernet ~30 watts
2008	FCoE	Fibre Channel over Ethernet
2008	DCE	Data Center Ethernet
		✓ Lossless Ethernet [Pause, Per Priority Pause]
		✓ Ethernet Congestion Management [ECN / (aka BCN)]
		✓ IEEE 802.1Qau, Priority Groups
		✓ IEEE support for (baby) giant frames
		✓ Layer 2 Multipath [Cisco DCE, TRILL]
2009	802.3at	PoE Plus, Power over Ethernet Plus ~60 watts
2010	higher speeds for Ethernet	- 40Gbps for server and storage, applications - 100Gbps for aggregation
2011	IEEE 802.3 EEE	Energy Efficient Ethernet Study Group (0BASE-T)

Evolution of Ethernet

Role of transport in enabling these technologies !

Mid 1980's

10Mb

UTP Cat 3



Late 2000's

10Gb

X2
SFP+ Cu
SFP+ Fiber
Cat 6/7 ??

Technology	Cable	Distance	Power (each side)	Transceiver Latency (link)
SFP+ CU Copper	Twinax	10m	0W normalized	~0.1µs
SFP+ USR ultra short reach	MM OM2 MM OM3	10m 100m	1W	~0
SFP+ SR short reach	MM 62.5µm MM 50µm	82m 300m	1W	~0
10GBASE-T	Cat6 Cat6a/7 Cat6a/7	55m 100m 30m	~8W ~8W ~4W	2.5µs 2.5µs 1.5µs

Ethernet Interface Evolution: 40G vs. 100G

	40G Muxed	40G Native	100G Native
IEEE Standard	None	None	Call for interest: July 2006. Expect ratification in 2010-2011.
Increased Bandwidth vs. 10GE	No, 4 x 10GE muxed solution	Yes, true 40G per interface	Yes, true 100G per interface
EtherChannel	2 links	8 links	8 links
Fiber savings	Yes	Yes	Yes
Approximate Availability	2008	2009	2010-11
Approximate Cost at FCS	2-3 x 10GE	10 x 10GE	At least 10 x 10GE

Density and Scalability Implications

Oversubscription and Uplinks

- What is the oversubscription ratio?

Develop an oversubscription reference model

Identify by application, tier, or other means

- Considerations

Server capacity

PCI and PCI-X: "Shared multi-drop bus architectures"

PCI Express: "Point to point dual simplex link"

PCIe Gen1: 2.5 Gbps/lane, PCIe Gen2: 5 Gbps/lane

PCIe maximum bandwidth of 132MB/s

Uplink choices available

Gigabit EtherChannel 10GE

10Gig EtherChannel

Flexibility in adjusting oversubscription ratio

Can I upgrade to 10GE easily?

10G EtherChannel?

Upgrade CPU and switch fabric? (sup1-2-720-?)

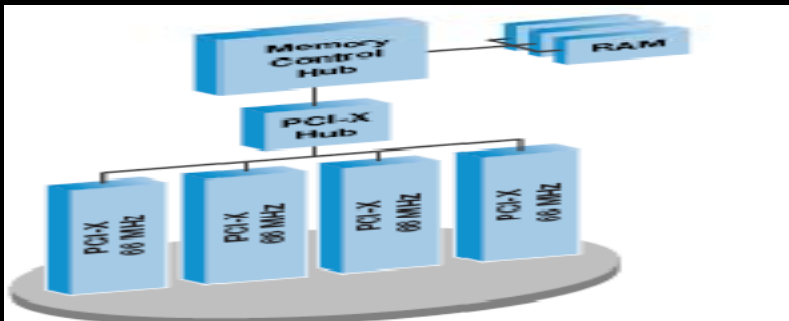
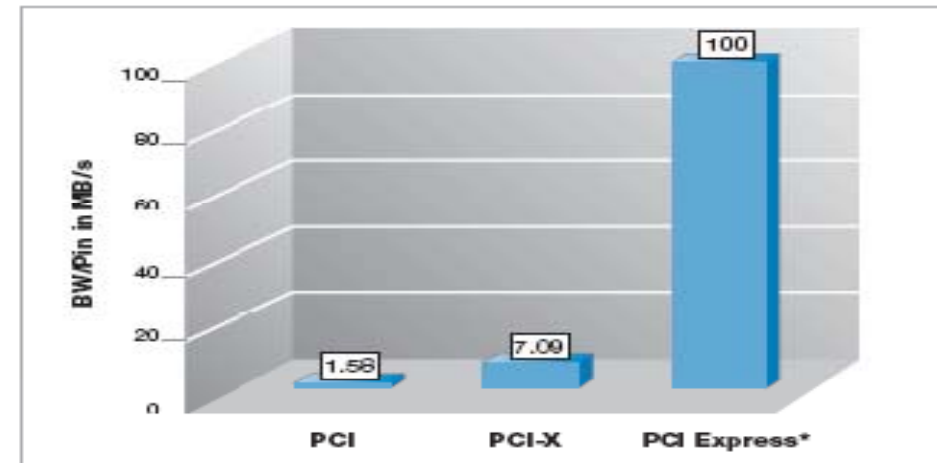


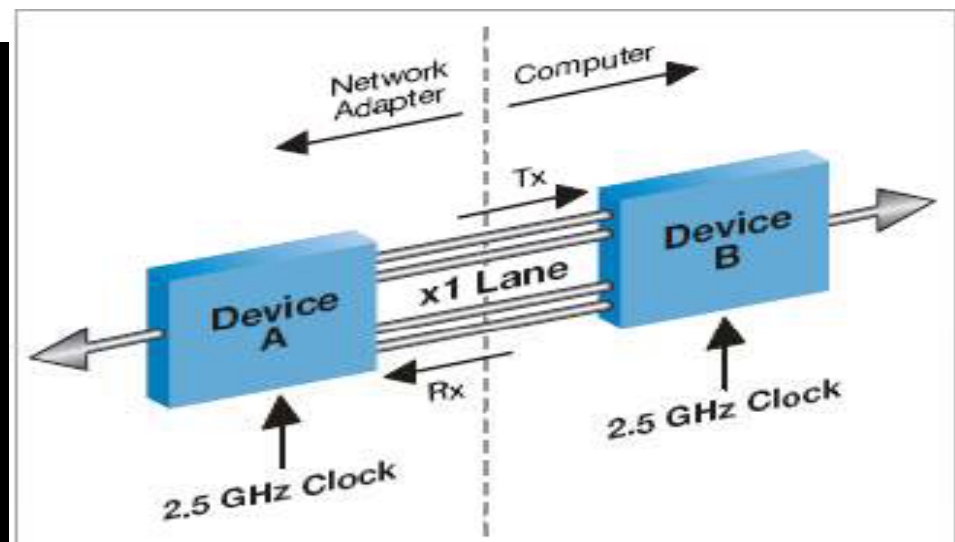
Figure 1. Bandwidth-per-pin comparison between PCI, PCI-X and PCI Express*.



PCI: 32 bits x 33 MHz and 84 pins = 1.58 MB/s

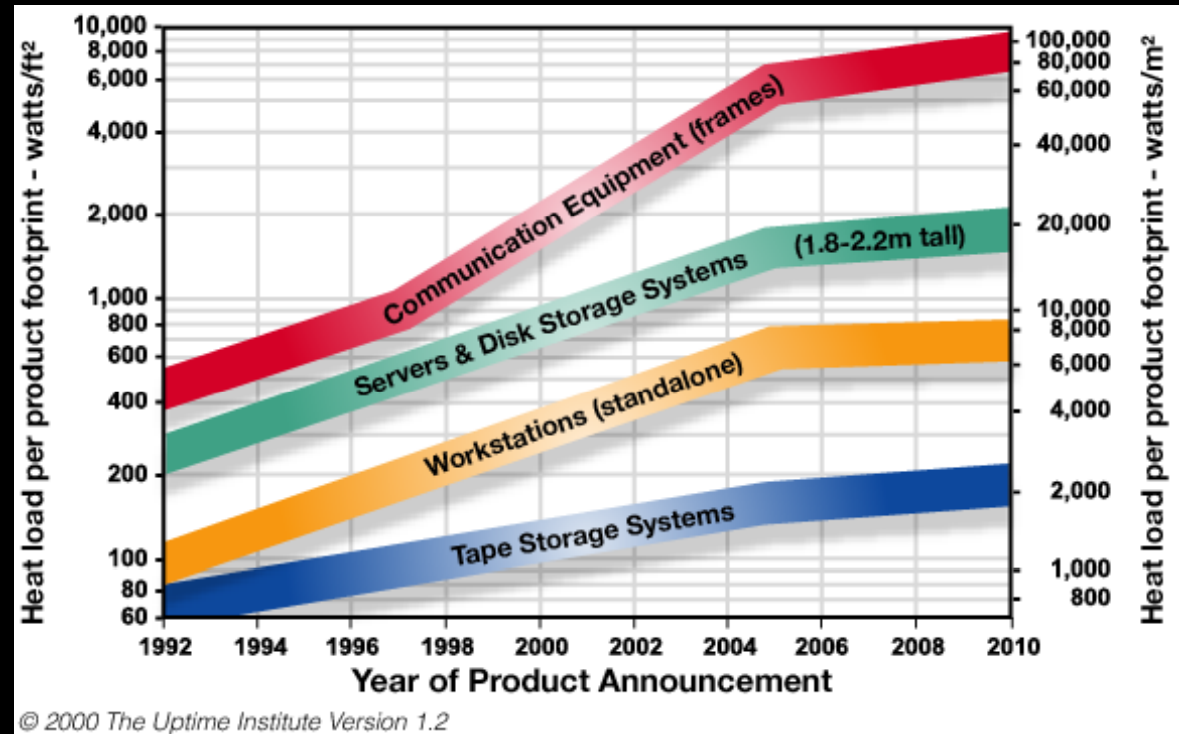
PCI-X: 64 bits x 133 MHz and 150 pins = 7.09 MB/s

PCIe: 8 bits/direction x 2.5 Gb/s direction and 40 pins = 100 MB/s



Cooling Issues and Solutions

- Today's products are hotter than yesterday's
- Tomorrow's products will be hotter than today's
- Data Center Managers prefer to tightly pack devices to conserve space
- **Measurements!**



Jonathan Koomey, Lawrence Berkeley National Laboratory, total electricity used by office and network equipment

- 74 TWh per year, about 2% of the total electricity use in the U.S
- 2005-Server-Energy bill in the US 2,7 Milliarden Dollar, 7,3 Milliarden Dollar worldwide
- 2005-Server-Power usage in the US 45 Millionen KWh, 123 Millionen KWh worldwide



What Is High Availability?

Availability	DPM	Downtime Per Year (24x365)		
99.000%	10000	3 Days	15 Hours	36 Minutes
99.500%	5000	1 Day	19 Hours	48 Minutes
99.900%	1000		8 Hours	46 Minutes
99.950%	500		4 Hours	23 Minutes
99.990%	100			53 Minutes
99.999%	10			5 Minutes
99.9999%	1			30 Seconds

- Measurement Criteria:**
- ✓ Application availability
 - ✓ Application response time
 - ✓ System availability
 - ✓ Network availability
 - ✓ RTO (Recovery Time Objective)
 - ✓ RPO (Recovery Point Objective)
 - ✓ MTTF (mean-time-to-failure)
 - ✓ MTTR (mean-time-to-repair)

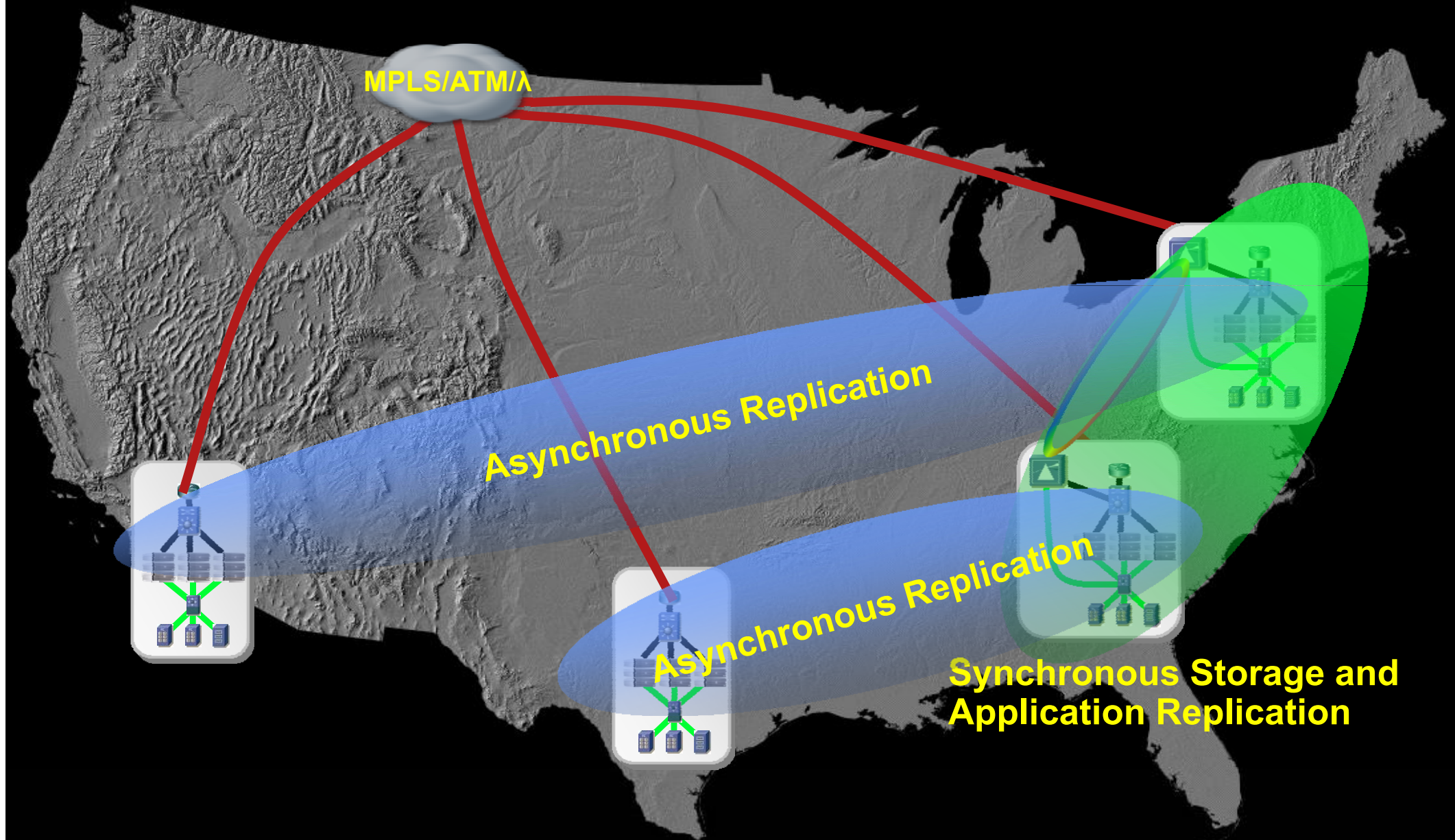
“High Availability”

**SubGroup: SJ-Transport/Data Center 12
Ops Review for September 2003**

SubGroup	Percent Avail	Average Daily Downtime	Adjusted Percent Avail	Adjusted Average Daily Downtime	Availability Target	
Production DC2	99.998	0.0 mins	99.999	0.0 mins	99.985	

Data Center Architectural Evolution

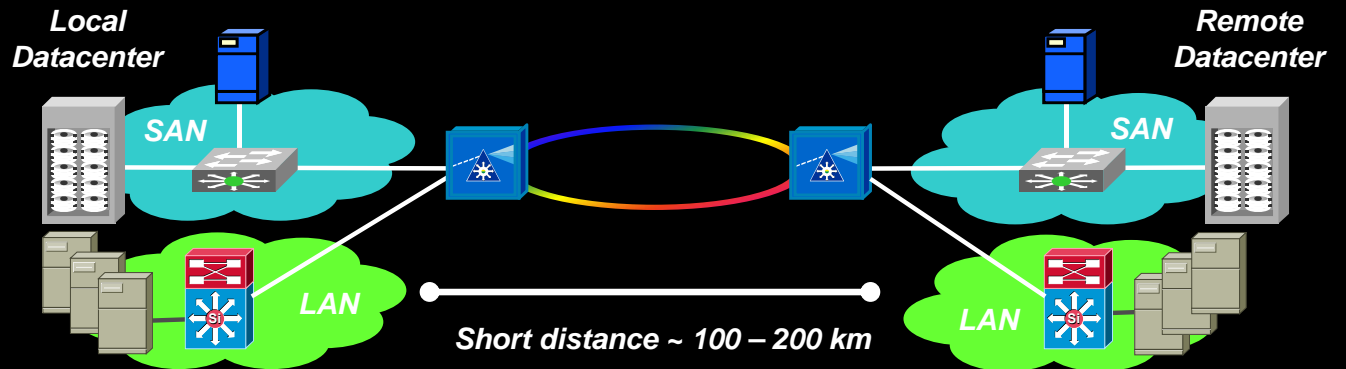
Business Continuance / Disaster Recovery



Datacenter interconnect options

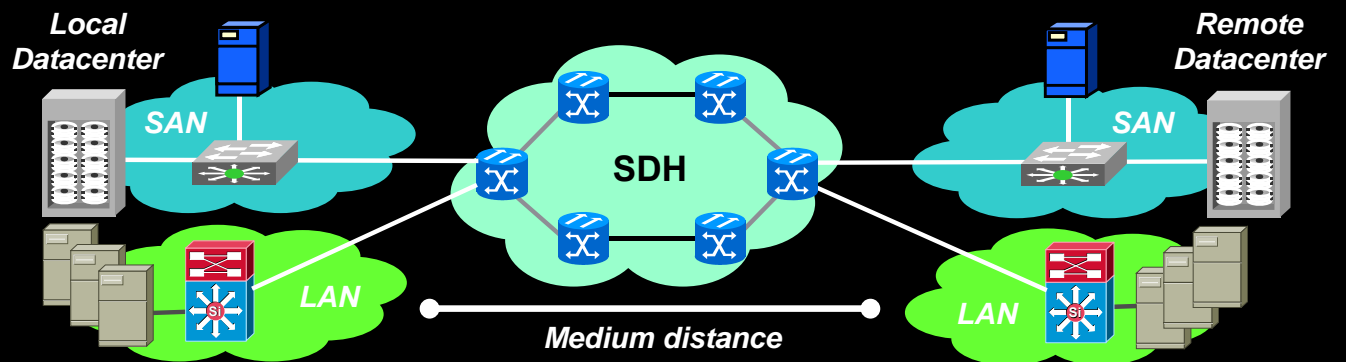
DWDM/CWDM

- most often short distance
- dark fiber must be available
- dedicated channels for LAN, SAN and other signals



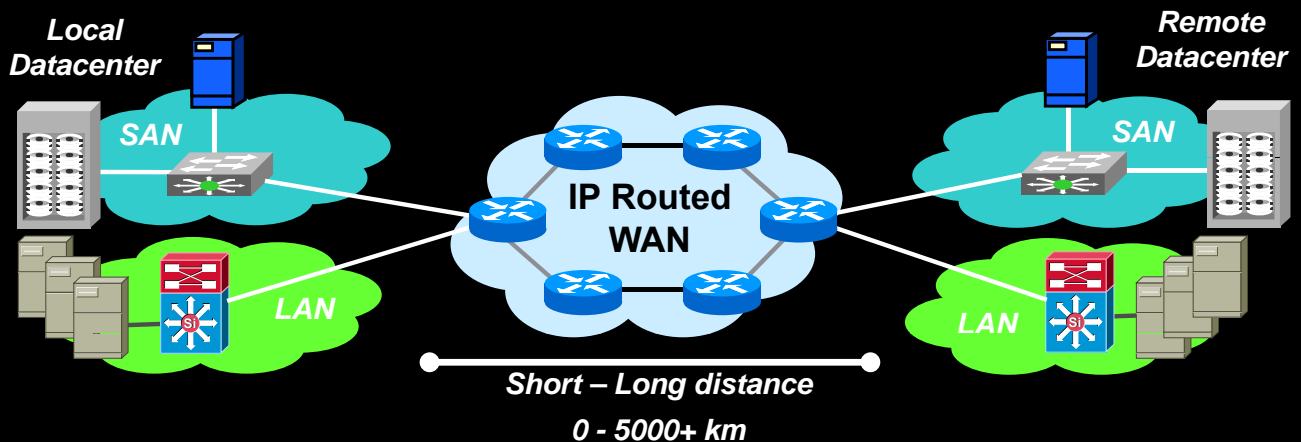
SONET/SDH

- most often short - intermediate distance
- dark fiber not avail. - distance, cost, exhaust
- links may be shared
- EoSDH and FCoSDH



IP, IP/MPLS, Metro Ethernet

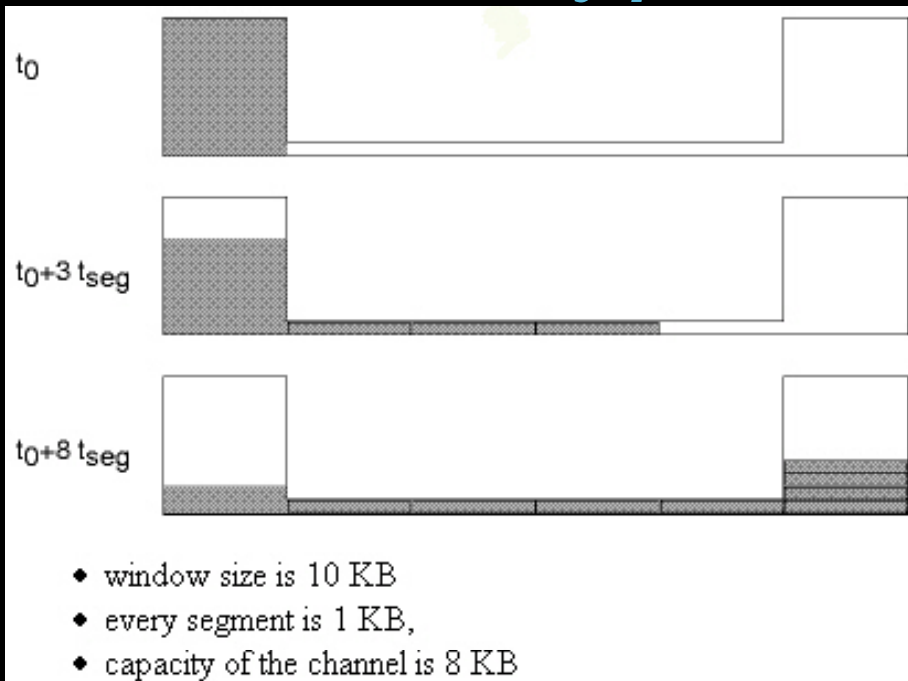
- short - long distance
- dark fiber not available
- links may be shared
- FCIP for FC and/or FICON



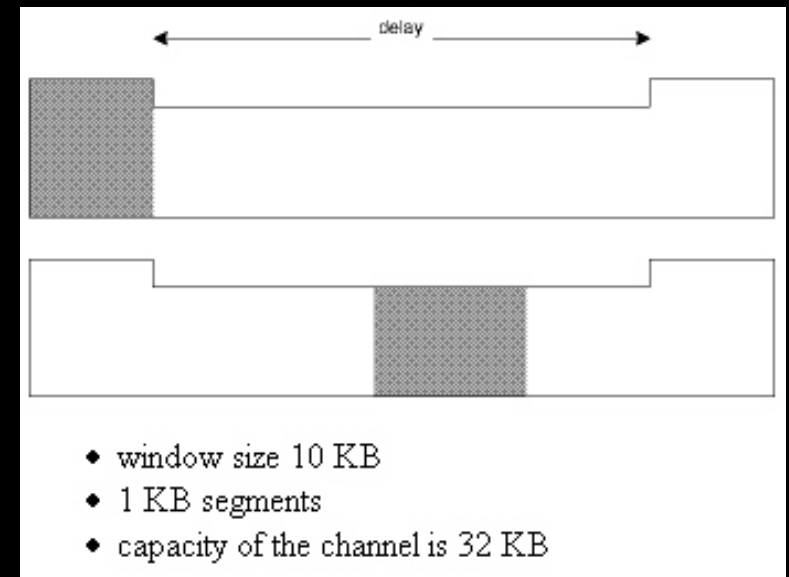
TCP/IP in Long Fat Networks

Bandwidth*Delay product

Slow Network



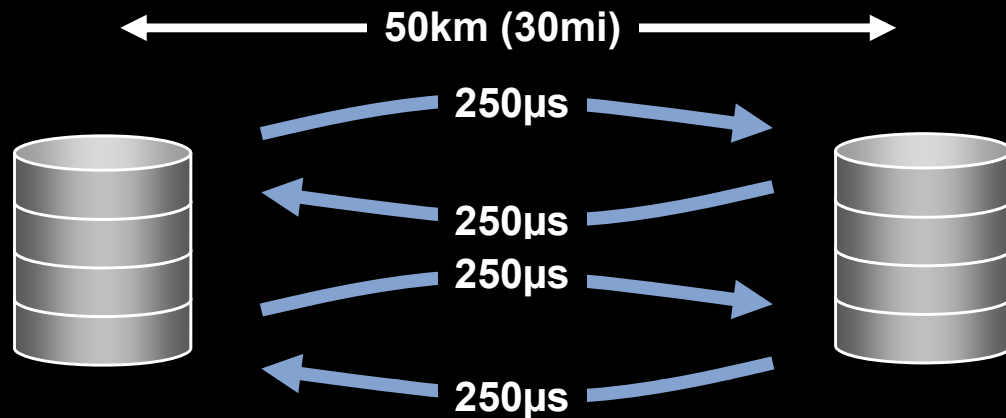
Long Fat Network



$$\text{Required Window Size} = \text{RTT} \times \text{Max Transfer Rate}$$

	5 km LAN	100 km MAN	2000 km WAN
Bytes in the Pipe:			
10 Mbit/s Ethernet	52	1040	20800
155 MBit/s ATM	703	14062	281250
1000 MBit/s GE	5208	104167	2083330
Typical Windows TCP/IP window size:			
<ul style="list-style-type: none"> ▪ below 1 Mbps = 8kB, 1 Mbps-100 Mbps = 17kB, greater than 100Mbps = 64kB ▪ if larger is needed use window scaling option - up to ~1GB (RFC1323) 			

Latency and Synchronous Replication



Speed of Light
 $c = 3 \times 10^8 \text{m/s}$ (vacuum) $\approx 3.3\mu\text{s/km}$
Speed through fiber $\approx \frac{2}{3} c \approx 5\mu\text{s/km}$

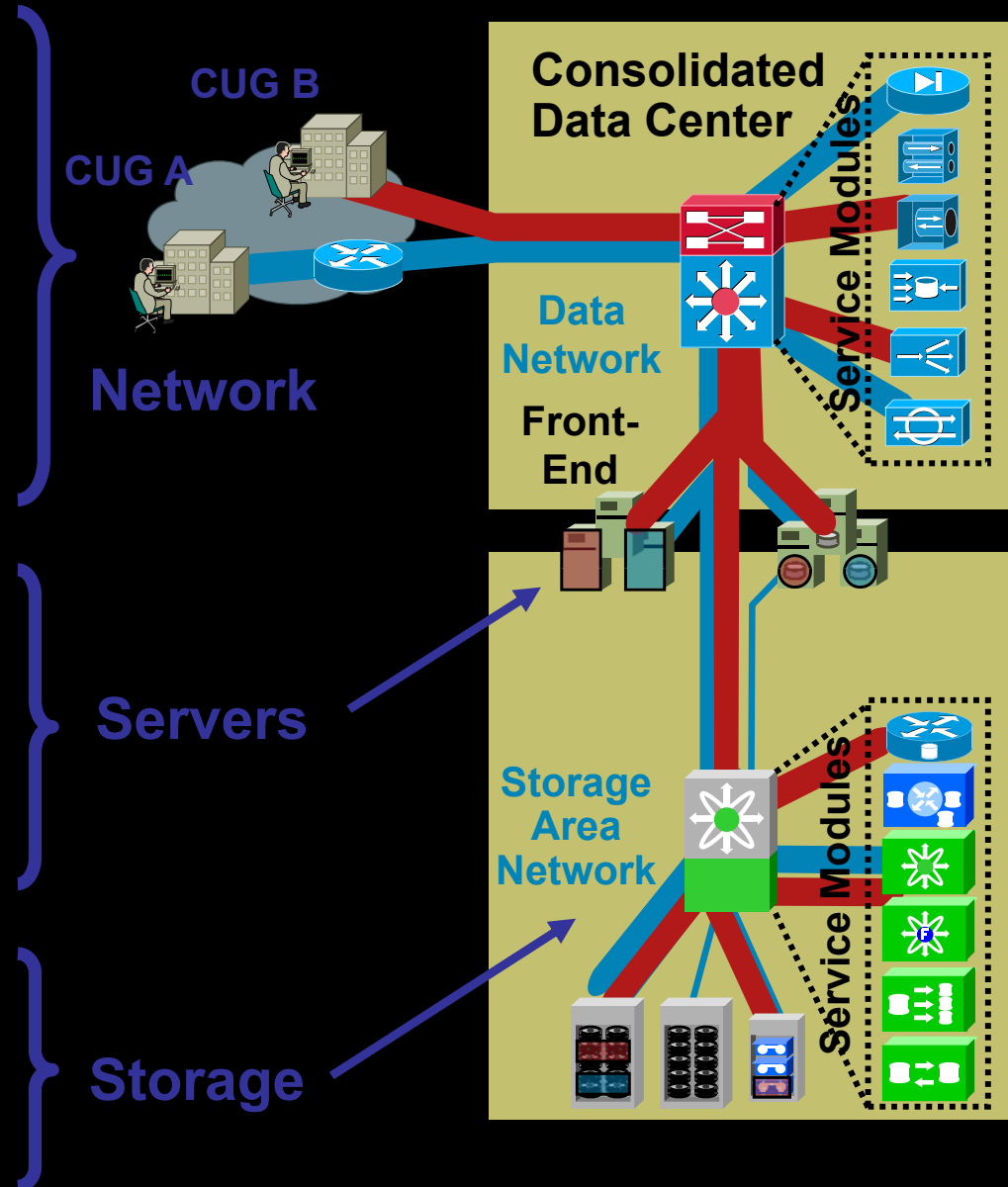
1 ms added on each mirrored write

- Two Round Trips between source and destination arrays per write I/O
2 x 2 x 5µs/km = 20µs/km additional latency
e.g. at 50km → additional 1000µs (1ms) I/O Service time (write) with Synchronous replication
Implementation dependent (ex. SRDF, DRM, HDS True copy)

“Network Virtualization” in the Data Center

One term, many contexts ...

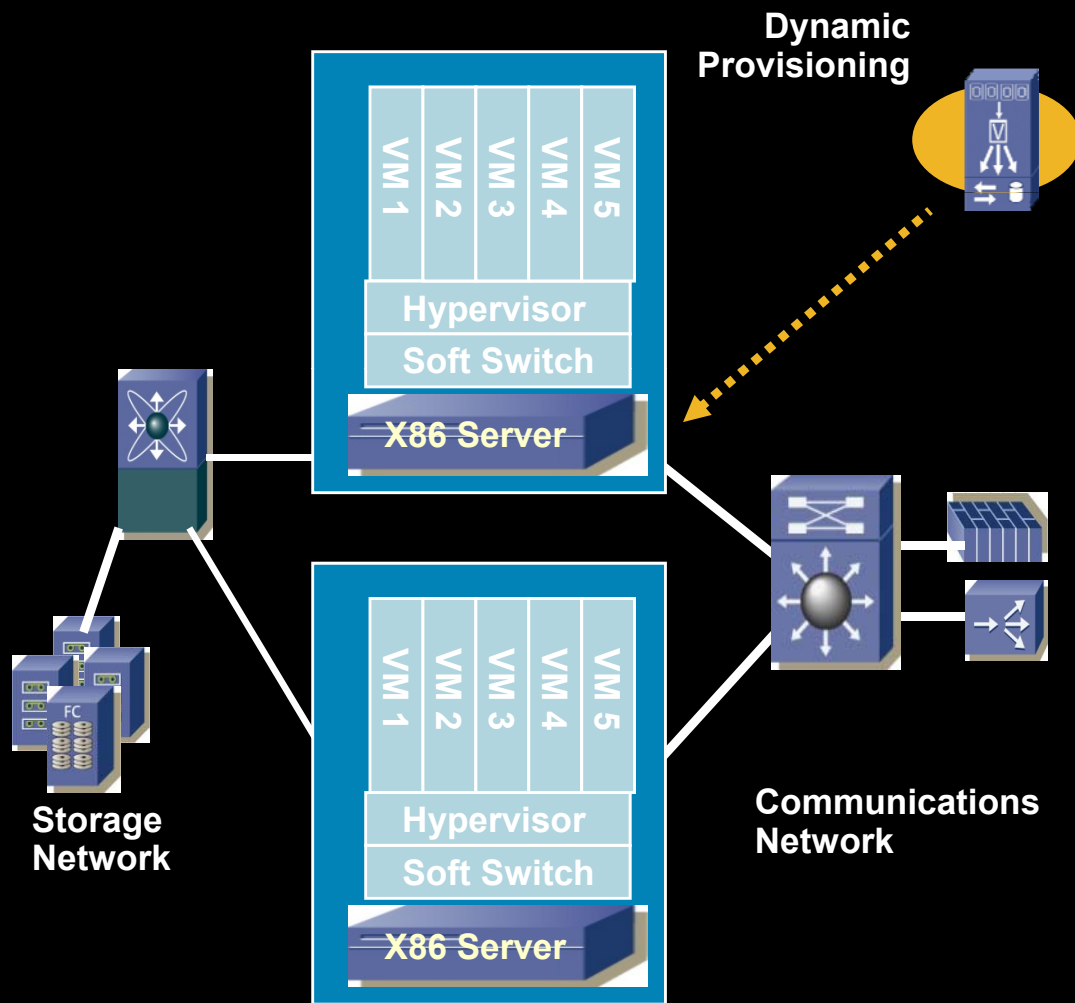
- Virtual connectivity services
 - IP/MPLS, L3 VPN, VRFs
 - L2 VPNs, VFI, PW
- Virtualized front-end
 - VLANs and private VLANs
 - Virtual intelligent services (Firewall, ACE, L4–7, etc.)
- Server virtualization
 - Clustering, GRID, Cisco VFrame™ Server, VMware ESX Server, SUN Logical Domains, Microsoft Virtual Server, OpenVZ, ...
 - Decoupling CPU, Mem and I/O functions from physical devices
- Virtualized storage
 - Virtual SANs (VSANs)
 - Network-hosted storage virtualization software



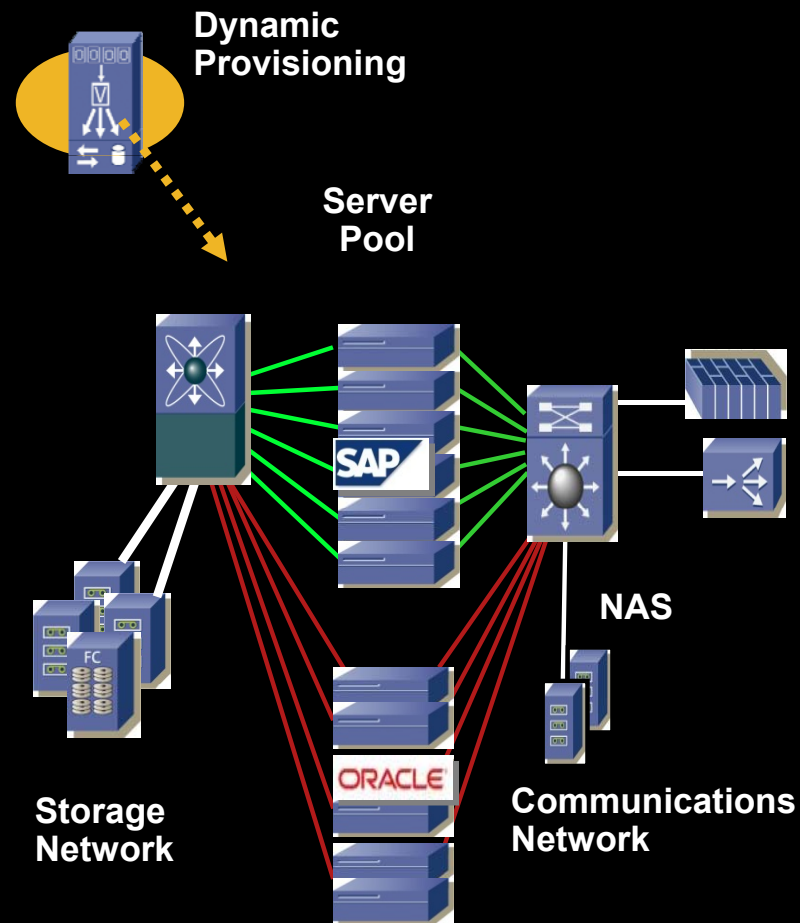
Server Virtualization

Virtualized Server Models

One Server to Many Virtualized Services



Many Servers to one Virtualized Service

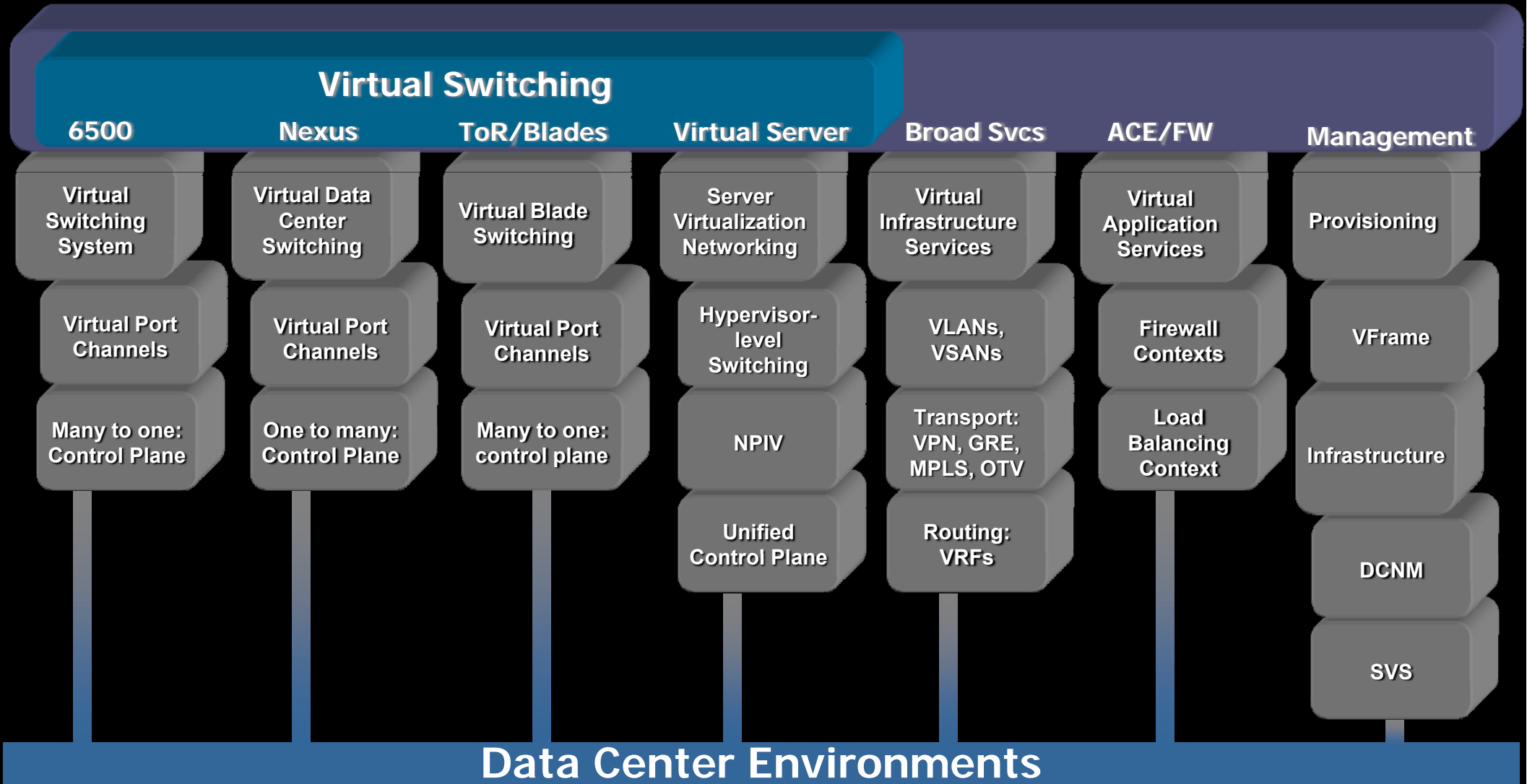


VMware Infrastructure DRS and HA clusters can include no more than **16 ESX Server hosts**, typical 8 to 10 hosts.

http://www.vmware.com/support/vi3/doc/releasenotes_esx302.html

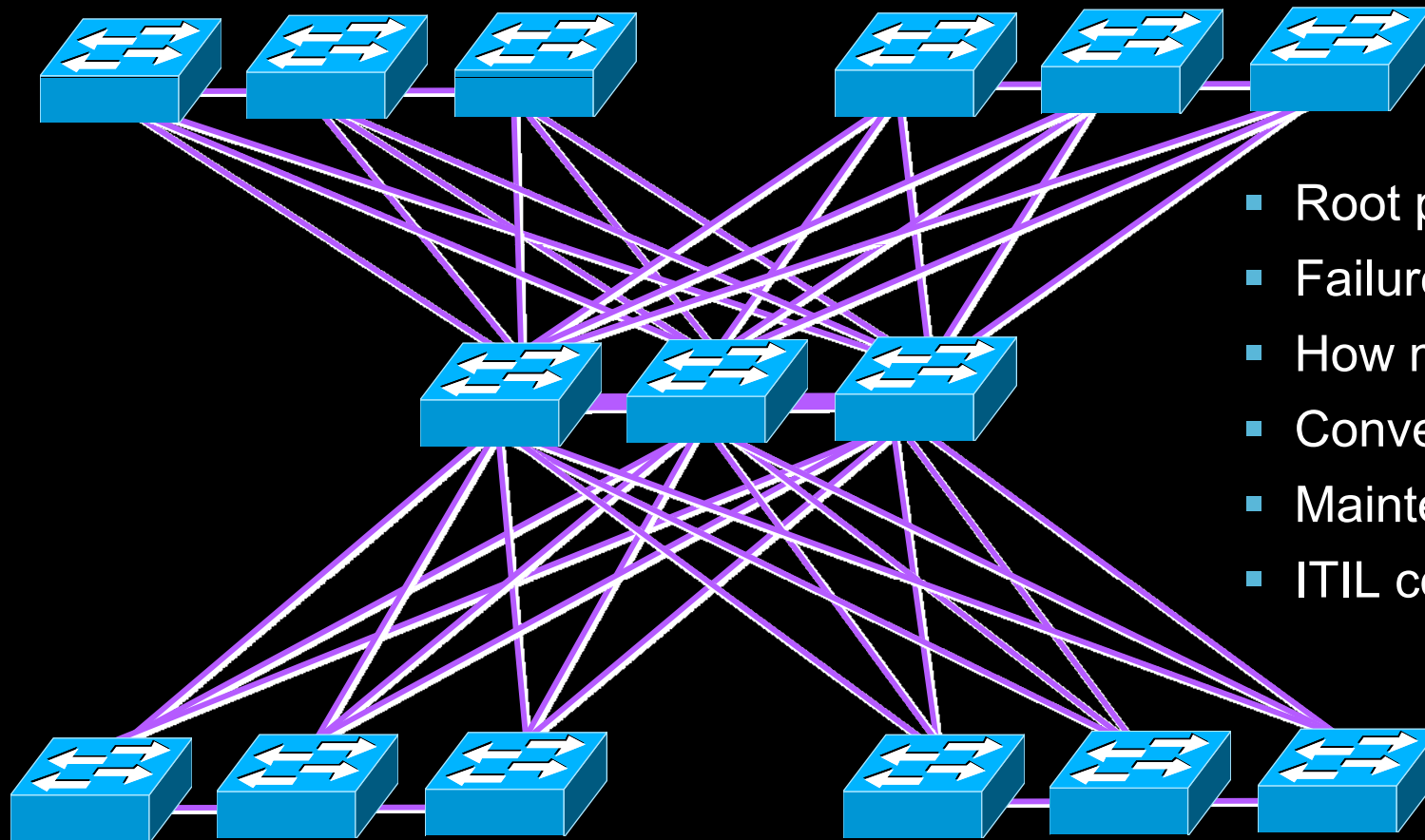
http://www.vmware.com/support/vi3/doc/vi3_esx35_vc25_rel_notes.html

Virtualization Technology



What's the problem ?

Most Datacenters are still looking like this...



- Root placement ?
- Failure deterministic?
- How many blocked links?
- Convergence ?
- Maintenance implications ?
- ITIL conformity ?



- Environmental Challenges
- What's up in the Datacenter
- DC Architecture Trends
- Fibre Channel compared to Ethernet
- I/O Consolidation
- Data Center Ethernet (DCE)
- Fibre Channel over Ethernet (FCoE)
- Unified I/O Use Case
- Summary

Network Equipment Distribution

End of Row and Middle of Row

End of Row

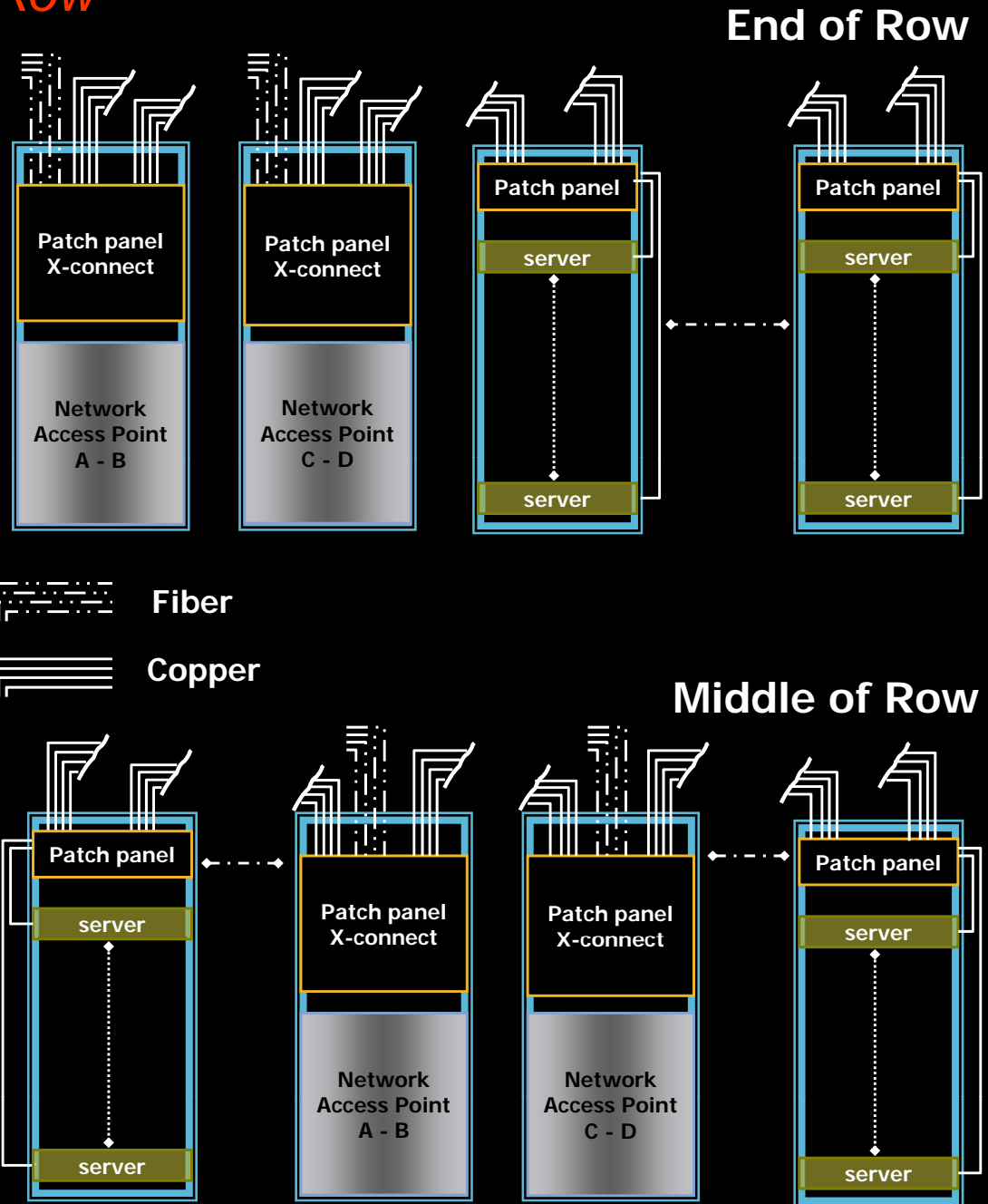
- Traditionally used
- Copper from server to access switches
- Poses challenges on highly dense server farms
 - Distance from farthest rack to access point
 - Row length may not lend itself well to switch port density

Common Characteristics

- Typically used for modular access
- Cabling is done at DC build-out
- Model evolving from EoR to MoR
 - Lower cabling distances (lower cost)
 - Allows denser access (better flexibility)
- 6-12 multi-RU servers per Rack
- 4-6 Kw per server rack, 10Kw-20Kw per network rack
- Subnets and VLANs: one or many per switch. Subnets tend to be medium and large

Middle of Row

- Use is starting to increase given EoR challenges
- Copper from servers to access switches
- Fiber may be used to aggregate ToR
- It addresses aggregation requirements for ToR access environments

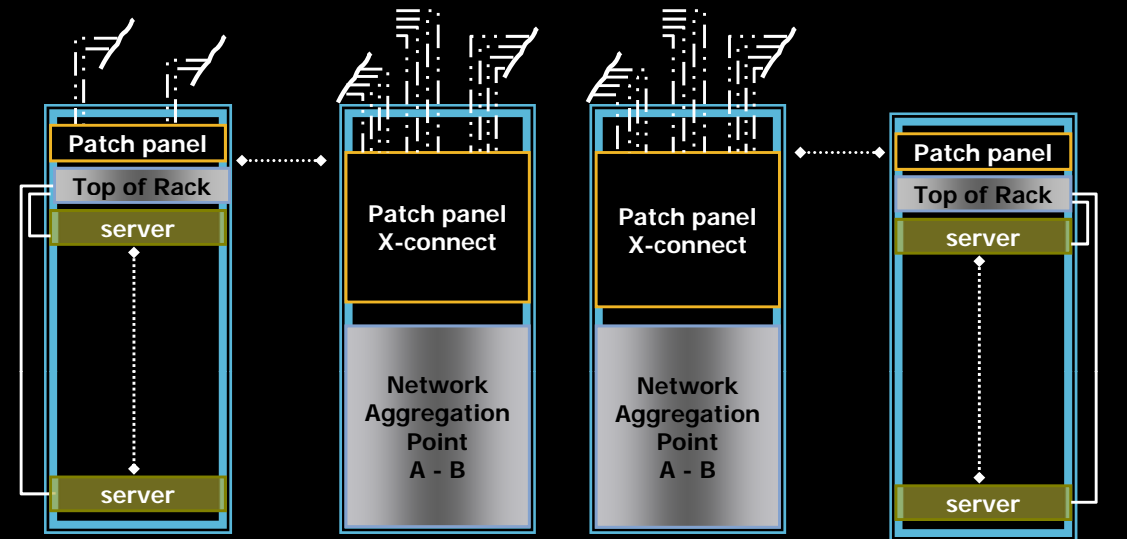


Network Equipment Distribution

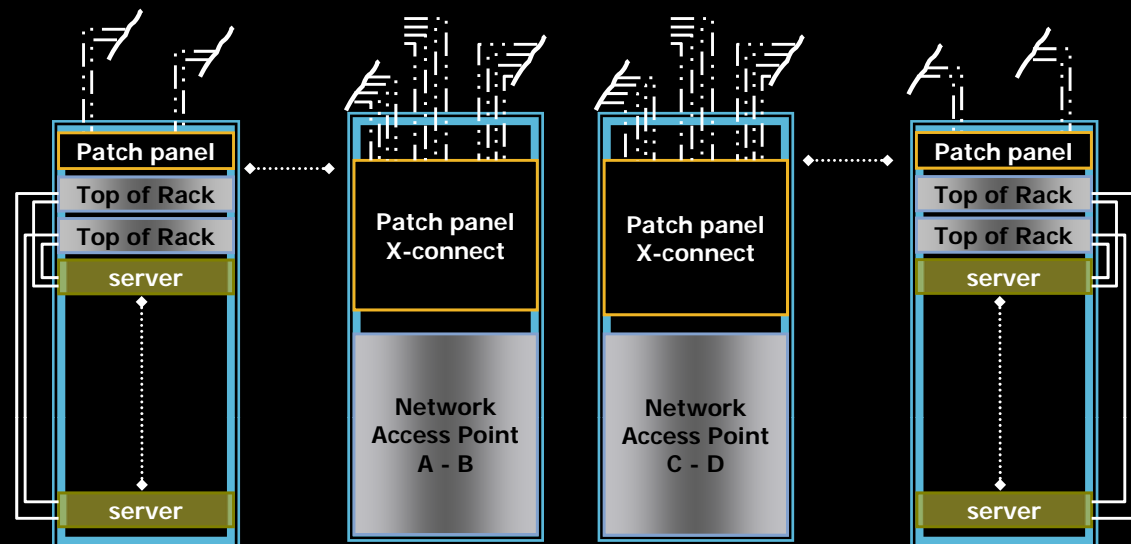
Top of Rack

ToR

- Used in conjunction with dense access racks (1U servers)
- Typically one access switch per rack
 - Some customers are considering two + cluster
- Typically:
 - ~10-15 servers per rack (enterprises)
 - ~15-30 servers per rack (SP)
- Use of either side of rack is gaining traction
- Cabling:
 - Within rack: Copper for server to access switch
 - Outside rack (uplink):
 - Copper (GE): needs a MoR model for fiber aggregation
 - Fiber (GE or 10GE): is more flexible and also requires aggregation model (MoR)
- **Subnets and VLANs:**
 - one or many subnets per access switch
 - Subnets tend to be small



To network core

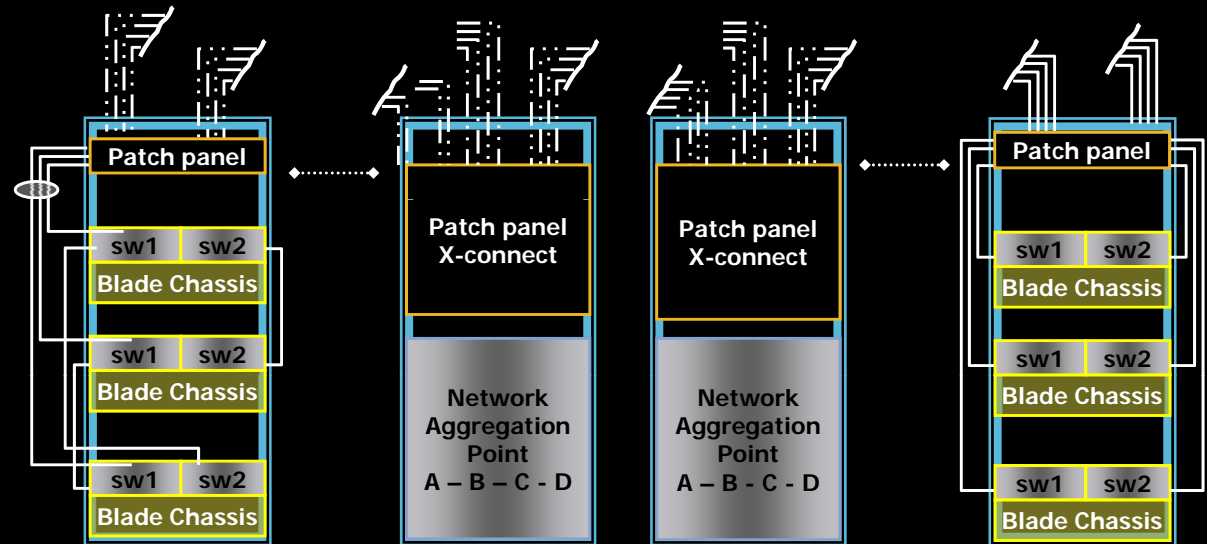


Network Equipment Distribution

Blade Chassis

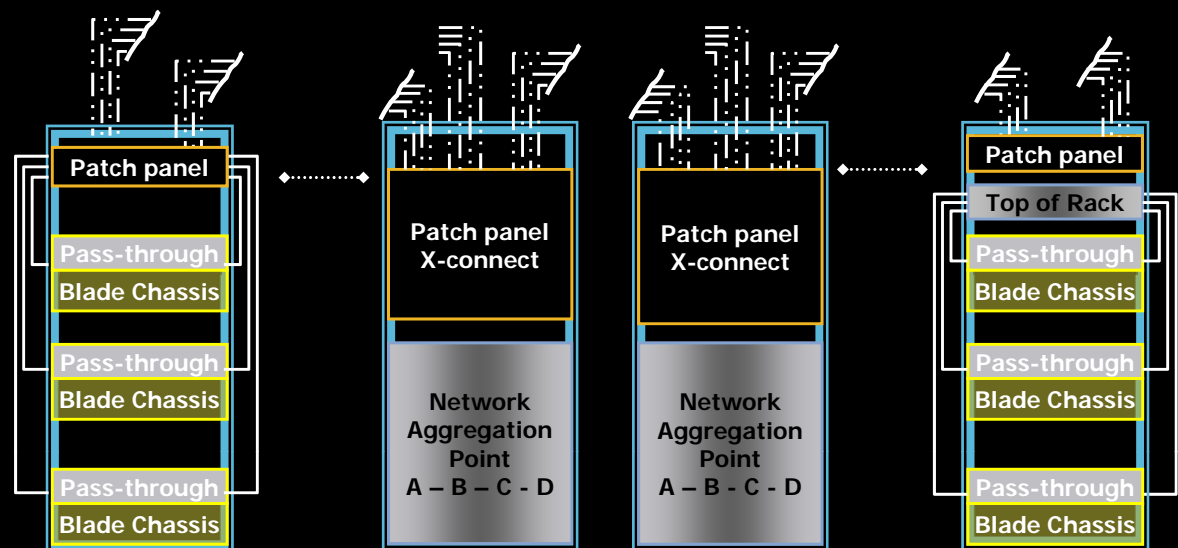
End of Row (Switch to Switch)

- Scales well for blade server racks (~3 blade chassis per rack)
- Most current uplinks are copper but the NG switches will offer fiber



Middle of Row (Pass-through)

- Scales well for pass-through blade racks
- Copper from servers to access switches

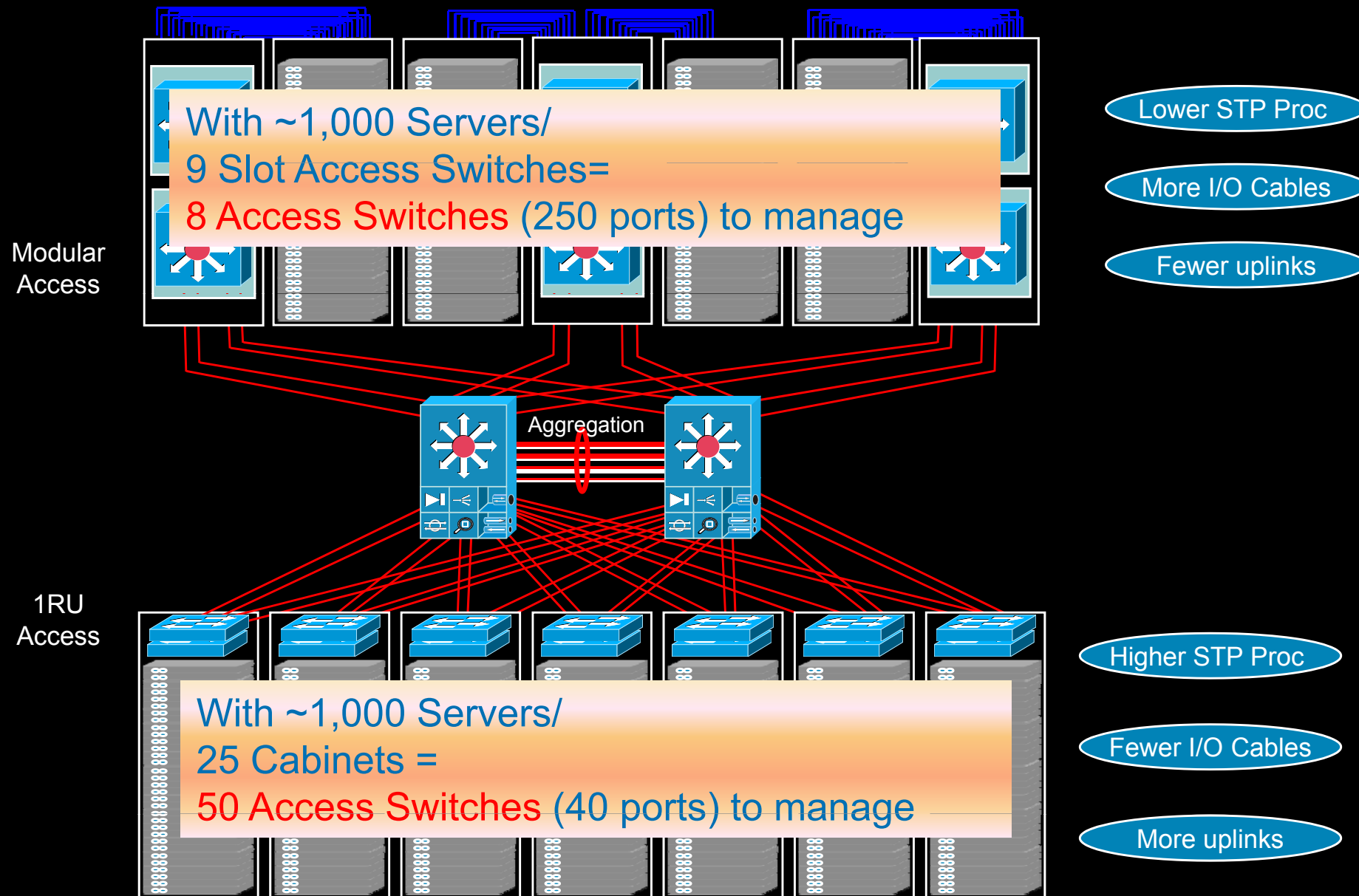


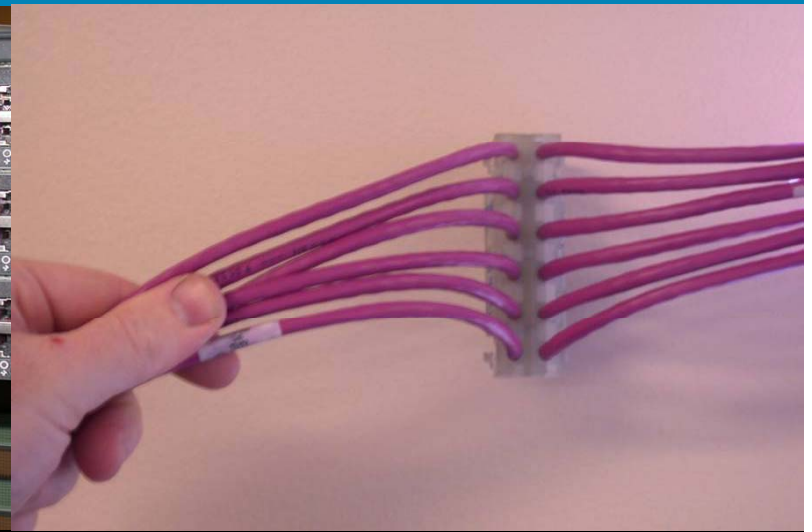
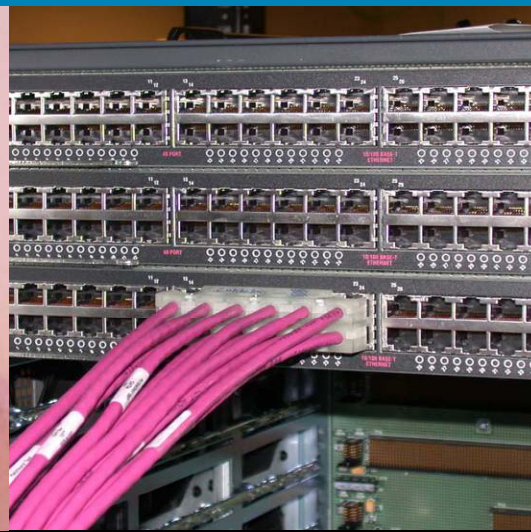
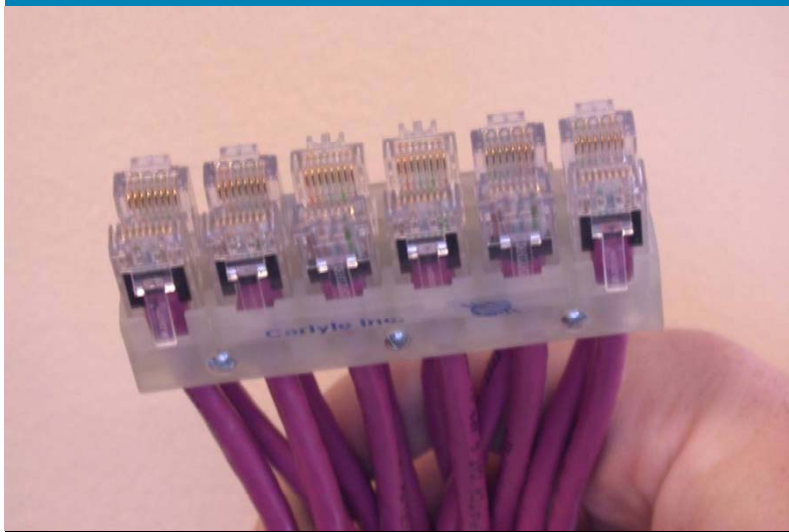
ToR

- Have not seen it used in conjunction with blade switches
- May be a viable option on pass-through environments if the access port count is right

Density and Scalability Implications

Top of Rack ./ Middle Of Row





- Environmental Challenges
- What's up in the Datacenter
- DC Architecture Trends
- Fibre Channel compared to Ethernet
- I/O Consolidation
- Data Center Ethernet (DCE)
- Fibre Channel over Ethernet (FCoE)
- Unified I/O Use Case
- Summary



LAN – SAN – Similarities

yes, this is not all technical correct but just to get the point...

LAN	SAN
Naming: MAC Address, FQDN	Naming: WWNN, WWPN
IANA Ethernet Address Block	Reserved FC_ID Addresses
Location: MAC + VLAN TAG, DHCP Option 82	Location: FC_ID
IP-Addressing: Network, Subnet, Host	FC-Addressing: Domain ID, Area ID, Port ID 239 Domains ID's (01-EF)
IPv4 32 Bit, IPv6 128 Bit	FC 24Bit
Routing: based on IP ~L3: OSPF with single AREA 0.0.0.0 L2: TRILL or other L2 Multipathing in the future	Routing: based on FC_ID FSPF
NAT (Network Address Translation)	IVR (Inter VSAN Routing)
QoS, RSVP	Fibre Channel Frame Classes, Class 1, Class 2, Class 3, Class 4, Class 6, Class F
Address Resolution: DNS Server / Router DNS -> name to IP ARP -> IP to MAC	Address Resolution: Directory/Name Server running on a switch WWPN to FCID
IEEE 802.3x, rate based flow control, link-level "Pause Frame"	BB_Credits, credit based flow control, link-level and end-to-end
switchport mode: access, dot1q-tunnel, dynamic, private-vlan, trunk, routed, routed p2p	Port Types: 'N' port, 'F' port, 'L' port, 'NL' port, 'E' port, 'GL' port, Auto Discovery
Trunking: dot1q, isl, negotiate	Inter-Switch Link (ISL): E_Port to E_Port, TE_Port and EISL
802.1x + Flexible Authentication	Extended Link Services (ELS): FLOGI, PLOGI, FAN, PRLI, PRLO, SCN, SCR, RSCN
~TCP Sequence Numbers	Fibre Channel Exchange
private VLAN's, ACL's	Fibre Channel Fabric Zoning
VLANs	VSANs
Frame Relay BCN, DCE BCN/QCN in the future	Fibre Channel Congestion Control (FCC)
BPDU, TCN – Topology Change Notification	SCN – State Change Notification

Some Math on Encoding

	Physical Layer	Encoding Method	Data Link Layer (Mbit/s)	Data Link Layer (MB/s)
FC 1G	1,062.50 Mb/s	8b/10	850 Mbit/s	106.25 MB/s
FC 2G	2,125.00 Mb/s	8b/10	1,700 Mbit/s	212.5 MB/s
FC 4G	4,250.00 Mb/s	8b/10	3,400 Mbit/s	425 MB/s
FC 10G	10,518.75 Mb/s	64b/66	10,200 Mbit/s	1,275 MB/s
FC 8G	8,500.00 Mb/s	8b/10	6,800 Mbit/s	850 MB/s
Ethernet 1G	1,250 Mb/s	8b/10	1,000 Mbit/s	125 MB/s
Ethernet 10G (LAN)	10,312.5 Mb/s	64b/66	10,000 Mbit/s	1,250 MB/s

What is Wire Rate?

FC Frame Payload Size	Link Rate	Frame Size (+header)	Inter-Frame Gap	Frame Rate (frames/sec)	Effective Throughput	
					(MB/s)	wire % efficiency
Small-Sized Frame (16 byte payload)	1G FC	52	24	1,398,026	22.368	21%
	2G FC	52	24	2,796,052	44.736	21%
	4G FC	52	24	5,592,105	89.473	21%
	8G FC	52	24	11,184,210	178.947	21%
	10G FC	52	8	21,250,000	240.000	26%
Maximum-Sized Frame (2112 byte payload)	1G FC	2148	24	48,918	200.000	97.2%
	2G FC	2148	24	97,836	400.000	97.2%
	4G FC	2148	24	195,672	413.26	97.2%
	8G FC	2148	24	391,344	826.52	97.2%
	10G FC	2148	8	591,373	1,248.98	99.9%

Improved efficiency versus 1G/2G/4G/8G due to lower IFG



- Environmental Challenges
- What's up in the Datacenter
- DC Architecture Trends
- Fibre Channel compared to Ethernet
- I/O Consolidation
- Data Center Ethernet (DCE)
- Fibre Channel over Ethernet (FCoE)
- Unified I/O Use Case
- Summary

Challenges in the Data Center

Everything is Growing! *And Growing!*

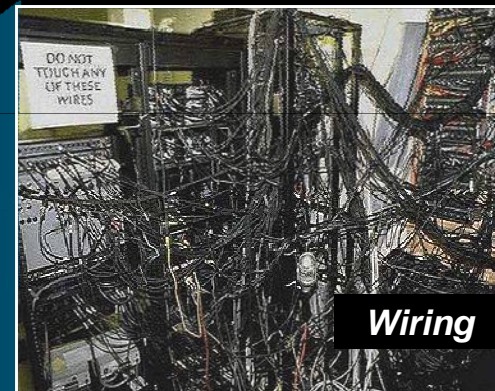
Virtual Machines



Servers



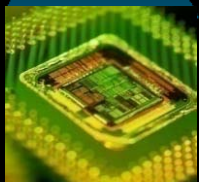
Storage



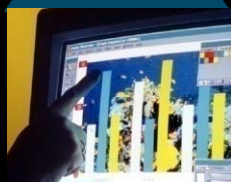
Wiring



Power



Multi Socket
Multi Core



Application
Demands

All of this growth is causing a significant problem with power, cabling and cooling

Today I/O resides in isolated Networks

- Block I/O

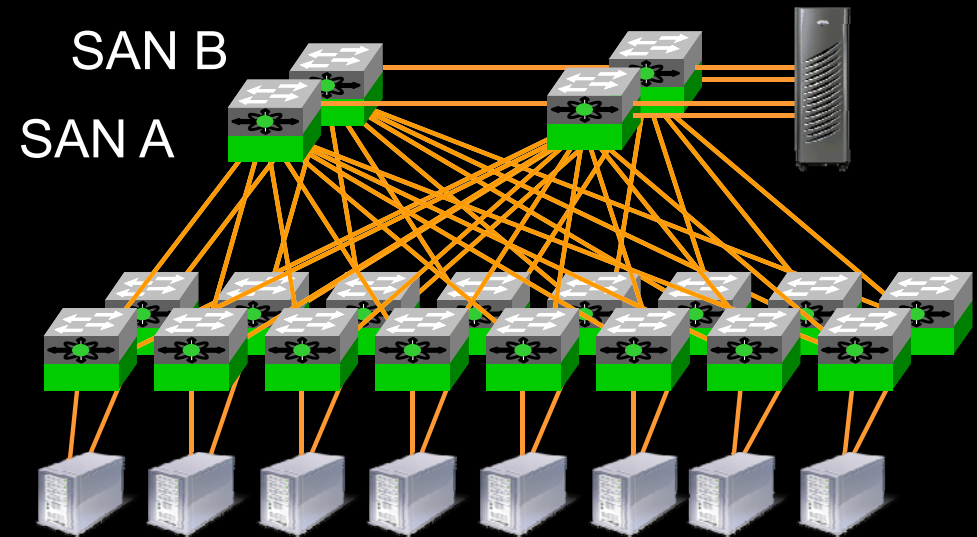
SAN/Storage Traffic

Lossless Fabric

Limited convergence on
dedicated Ethernet
Networks

FCIP

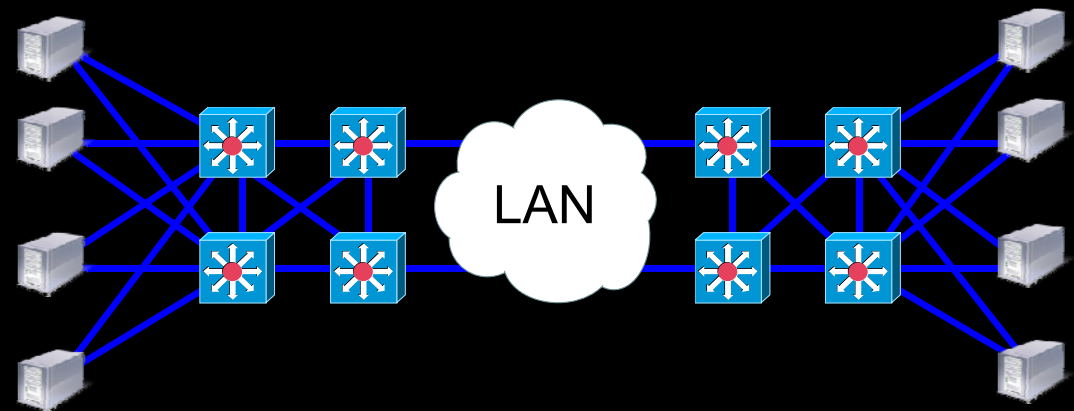
iSCSI



- Ethernet

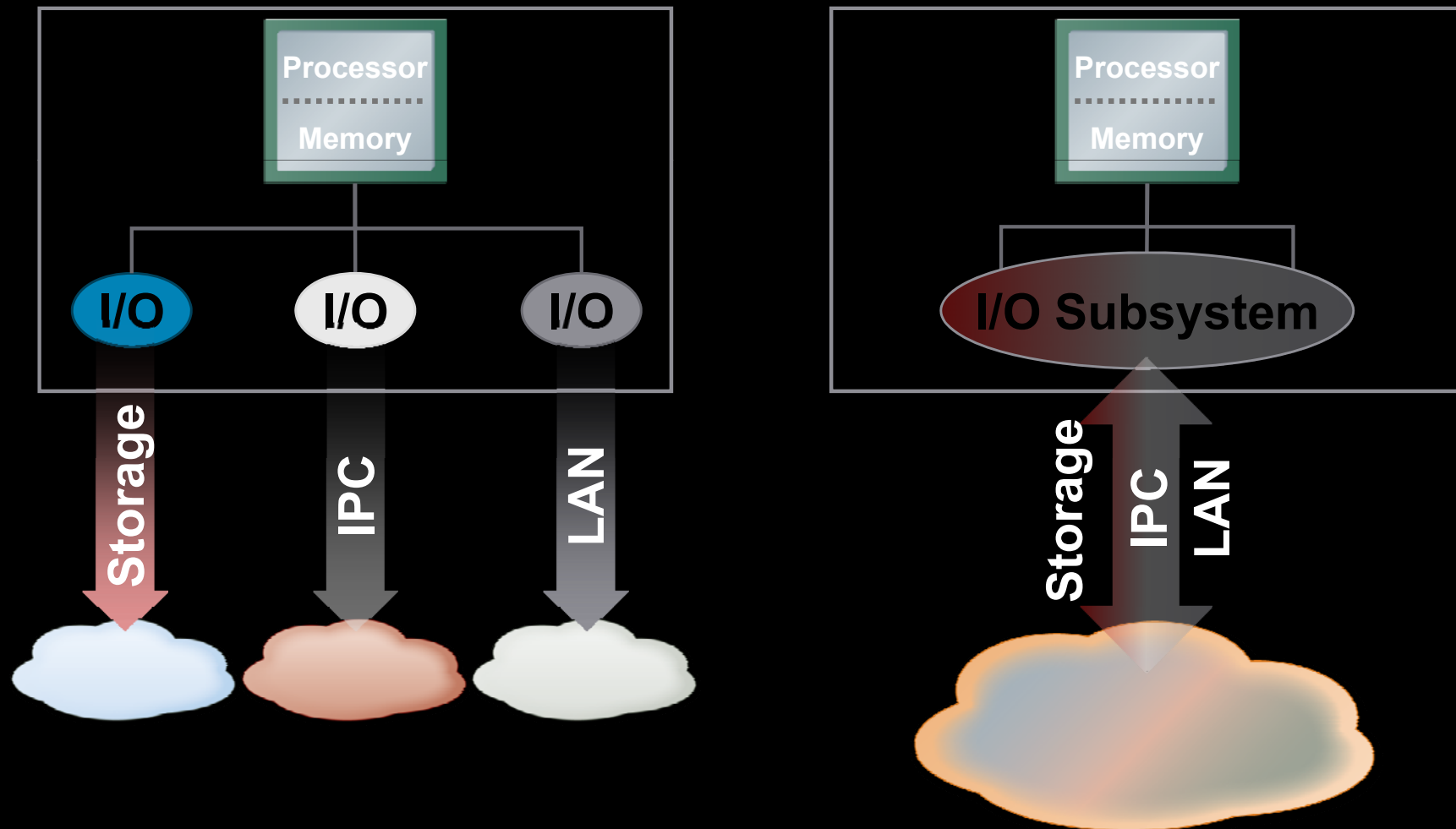
LAN Traffic

“Packet Drop” flow control



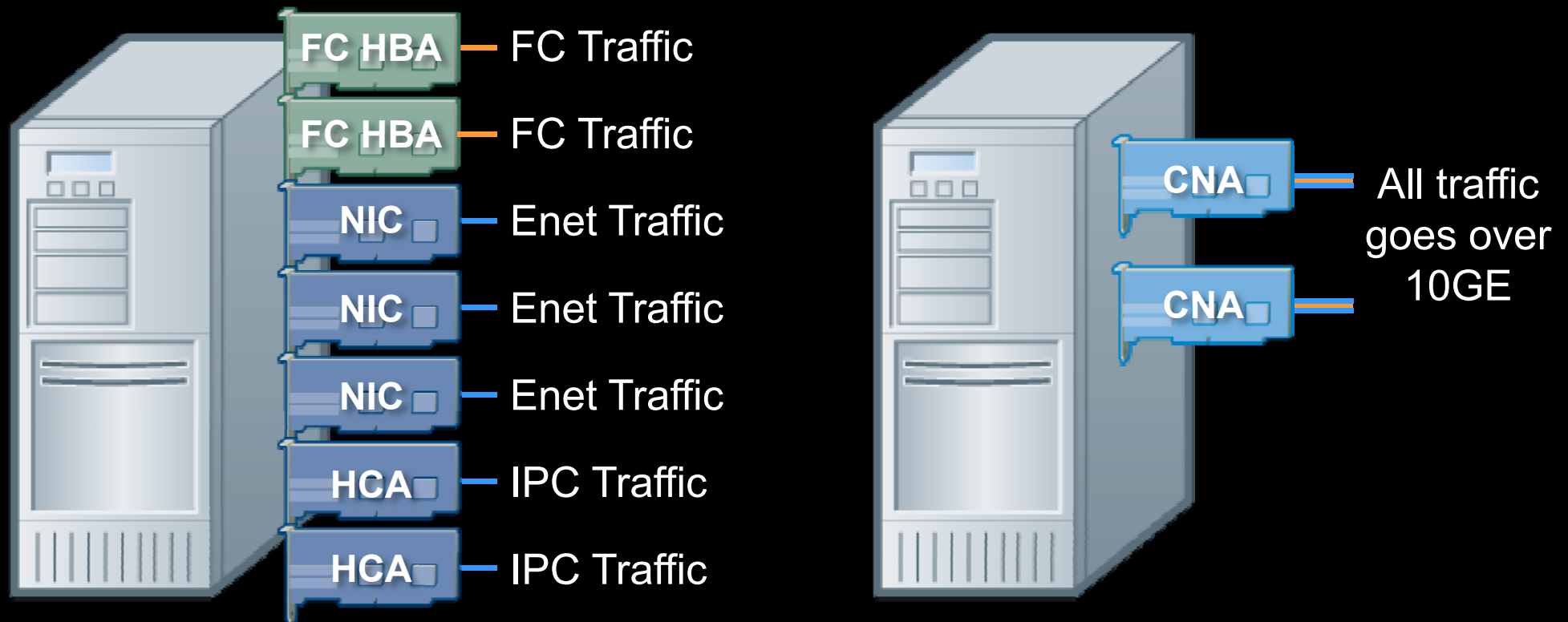
I/O Consolidation in the Network

A single network, instead of three!



I/O Consolidation in the Host

- Fewer CNAs (Converged Network adapters) instead of NICs, HBAs and HCAs
- Limited number of interfaces for Blade Servers

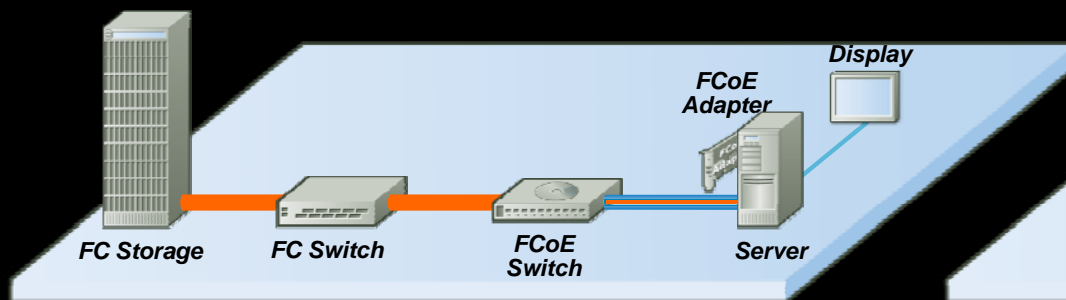


I/O Consolidation: Benefits to Customers



Fewer CNAs and Cables

Storage keeps the same Management Model as Native FC



No Gateway



Less Power and Cooling

Why Has it not Succeeded Yet?

- Previous attempts

 - Fibre Channel

 - Never credible

 - Infiniband

 - Not Ethernet

 - iSCSI

 - Not Fibre Channel

- Before PCI-Express there was not enough I/O bandwidth in the servers

- It needs to be Ethernet, but ...

 - 1 GE didn't have enough bandwidth



- Environmental Challenges
- What's up in the Datacenter
- DC Architecture Trends
- Fibre Channel compared to Ethernet
- I/O Consolidation
- Data Center Ethernet (DCE)
- Fibre Channel over Ethernet (FCoE)
- Unified I/O Use Case
- Summary

What is Data Center Ethernet (DCE)?



Data Center Ethernet is an architectural collection of Ethernet extensions designed to improve Ethernet networking and management in the Data Center.

Cisco is showing innovation while working through the standardization process with these extensions in open standards forums.

Ethernet enhancements, DCB, CEE and DCE™

Different names for different collections?

Different organizations created different names to identify the collection of the specifications, all based on the **same core** specifications

■ PFC, ETS, DCBX

- IEEE “Data Center Bridging (DCB)”

No way to identify the group of specifications, therefore grouped the specifications into DCB

- IBM “Converged Enhanced Ethernet (CEE)”

■ PFC, ETS, DCBX, L2MP, BCN/QCN, lossless fabric

- CISCO “Data Center Ethernet (DCE)”

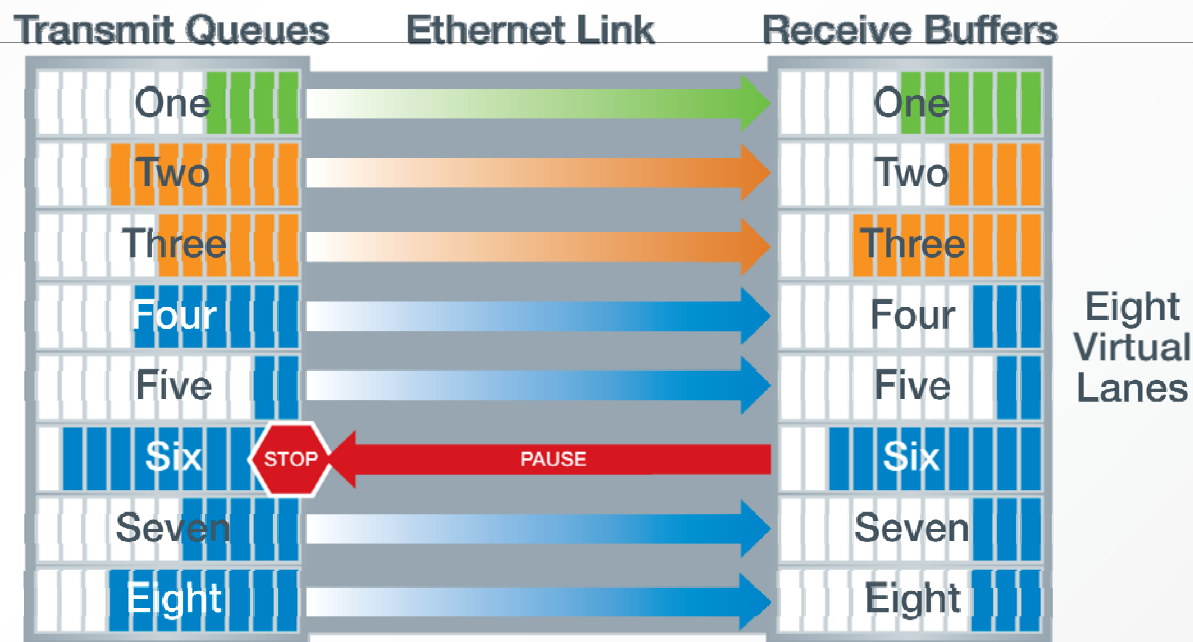
Superset of the CEE and DCB proposals, using the same three specifications
Includes L2MP, lossless fabric, and congestion notification

Data Center Ethernet Features Overview

Feature	Benefit
Priority-based Flow Control (PFC)	Provides class of service flow control. Ability to support storage traffic
CoS Based BW Management (ETS)	Grouping classes of traffic into "Service Lanes" IEEE 802.1Qaz, CoS based Enhanced Transmission
Congestion Notification (BCN/QCN)	End to End Congestion Management for L2 network
Data Center Bridging Exchange (DCBX)	Auto-negotiation for Enhanced Ethernet capabilities DCBX (Switch to NIC)
L2 Multi-path for Unicast & Multicast (L2MP)	Eliminate Spanning Tree for L2 topologies Utilize full Bi-Sectional bandwidth with ECMP
Lossless Service	Provides ability to transport various traffic types (e.g. Storage, RDMA)

Data Center Ethernet Features (PFC)

Priority-Based Flow Control (PFC)

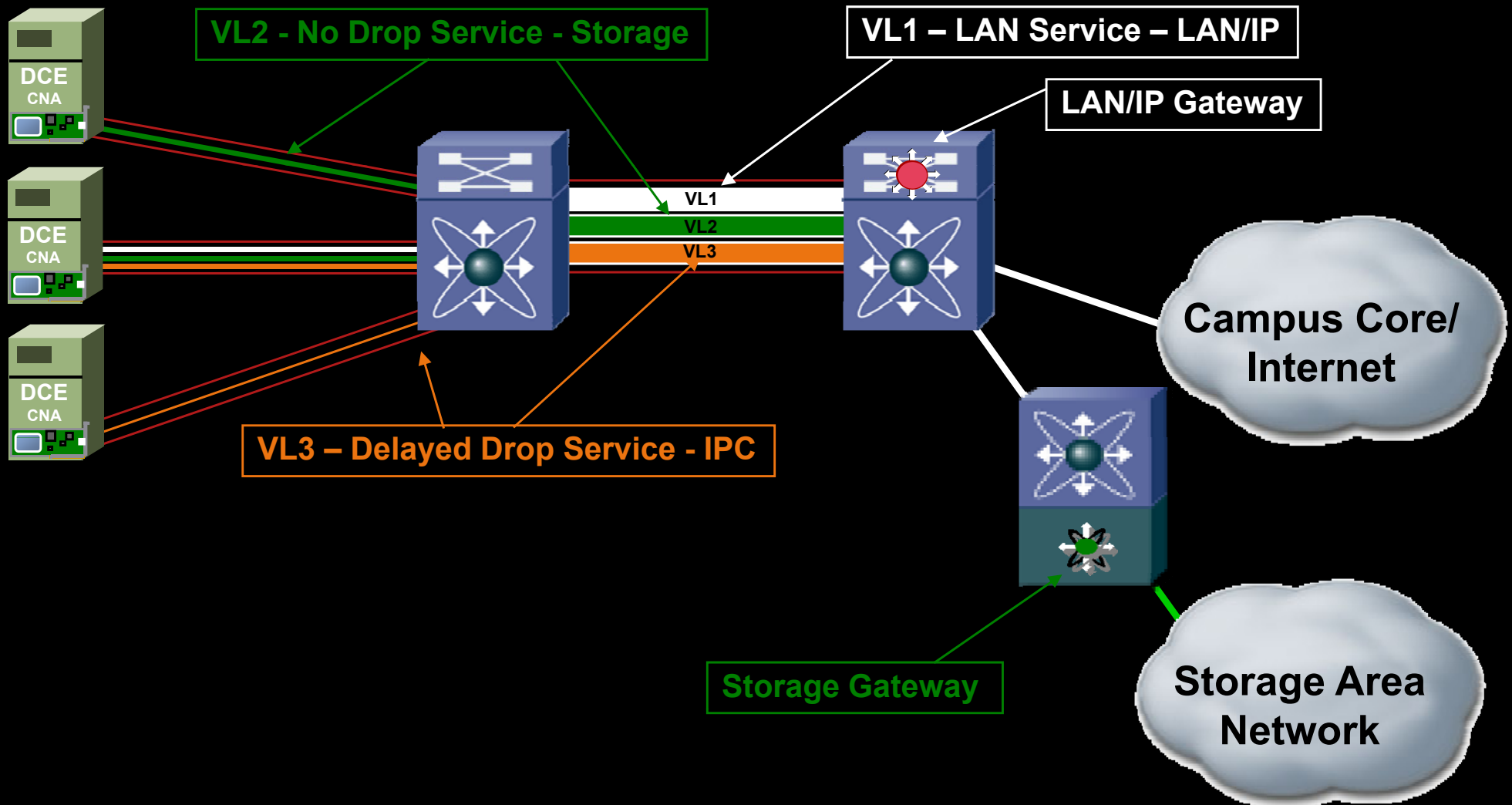


- Enables lossless Fabrics for each class of service
- PAUSE sent per virtual lane when buffers limit exceeded
- Network resources are partitioned between VL's (E.g. input buffer and output queue)
- The switch behavior is negotiable per VL

Data Center Ethernet Features (VL)

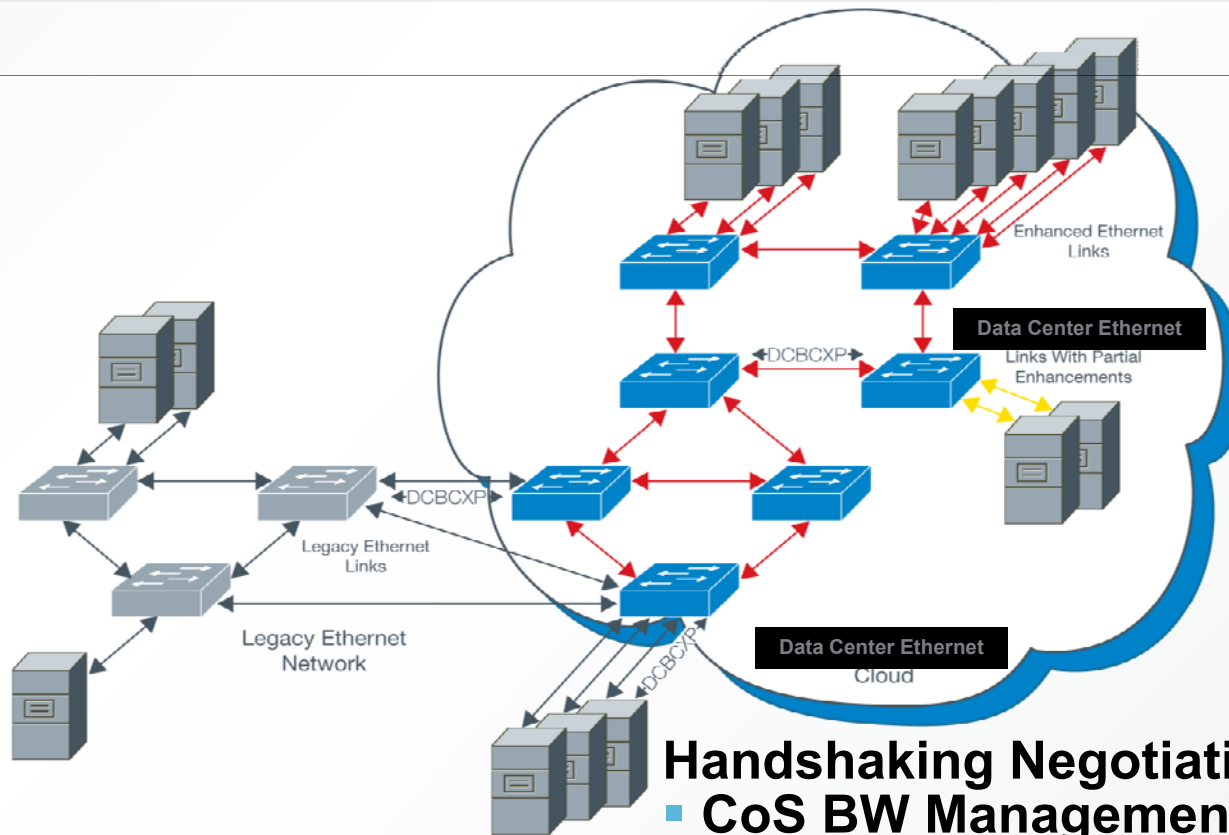
An example...

Up to 8 VL's per physical link
Ability to support QoS queues within the lanes



Data Center Ethernet Features (DCBX)

Data Center Bridging eXchange

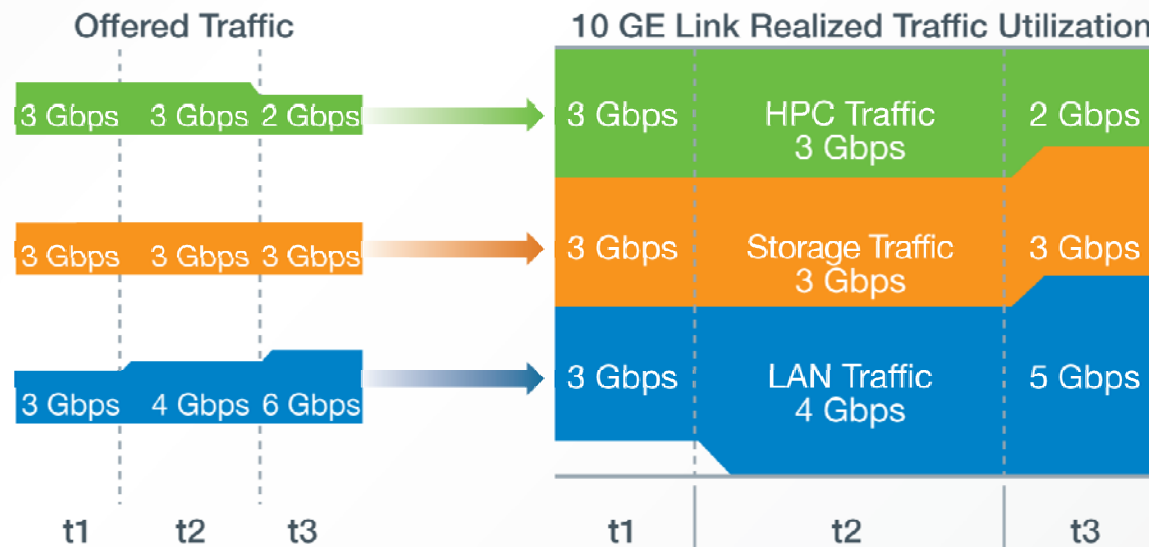


Handshaking Negotiation for:

- CoS BW Management
- Class Based Flow Control
- Congestion Management (BCN/QCN)
- Application (user_priority usage)
- Logical Link Down

Data Center Ethernet Features (ETS)

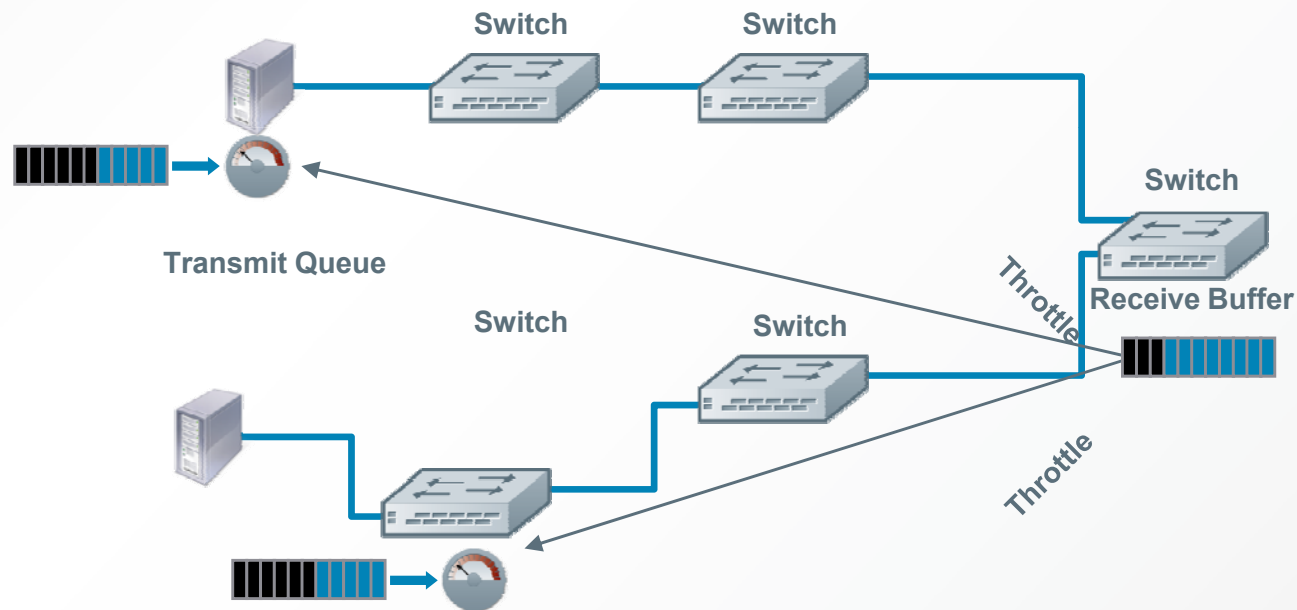
Enhanced Transmission Selection (ETS)



- Enables Intelligent sharing of bandwidth between traffic classes control of bandwidth
- Being Standardized in IEEE 802.1Qaz
- Also known as Priority Grouping

Data Center Ethernet Features (BCN/QCN)

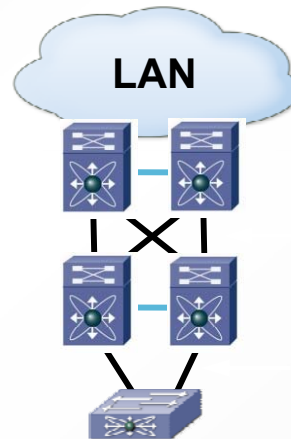
Congestion Management



- Moves congestion out of the core to avoid congestion spreading
- Allows End-to-End congestion management
- Standards track in 802.1Qau

Data Center Ethernet Features (L2MP)

Layer 2 Multi-Pathing



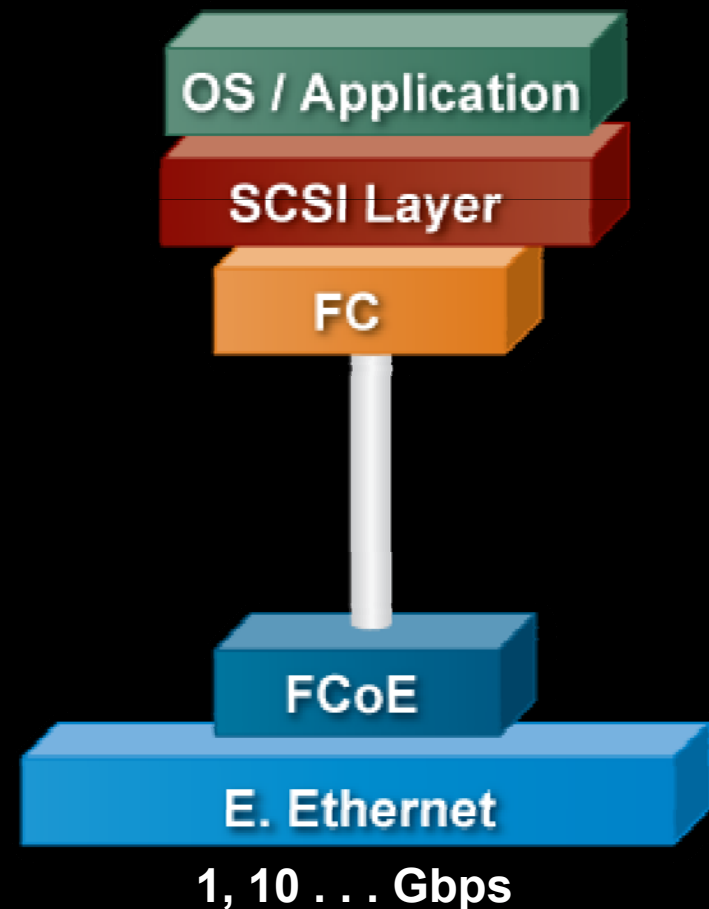
- Uses Layer 3 protocol at Layer 2
- Equal Cost Multi-Pathing
- Eliminates STP from L2 domain
- Preferred path selection
- Load Balancing across available bisectional Bandwidth



- Environmental Challenges
- What's up in the Datacenter
- DC Architecture Trends
- Fibre Channel compared to Ethernet
- I/O Consolidation
- Data Center Ethernet (DCE)
- Fibre Channel over Ethernet (FCoE)
- Unified I/O Use Case
- Summary

What is Fibre Channel over Ethernet?

- From a Fibre Channel standpoint it's
FC connectivity over a new type of cable called... an Ethernet cloud
- From an Ethernet standpoints it's
Just another ULP (Upper Layer Protocol) to be transported, but... a challenging one!
- And technically...

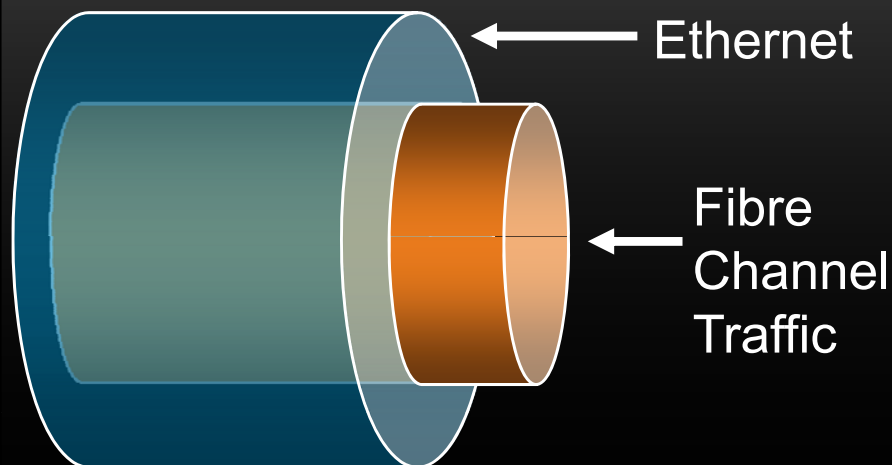


FCoE is an extension of Fibre Channel onto a Lossless Ethernet fabric

FC over Ethernet (FCoE)

FCoE

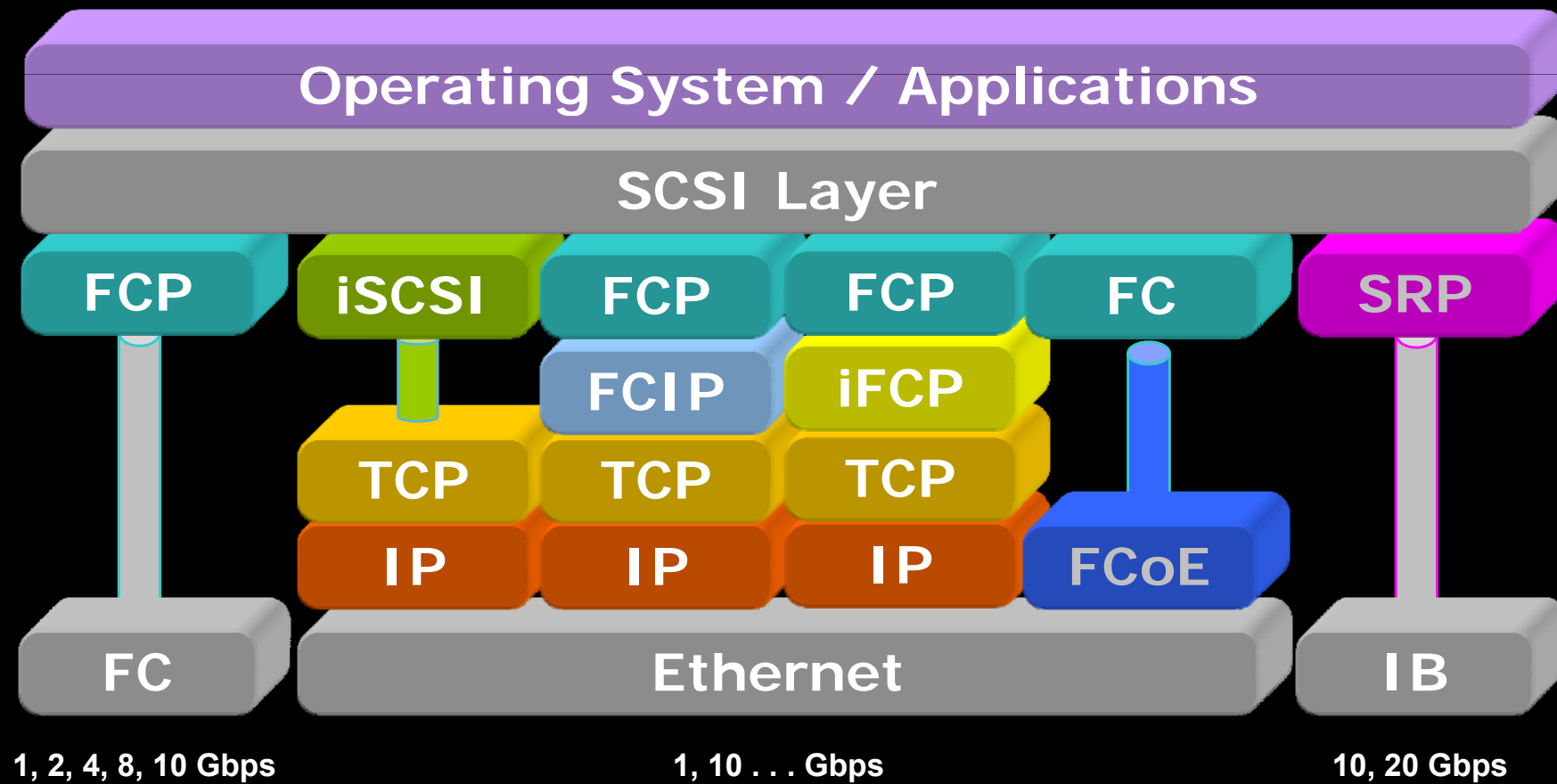
- Mapping of FC Frames over Ethernet
- Enables FC to Run on a Lossless Ethernet Network



Benefits

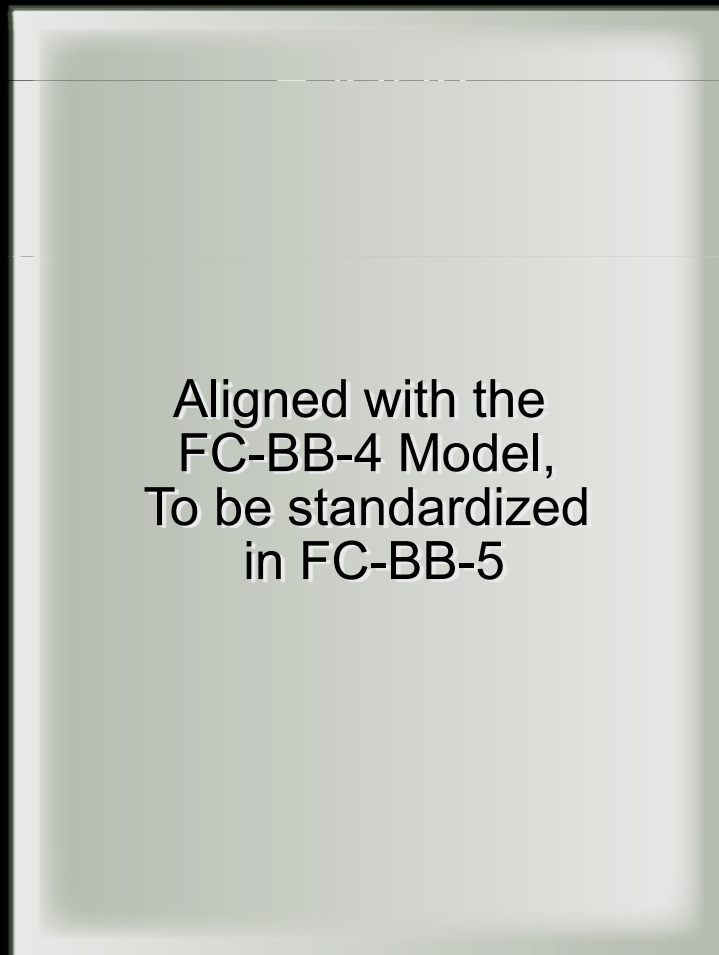
- Fewer Cables
 - Both block I/O & Ethernet traffic co-exist on same cable
- Fewer adapters needed
- Overall less power
- Interoperates with existing SAN's
- No Gateway - Stateless

FC over Ethernet (FCoE)



FCoE Advantages

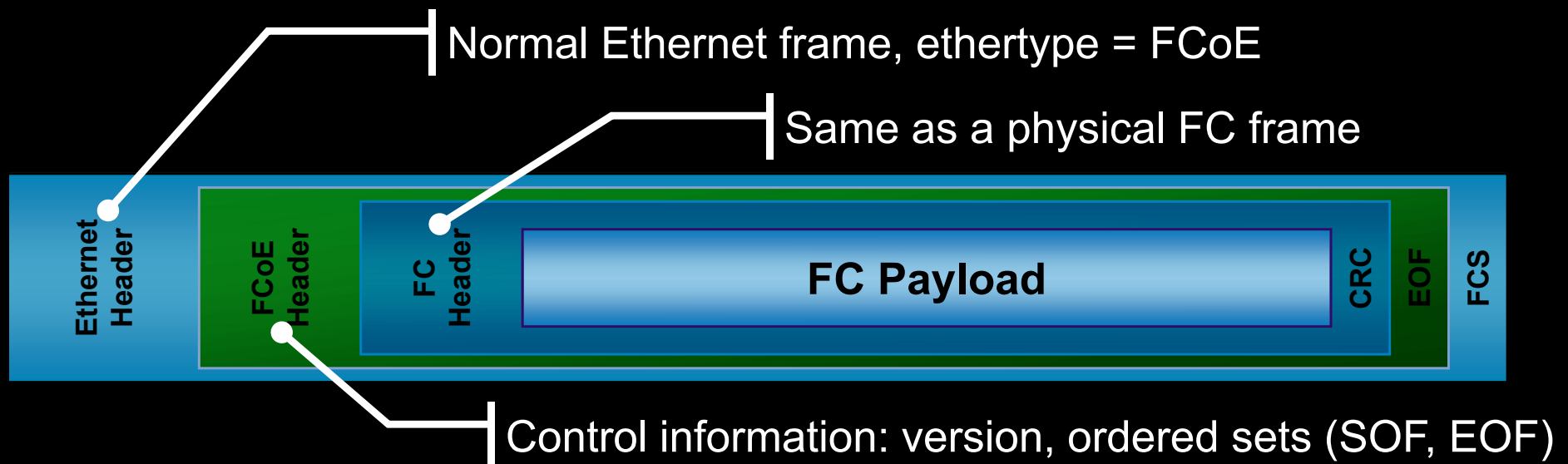
FCoE is managed like FC at initiator, target, and switch level



- Completely based on the FC model
- Same host-to-switch and switch-to-switch behavior of FC
- E.g., in order delivery or FSPF load balancing
- WWNs, FC-IDs, hard/soft zoning, DNS, RSCN

FCoE Enablers

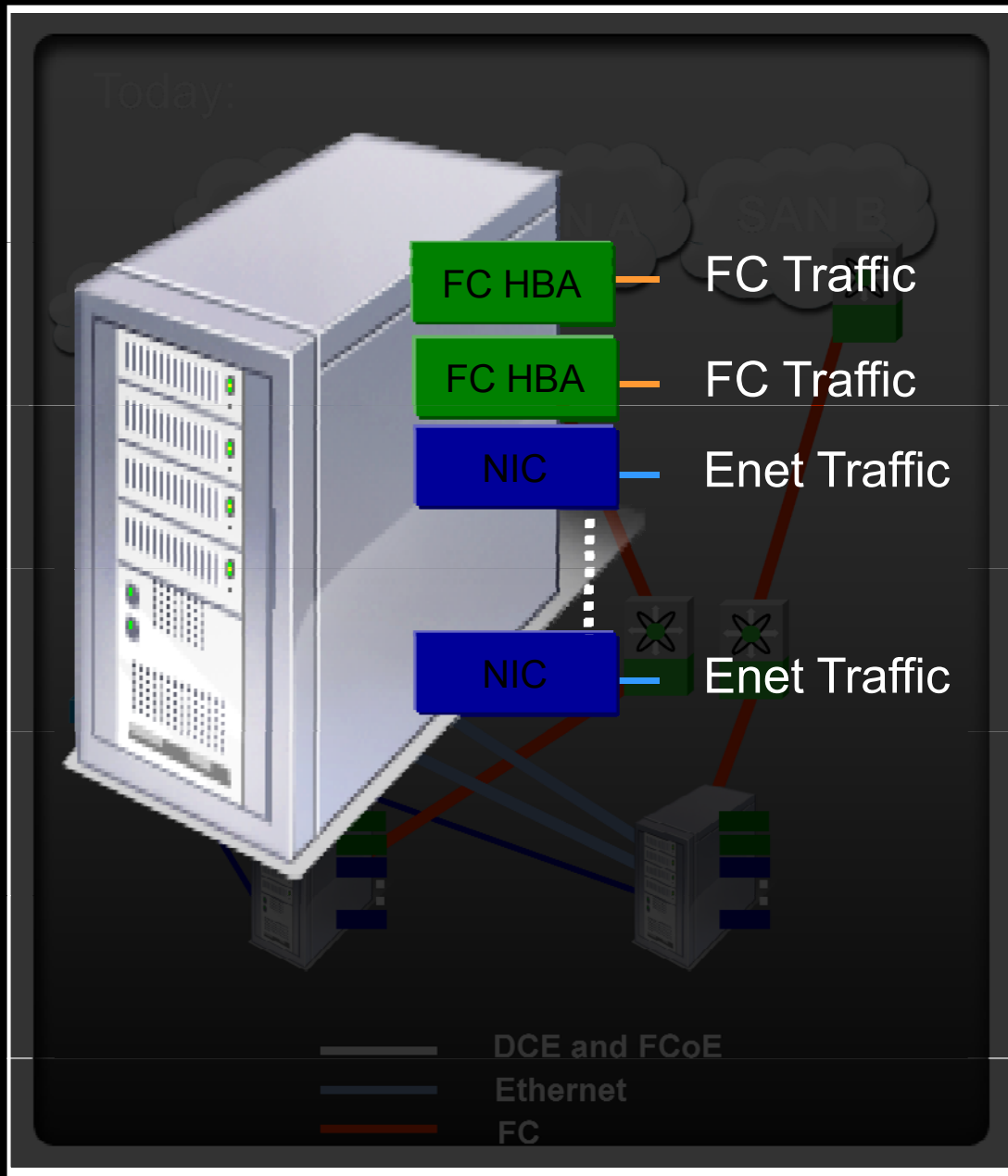
- 10Gbps Ethernet
- Lossless Ethernet
 - Matches the lossless behavior guaranteed in FC by B2B credits
- Ethernet jumbo frames
 - Max FC frame = 2112 bytes





- Environmental Challenges
- What's up in the Datacenter
- DC Architecture Trends
- Fibre Channel compared to Ethernet
- I/O Consolidation
- Data Center Ethernet (DCE)
- Fibre Channel over Ethernet (FCoE)
- Unified I/O Use Case
- Summary

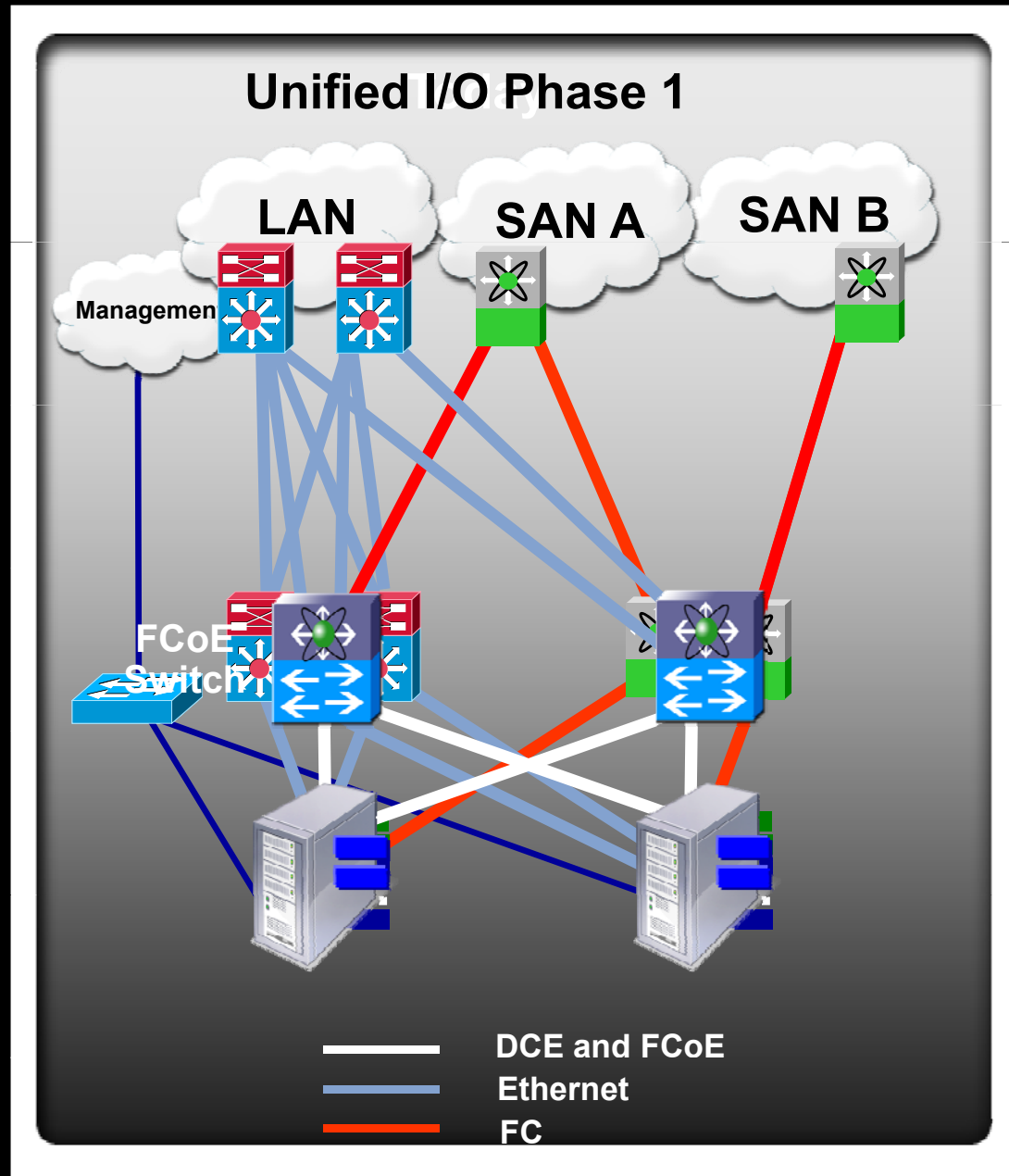
Unified I/O Use Case



Today:

- 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 – firmware, driver-patching, versioning

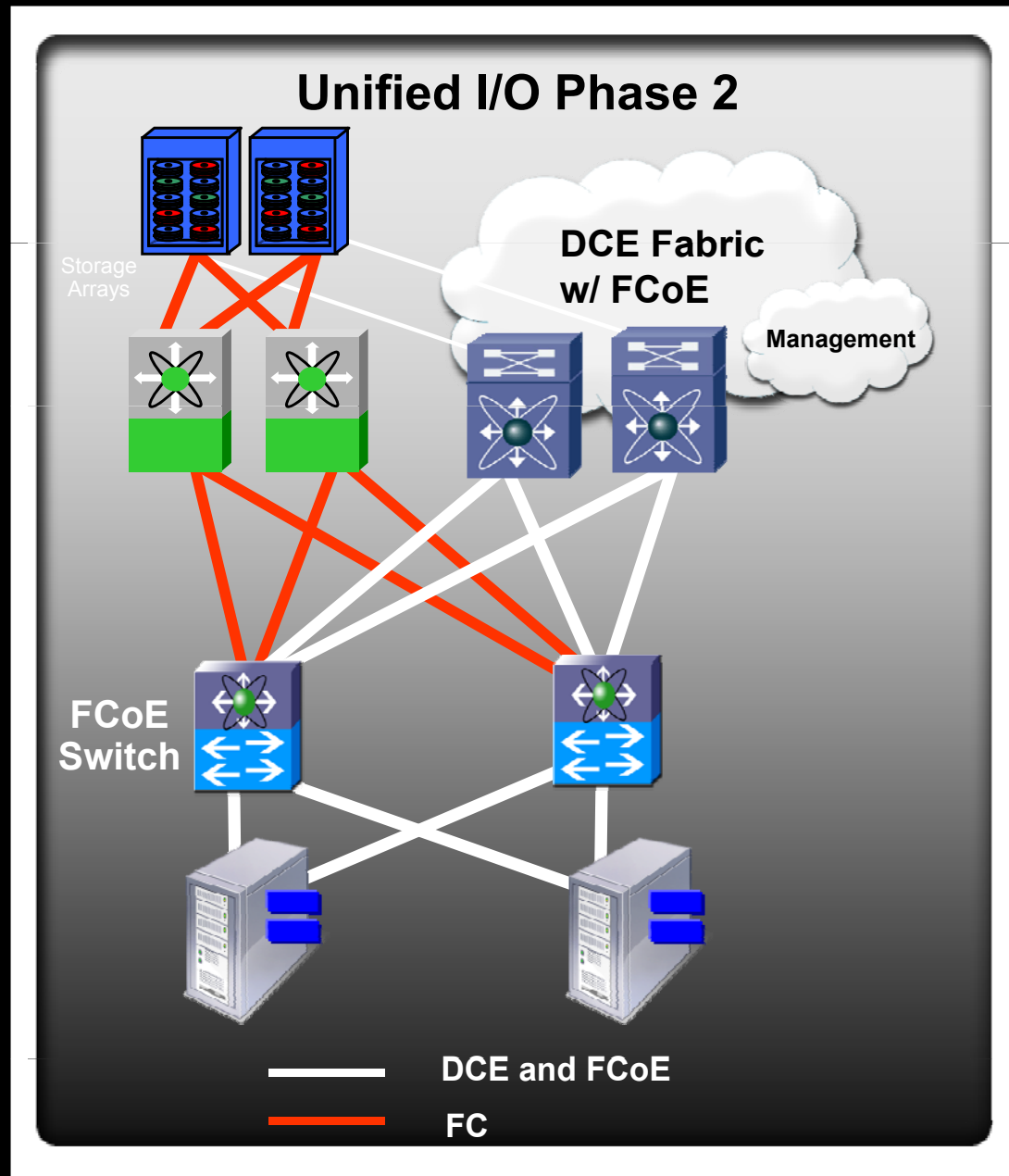
Unified I/O Use Case



Today Unified I/O Phase 1

- Parallel LAN/SAN Infrastructure
- Reduction of server adapters
- Inefficient use of Network Infrastructure
- Simple configuration of access layer & cabling
- Gateway for server implementation might be in installed base of existing LAN and SAN adapter and cabling costs
- L2 Multipathing Access – Distribution Adds downstream port costs,
- Low capex and op-ex
- Fewer Cables
- Each connection adds additional points of failure in the fabric
- Investment Protection (LANs and SANs)
- Consistent Operations and Provisioning
- Multiple fault domains – complex diagnostics
- Management complexity – firmware, driver-patching, versioning

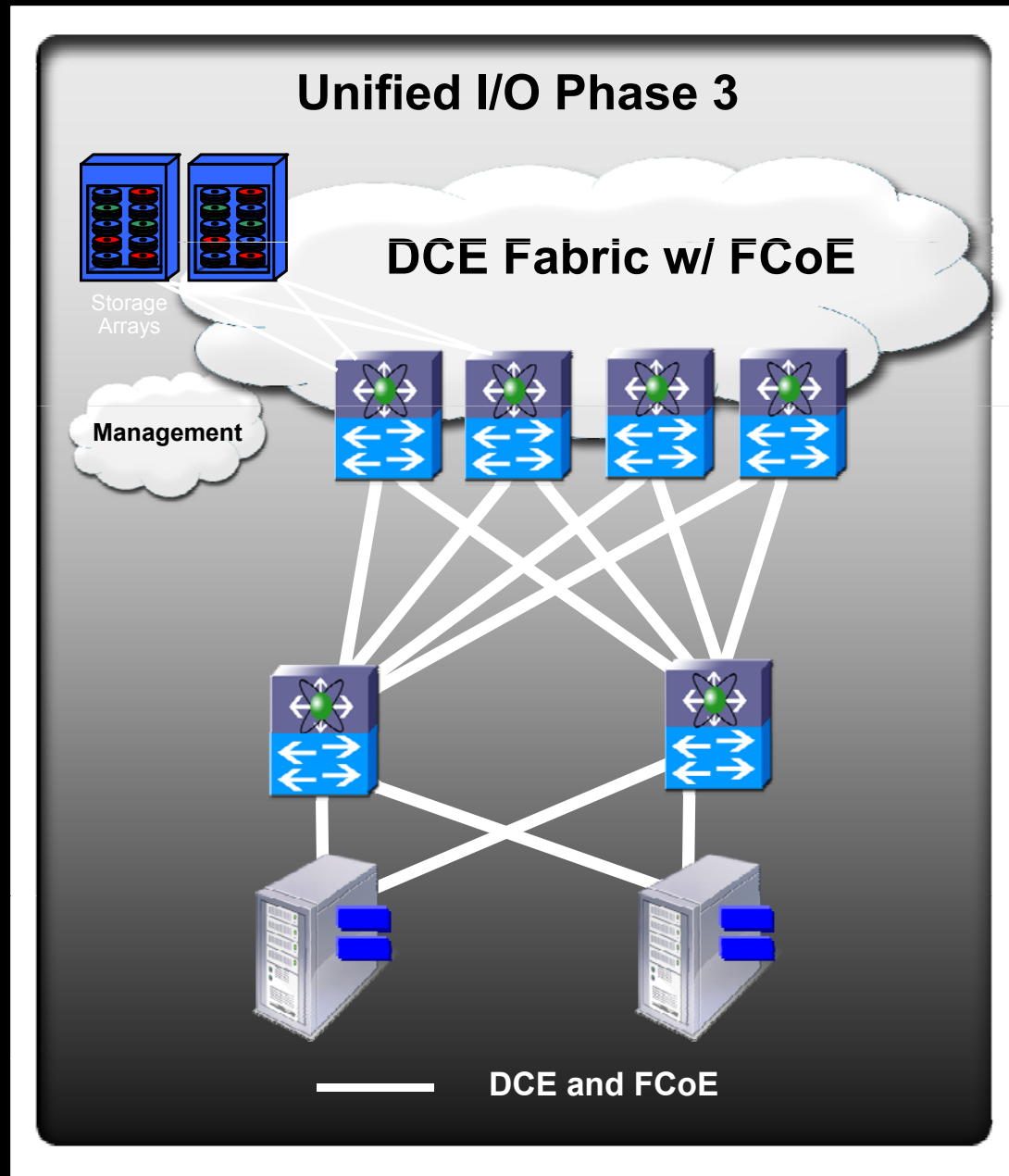
Unified I/O Use Case



Unified I/O Phase 2

- Elimination of parallel network infrastructure
- L2/L3 Multipathing end to end
- Faster infrastructure provisioning
- Lower TCO
- Disk array access via DCE or Native FC

Unified I/O Use Case



Unified I/O Phase 3

- Datacenter wide Unified Fabric for LAN and SAN
- L2/L3 Multipathing end to end
- Consistent network policies across datacenter
- Lower TCO



Cisco Expo
2008

Summary



Data Center 3.0 Infrastructure Portfolio

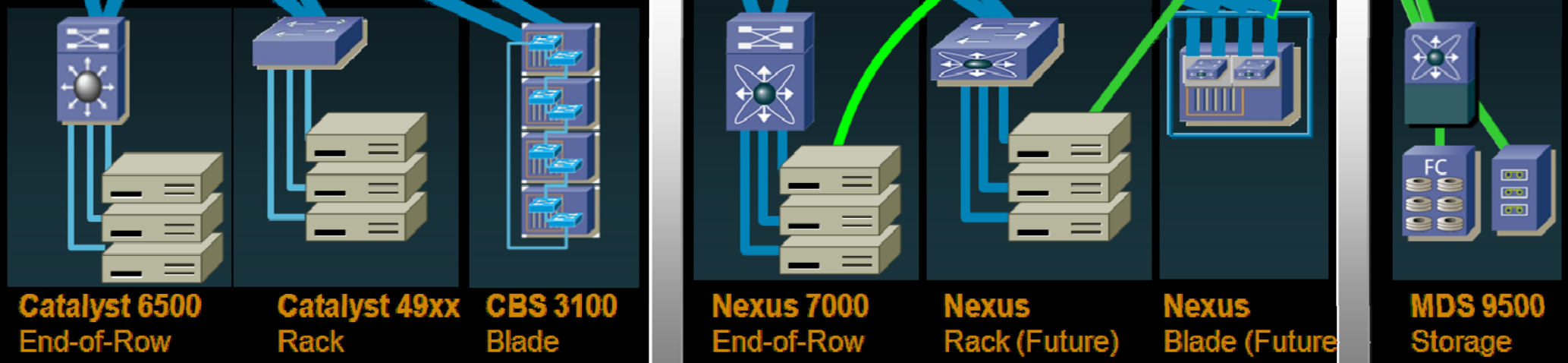
DC Core



DC Aggregation



DC Access



1Gb Server Access

10Gb Server Access

Storage

Data Center 3.0 Infrastructure Portfolio

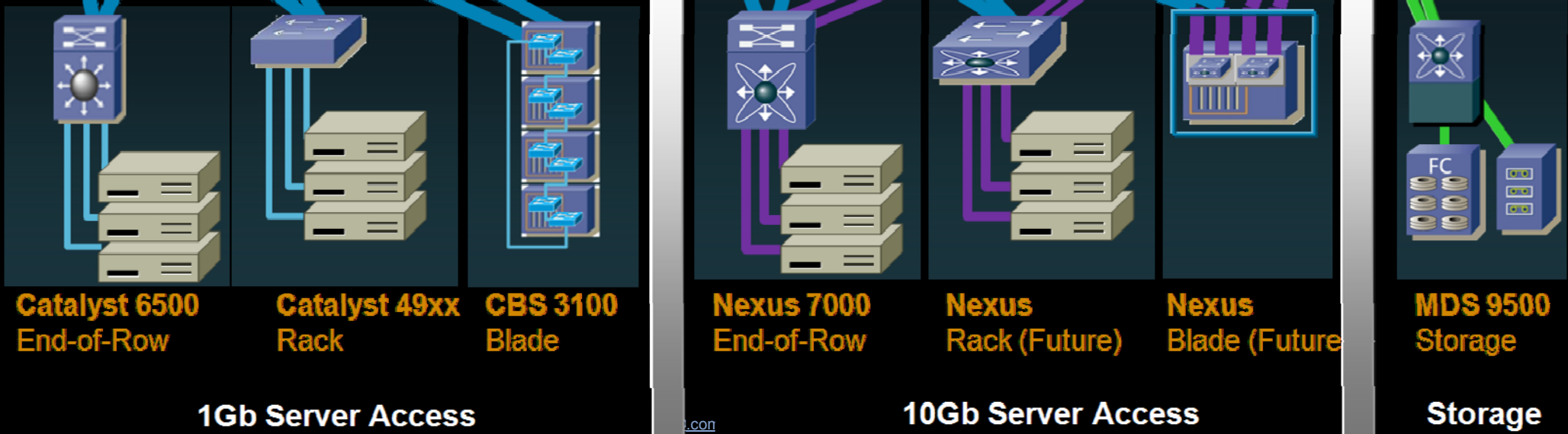
DC Core



DC Aggregation



DC Access



Q and A



Complete Your Online Session Evaluation

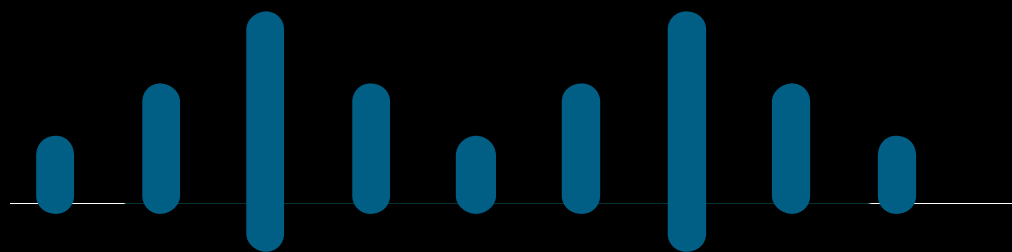
Please complete the online evaluation under

www.cisco.at/expo2008/feedback

The first 100 to complete the survey will receive a copy of Don Tapscott's book "Wikinomics".

We very much appreciate and value your feedback, many thanks!





CISCO