Unifikovaná komunikační infrastruktura datového centra

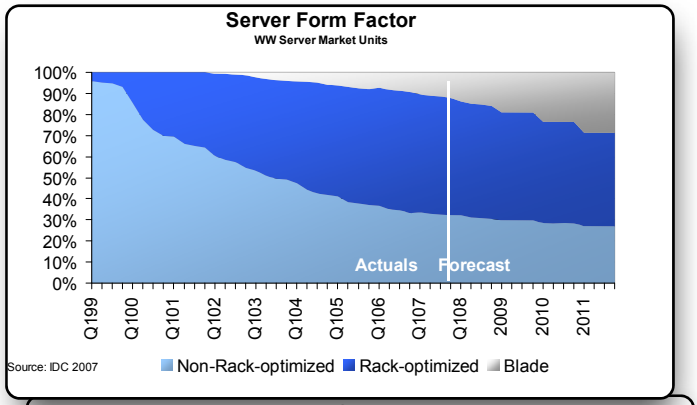
Tomáš Michaeli

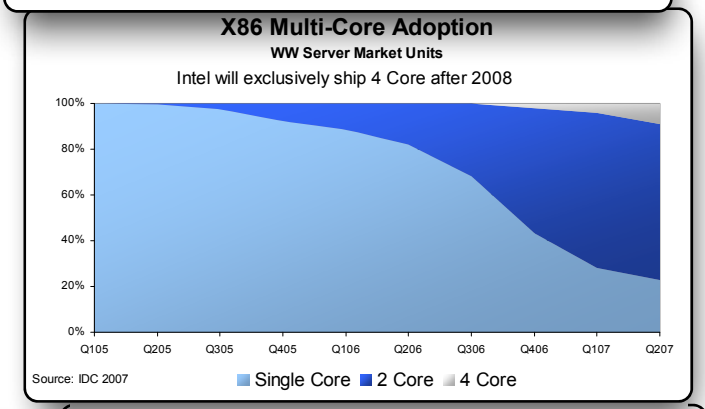
Consulting System Engineer, Data Center/Storage Networking

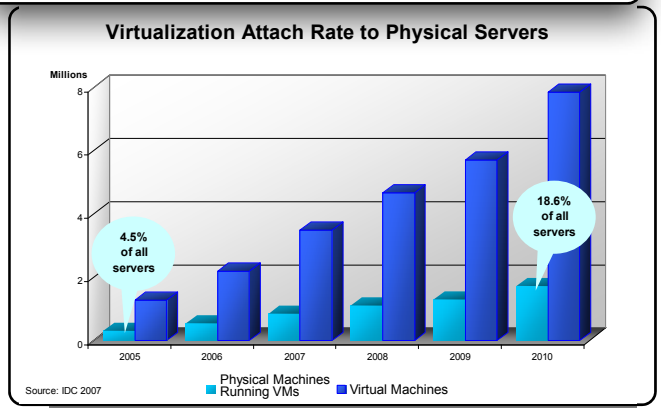
tomichae@cisco.com

Bandwidth drivers in the Data Center

- Servers moving to dense rack chassis
- Rapid Adoption of multi-core
  - Post 2008 Intel will ship exclusively 4+ cores servers
- Growth of virtualization exceeds growth of physical servers
- Power, cooling and cabling cost
- 10GbE, 8GbFC → Unified fabric

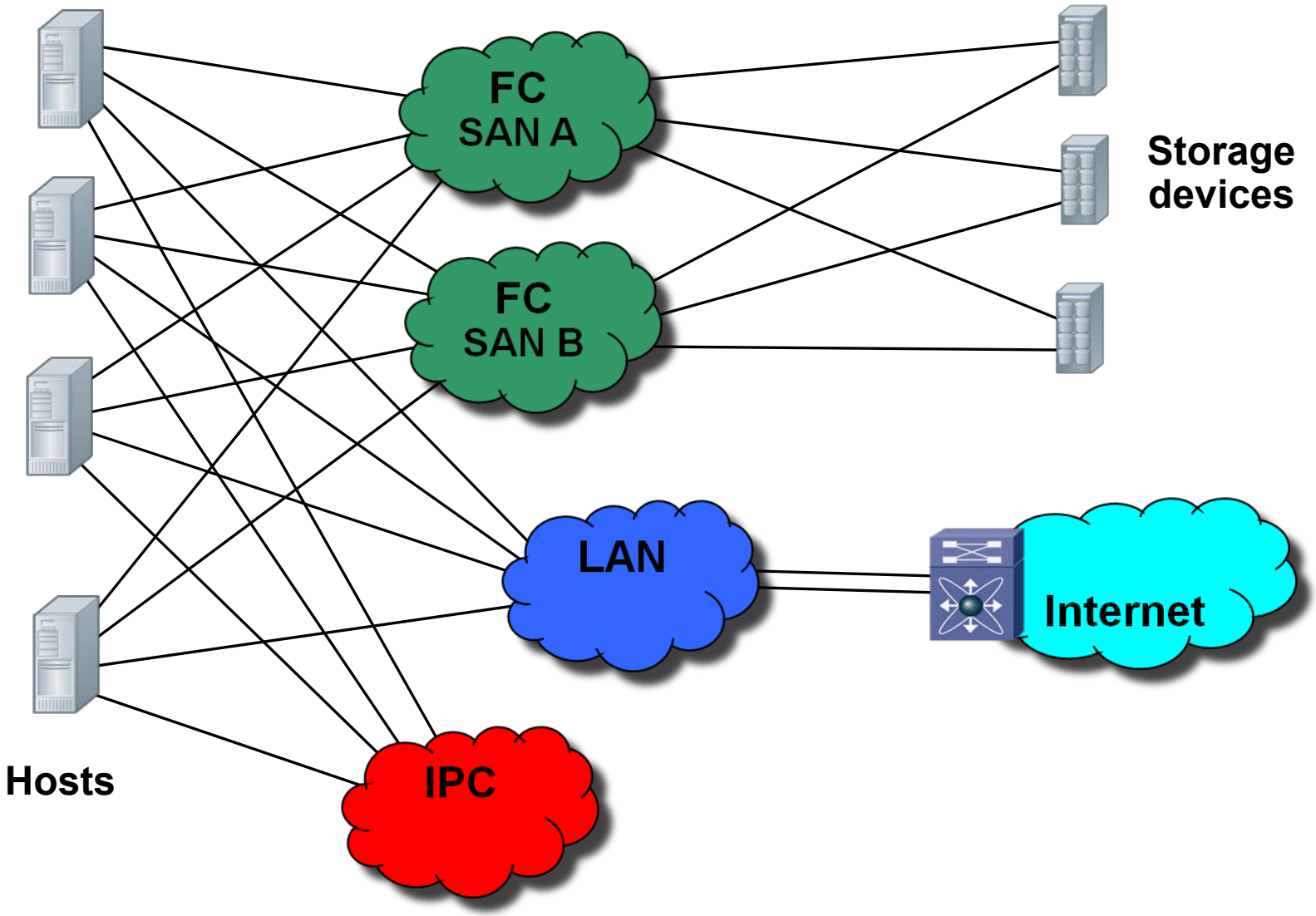




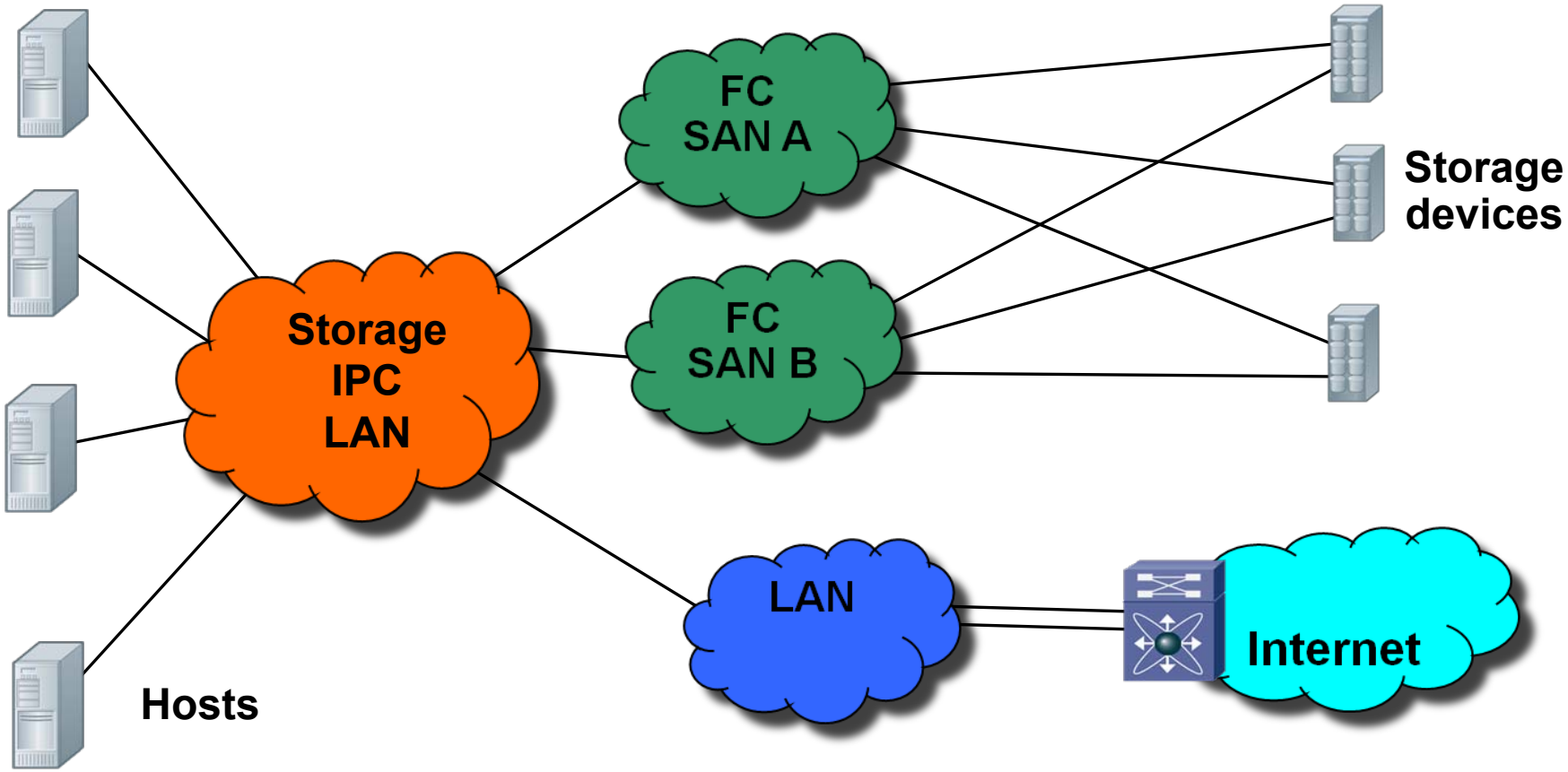


Multi-Core CPUs and Server Virtualization driving the demand for higher bandwidth network connections

Current Data Center Structure

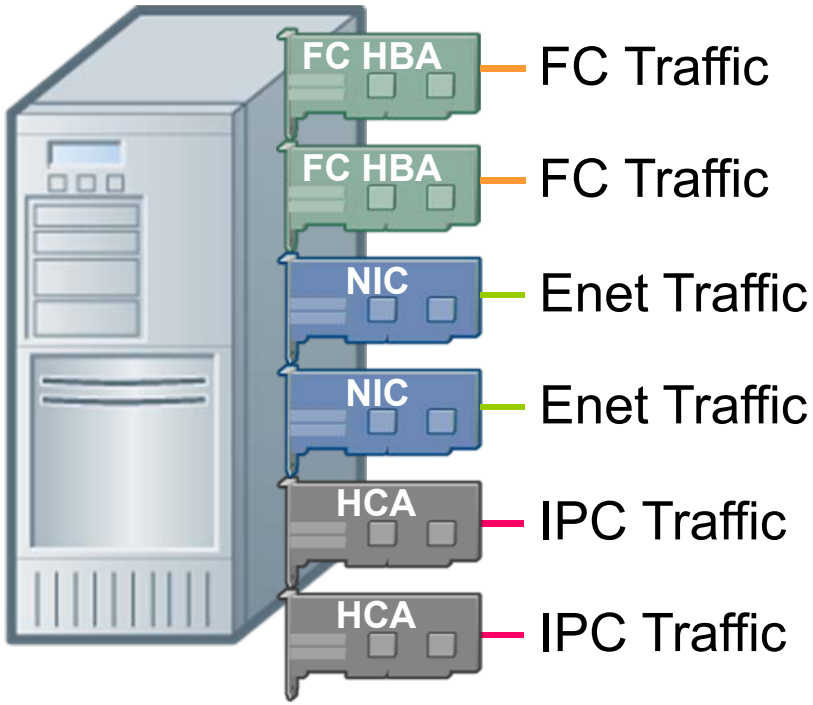


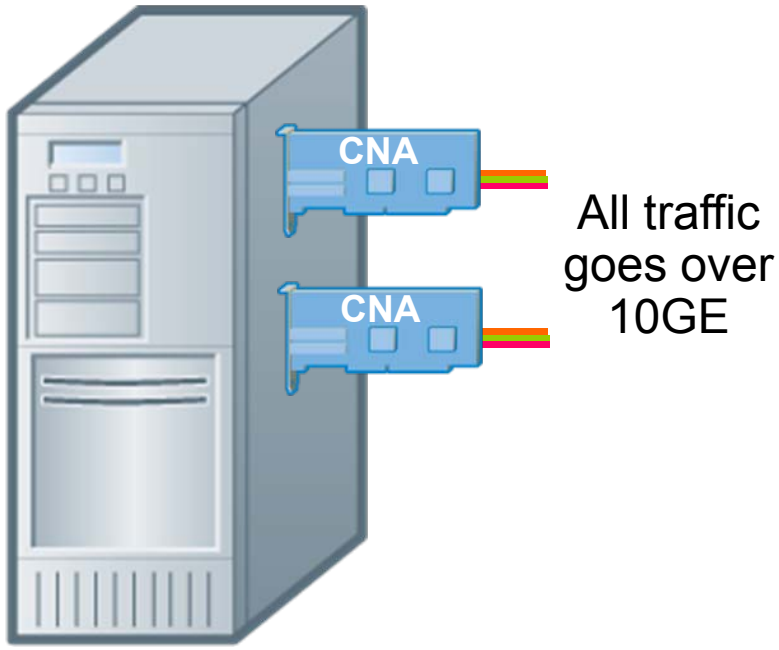
Pragmatic Consolidated Data Center Structure



I/O Consolidation – in the Host

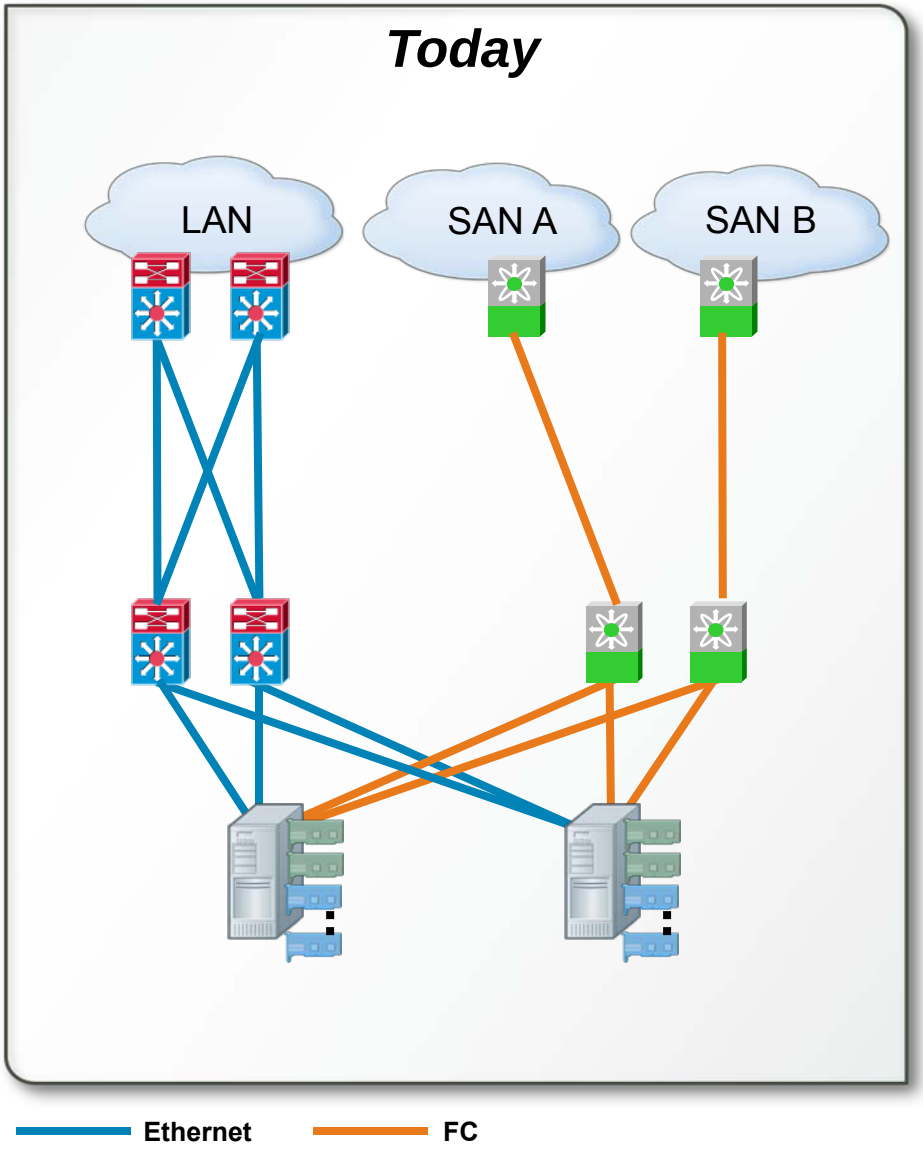
- Fewer CNAs (Converged Network Adapters) instead of NICs, HBAs and HCAs
- Less power consumption, less cables, better cooling





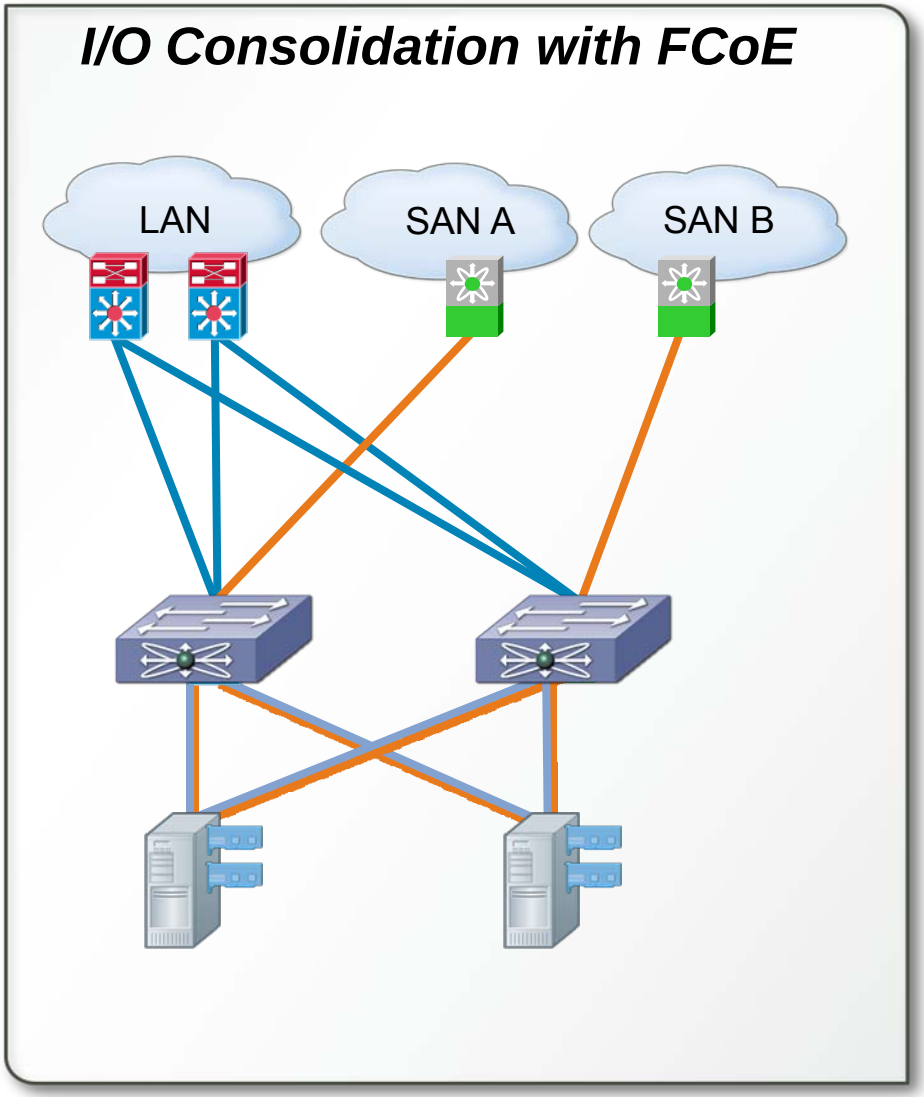
I/O Consolidation – In the network

- 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
- Management complexity



I/O Consolidation - Phase 1

- Reduction of server adapters
- Simplification of access layer & cabling
- Gateway free implementation
- L2 Multipathing Access – Distribution
- Lower TCO
- Investment Protection (LANs and SANs)
- Consistent Operational Model



I/O Consolidation Requirements

- Ethernet based
  - Lot of investment in data centers
- More bandwidth
  - PCI Express 8x and 16x solves the bandwidth bottleneck in server buses to support 10G interfaces
- 10GE is the best candidate for I/O consolidation
  - High bandwidth, Full duplex links
  - Competitive cost
    - Optical interface cost dramatically reduced by SFP+ modules
    - Twinax cable in alternative to Cat5e/6/6a cables as low cost electric interface

10GE SFP+

Lossless Ethernet

- Ethernet may be extended to become lossless:
  - Collisions do not exist in full duplex links
    - The only ones deployed in data centers
  - Link errors are rare in data centers
    - They are present also in native Fibre Channel
  - Congestion can be dealt with extensions
    - Pause-based Flow Control Priority-based Flow Control
- Lossless Ethernet allows a direct mapping of Fibre Channel over Ethernet (FCoE)

Priority-based Flow Control (PFC)

- Also called Per-Priority Pause (PPP)
- PAUSE functionality per Ethernet priority

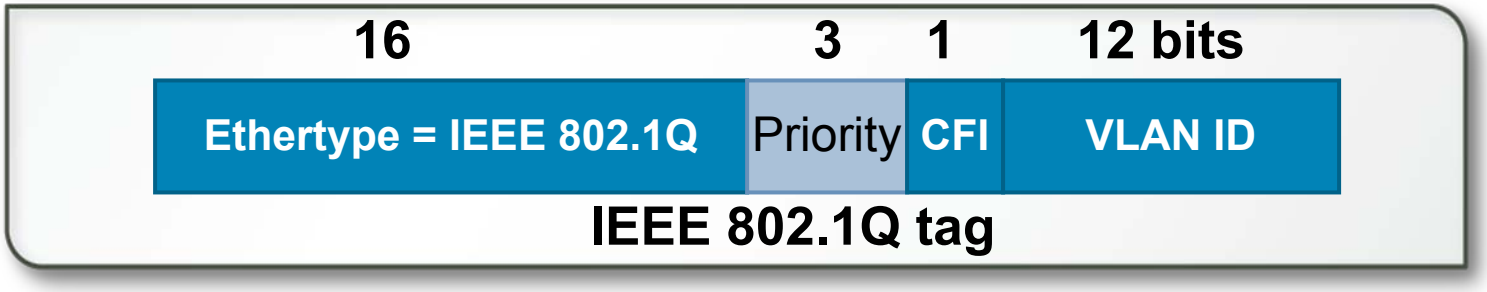
  IEEE 802.1Q defines 8 priorities

  Traffic classes are mapped to different priorities:

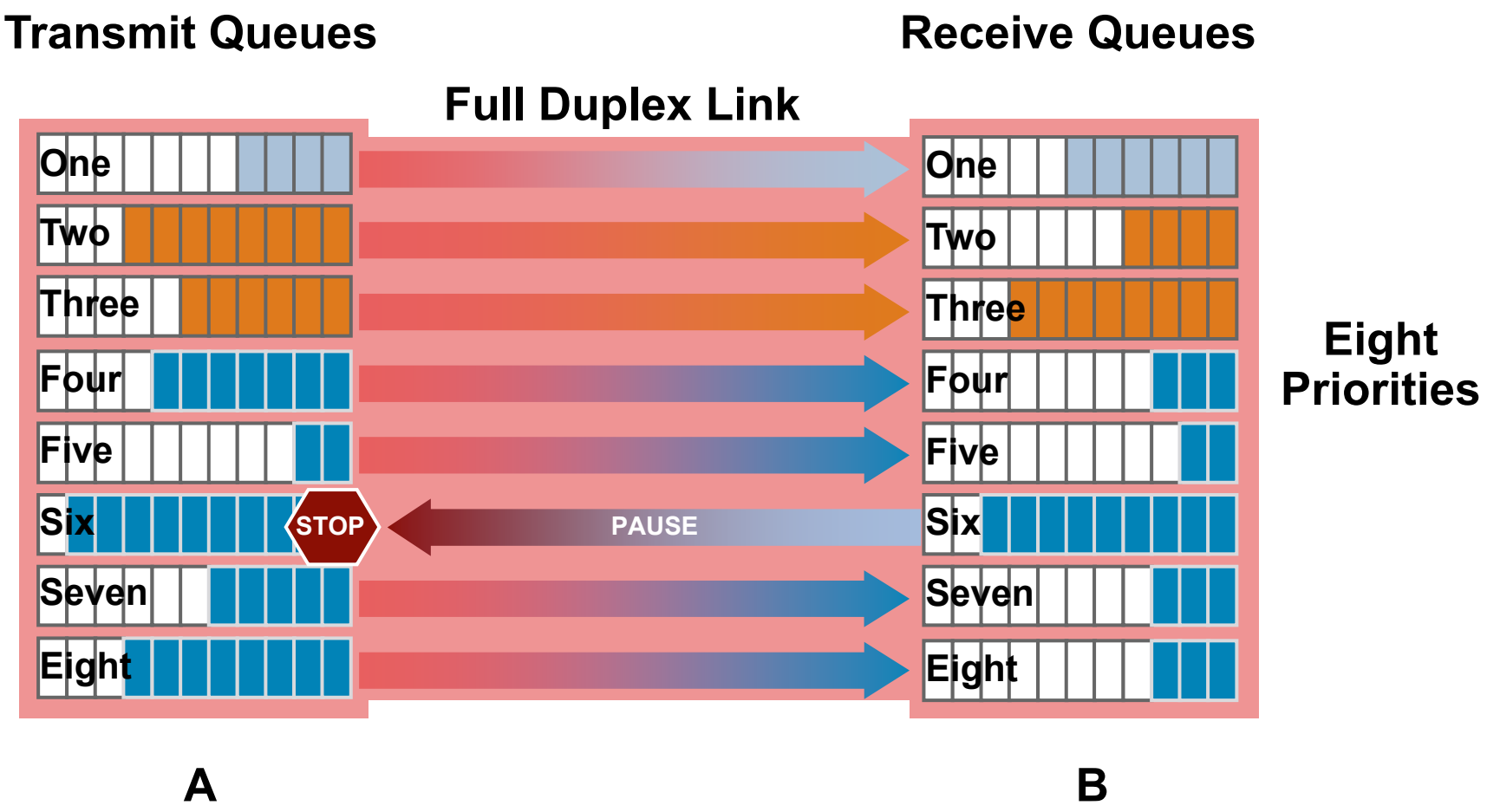
  Storage traffic may be paused while IP traffic is being forwarded or, vice versa

  Requires independent resources per priority (buffers)
- High level of industry support

  Standard Track in IEEE 802.1Qbb



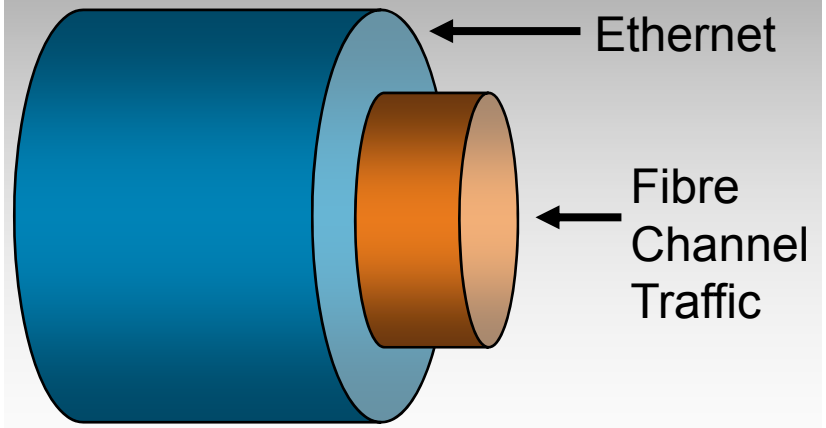
Priority-based Flow Control



FC over Ethernet (FCoE)

FCoE

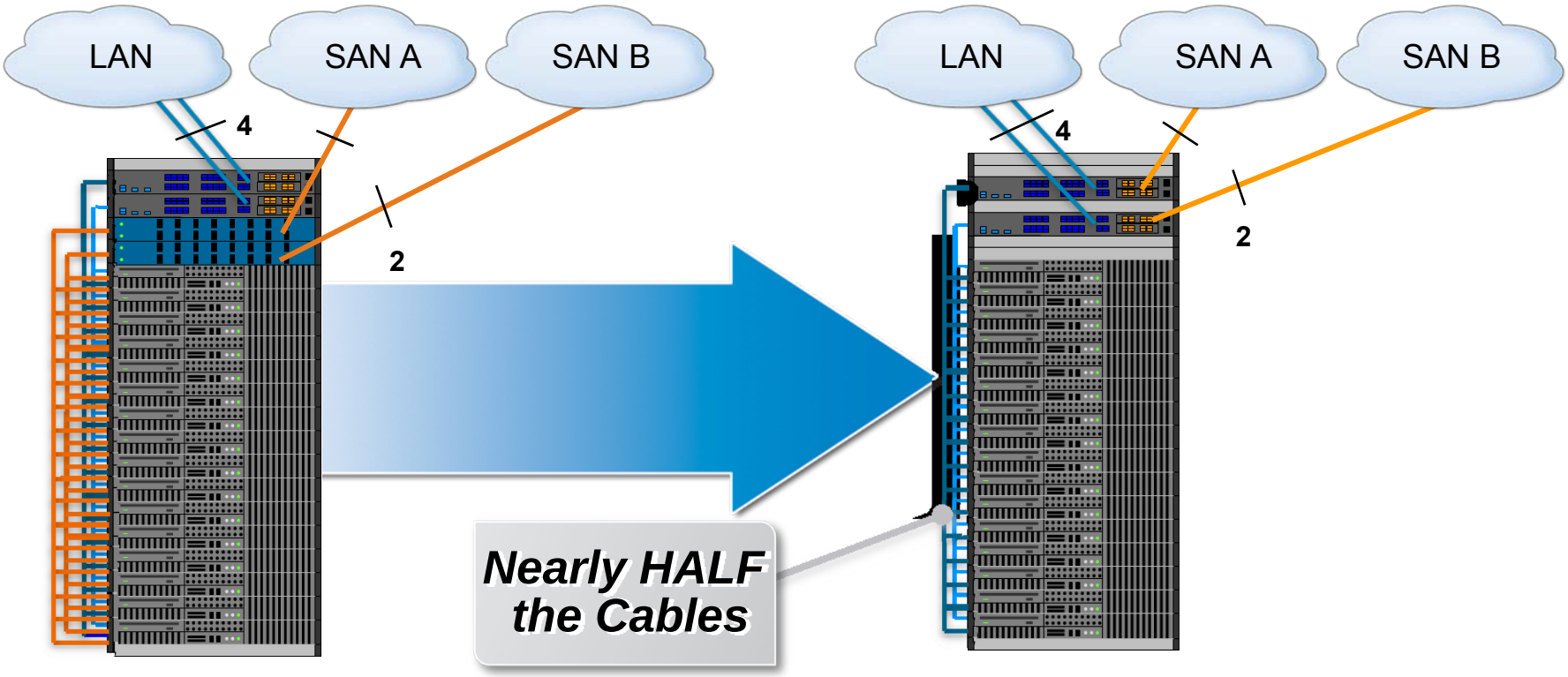
- Mapping of FC Frames over Ethernet
- Enables FC to Run on a Lossless Ethernet Network



Benefits

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

FCoE Benefits – cable, adapter reduction

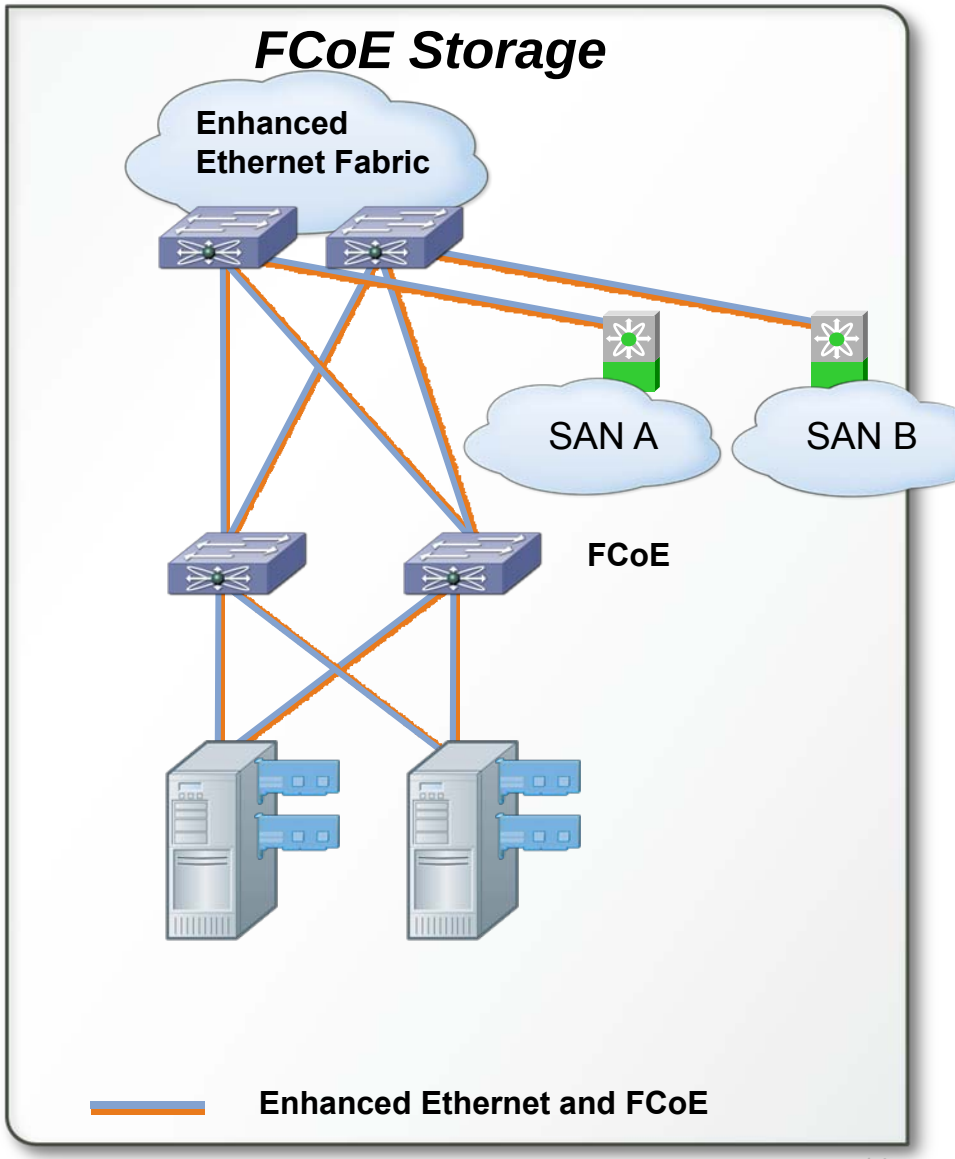


16 Servers	Enet	FC	Total
Adapters	16	16	32
Switches	2	2	4
Cables	36	36	72
Mgmt Pts	2	2	4

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

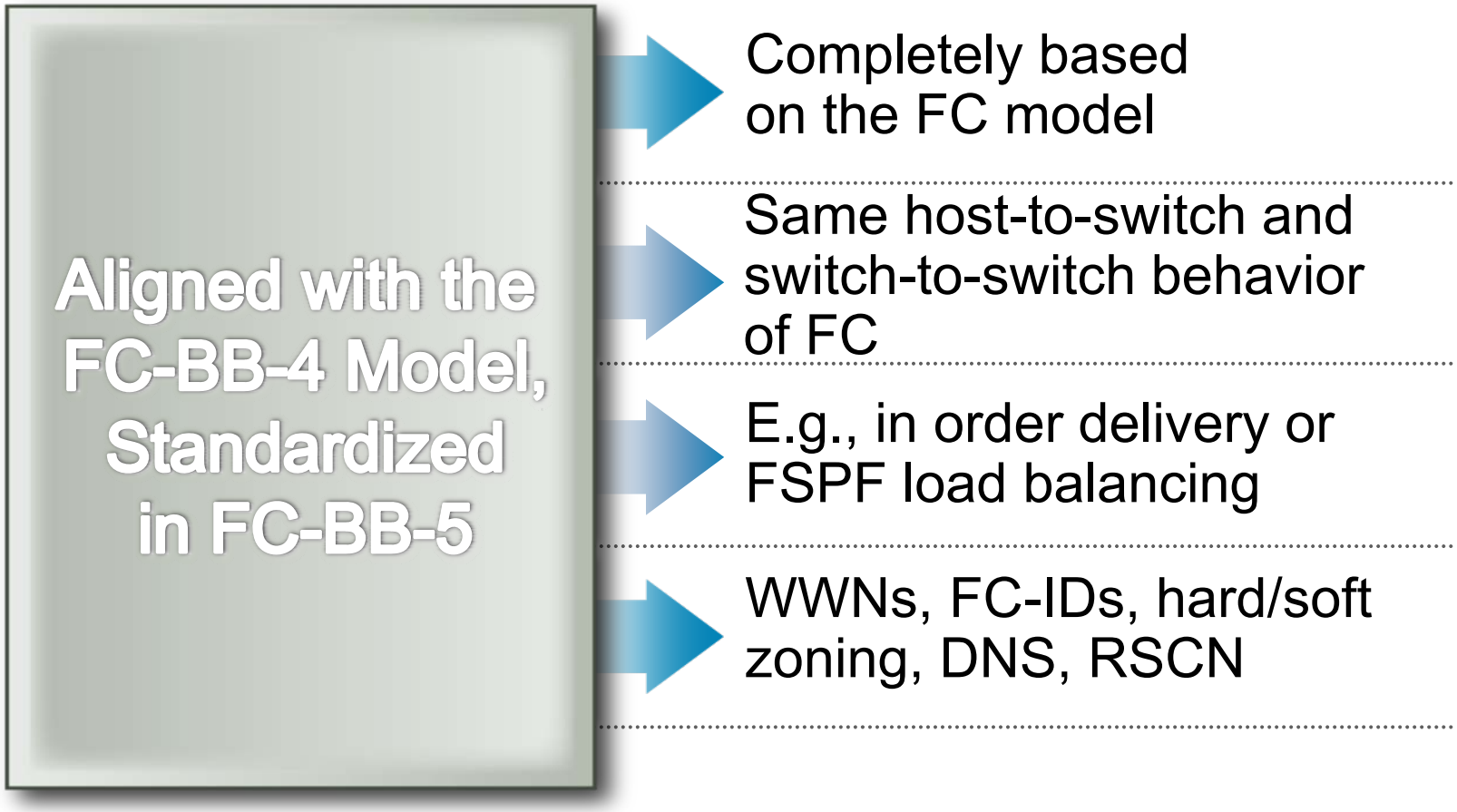
I/O Consolidation - Phase 2

- Multi hop DCE might be possible by keeping FCoE encapsulation
- Direct attached FCoE Storage
- Unified FCoE network environment
- Maintains architecture of SAN separation for high availability

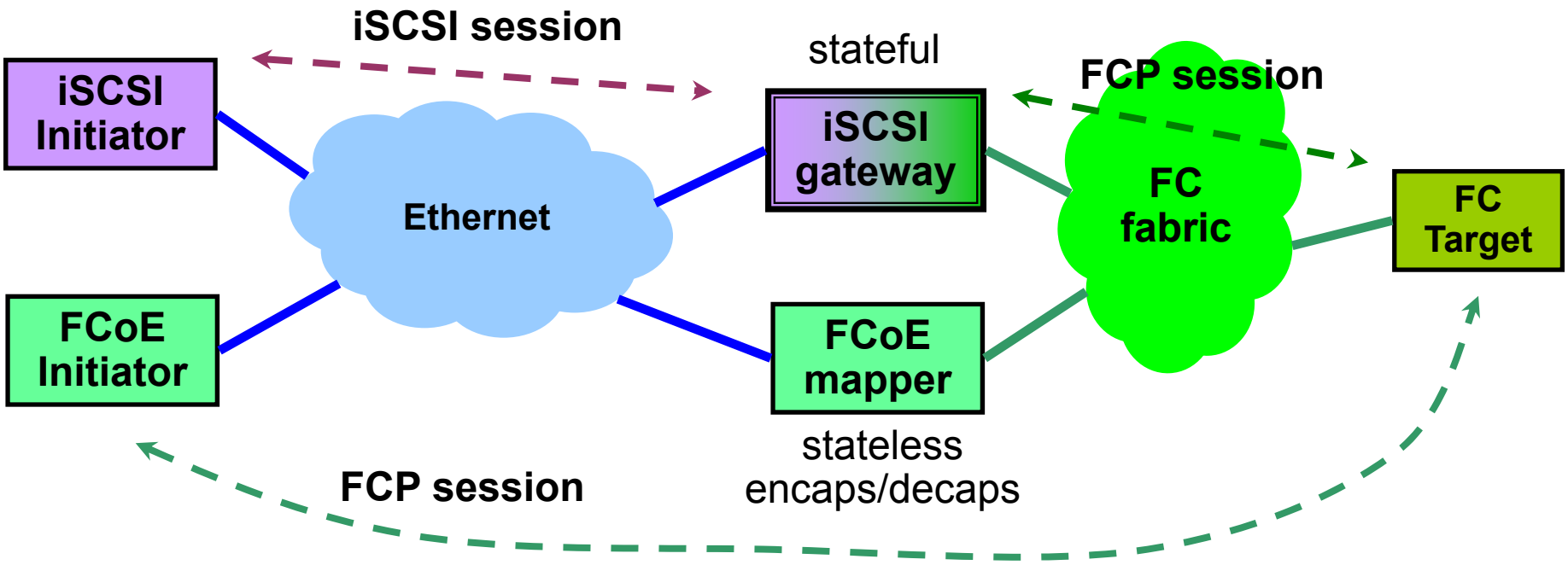


FCoE is Fibre Channel

FCoE is Fibre Channel at the host and switch level

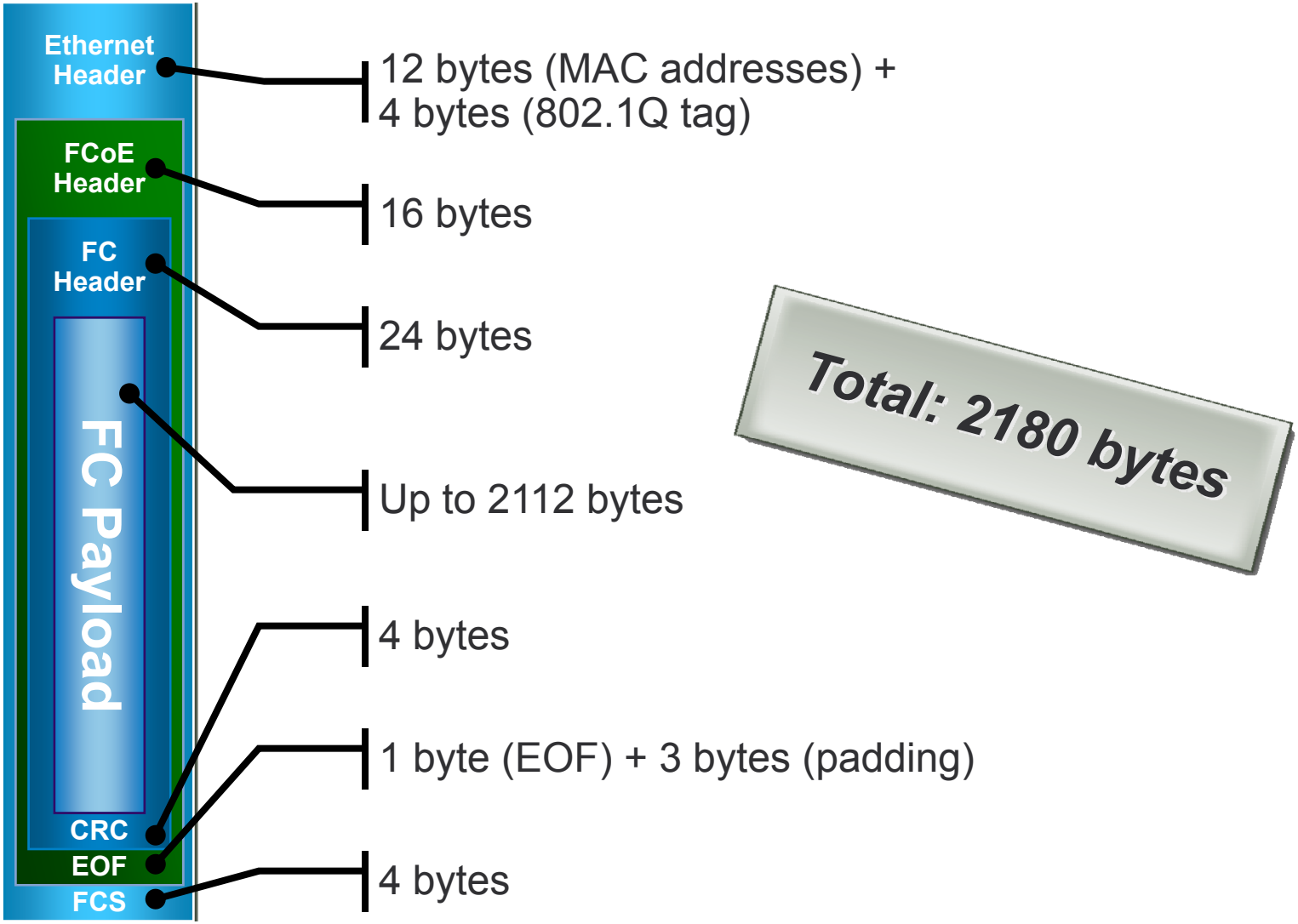


Gateway-less FCoE



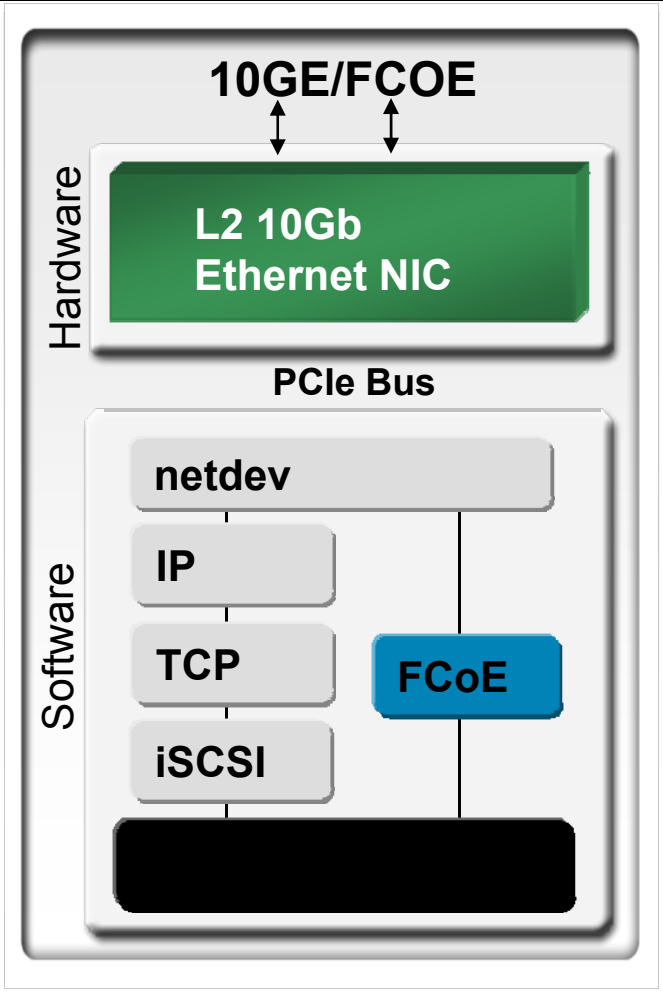
- Stateful gateway issues:
  - Single point of failure
  - Limited scalability

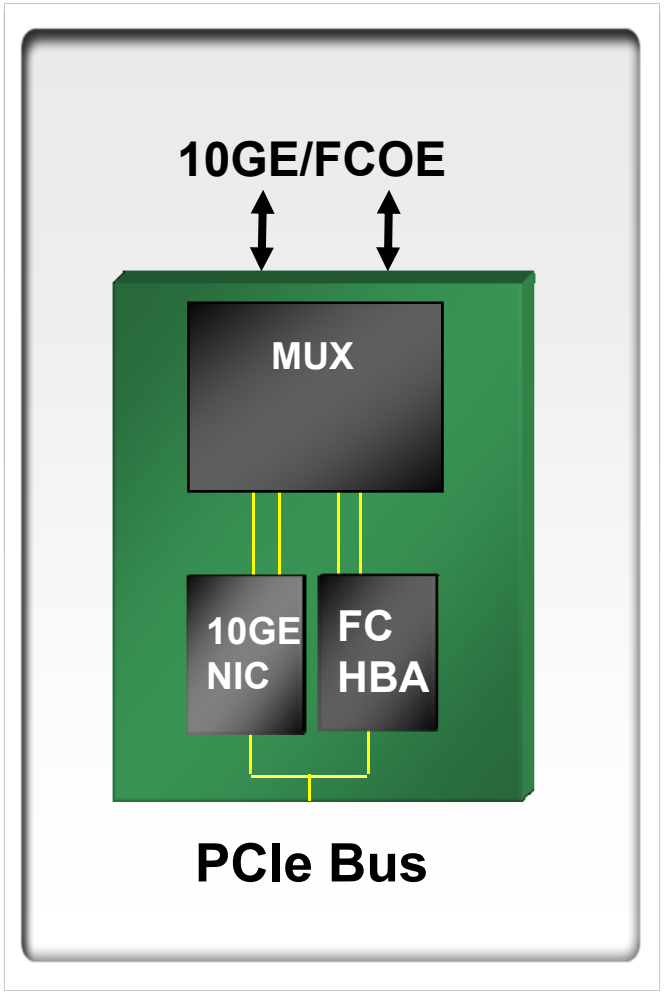
FCoE frame size



Driver Software

CNA



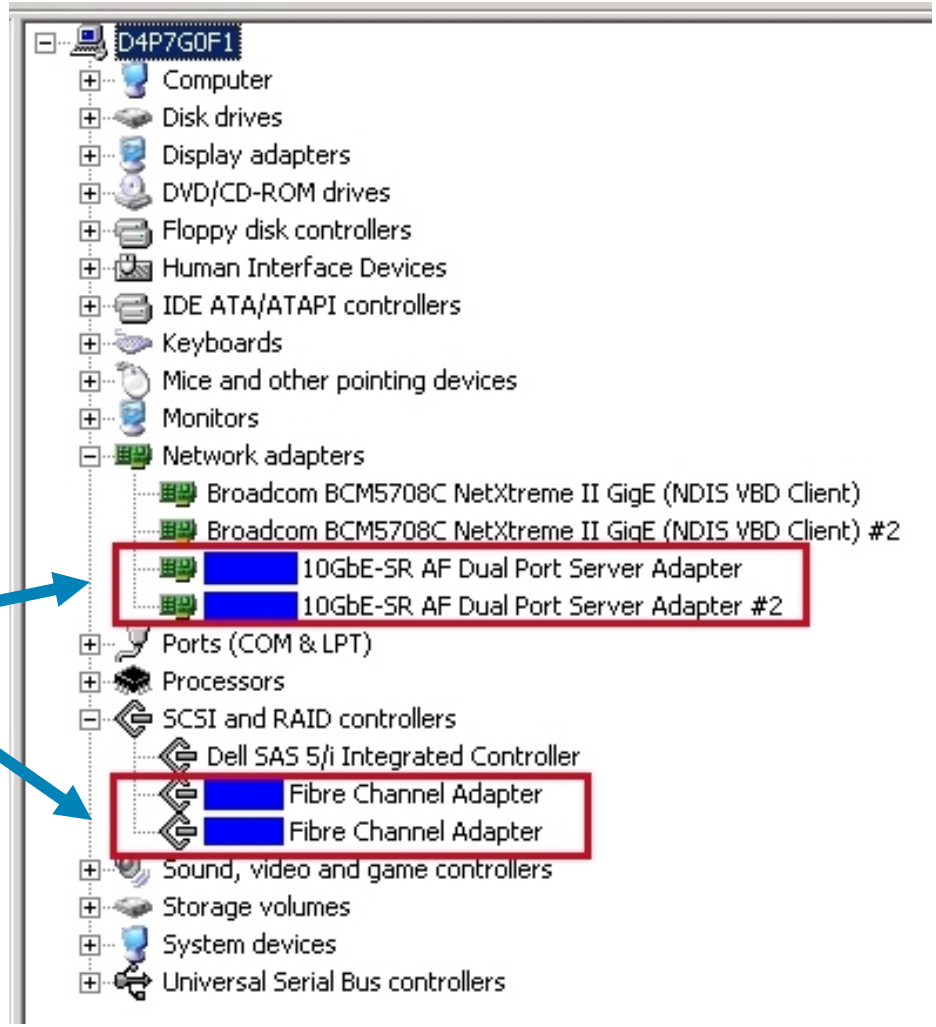


View from Operating System

- Standard drivers
- Same management
- Operating System sees:

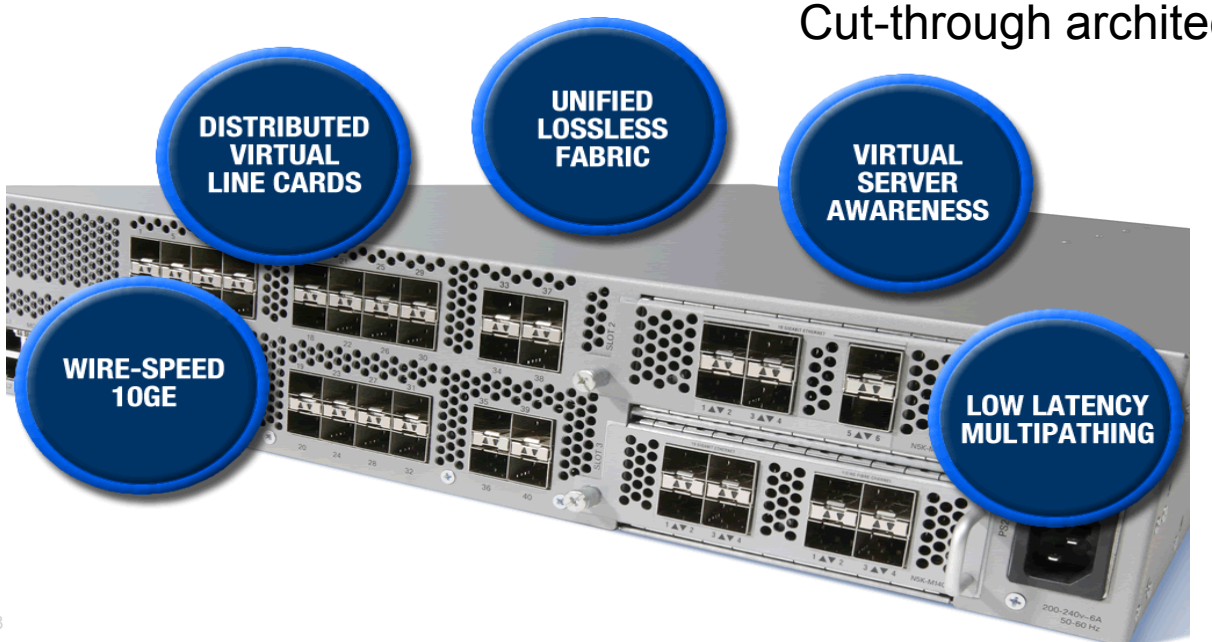
  Dual port 10 Gigabit Ethernet adapter

  Dual Port 4 Gbps Fibre Channel HBAs



Nexus Data Center switch family

- **Simplified Serviceability**
  - Front servicing
  - Reduce impact to other devices
- **High Availability**
  - Hot plug power supplies and fans
  - Fans N+1 redundant
  - NX-OS modular Operating System
- **Front to Back Cooling**
- **Simplified Cabling**
  - Ports in back
  - Cost optimized cabling – Twinax, USR
  - Selection of optical interfaces
- **Low Latency**
  - 3.2µs (port to port) with services turned on
  - Consistent for all packet sizes
  - Cut-through architecture



Open Standards based effort for DCE

Feature / Standard	Standards Status
Priority Flow Control (PFC) IEEE 802.1Qbb	PAR Approved, Editor Claudio DeSanti (Cisco) 1st Draft Published Expected Approval 04/2009
Bandwidth Management IEEE 802.1Qaz	PAR Approved, Editor Craig Carlson (QLogic) 1st Draft Published Expected Approval 07/2009
Congestion Management IEEE 802.1Qau	PAR Approved, Editor Norman Finn (Cisco) Advanced Development Draft 1.2 Published Expected Approval 04/2009
Data Center Bridge Exchange Protocol DCBX	This is part of Bandwidth Management (IEEE 802.1Qaz)
Fibre Channel over Ethernet FCoE	In T11 FC-BB-5 Letter Ballot Standard 06/2009
L2 Multipath for Unicast & Multicast	TRILL is an IETF Working Group Last Call Expected 11/2008

TCO – 1GE/4GFC vs. Consolidated I/O

Cisco Nexus 5000 Unified Fabric TCO Calculator

Assumptions

Number of servers:	1,000
Traditional network server connections to LAN:	6
Traditional network server connections to SAN:	2
Span of analysis (years):	4

The following information can be customized if you click through the levels by clicking on the + button.

Unconsolidated LAN +

Unconsolidated SAN +

Consolidated Network +

Service +

Cabling +

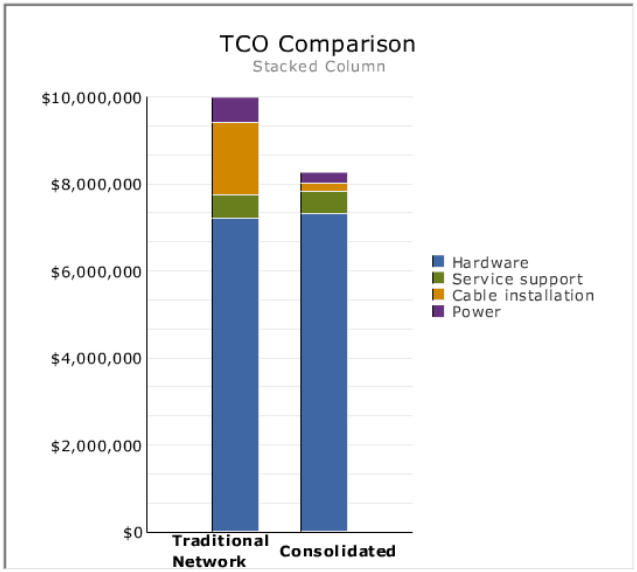
Power +

Cisco Nexus 5000 Unified Fabric TCO Calculator Scope and Assumptions

Model Scope:
The model compares a Unified Fabric implementation where LAN and SAN I/O are combined over 10GbE to traditional, unconsolidated I/O technology based on separate LAN and SAN technologies. The objective is to economically quantify the comparative value of these approaches to constructing data center networks. Due to the newness of Unified Fabric technology there are many benefit areas that cannot be quantified as yet but are expected to show significant

Detailed assumptions | For more info read this TCO White Paper

Cost	Traditional Network	Consolidated	Difference
Hardware	$7,220,500	$7,311,500	
Service support	$515,600	$526,000	
Cable installation	$1,690,000	$190,000	
Power	$559,354	$227,375	
Total Cost	$9,985,454	$8,254,875	17%
Inter-rack cables	8,450	450	95%
Power utilized (kw-hr)	2,644,994	1,184,352	55%



TCO – 10GE/4GFC vs. Consolidated I/O

Cisco Nexus 5000 Unified Fabric TCO Calculator

Assumptions

Number of servers:	1,000
Traditional network server connections to LAN:	4
Traditional network server connections to SAN:	2
Span of analysis (years):	4

The following information can be customized if you click through the levels by clicking on the + button.

Unconsolidated LAN +

Unconsolidated SAN +

Consolidated Network +

Service +

Cabling +

Power +

Cisco Nexus 5000 Unified Fabric TCO Calculator Scope and Assumptions

Model Scope:
The model compares a Unified Fabric implementation where LAN and SAN I/O are combined over 10GbE to traditional, unconsolidated I/O technology based on separate LAN and SAN technologies. The objective is to economically quantify the comparative value of these approaches to constructing data center networks. Due to the newness of Unified Fabric technology there are many benefit areas that cannot be quantified as yet but are expected to show significant

Detailed assumptions | For more info read this TCO White Paper

Cost	Traditional Network	Consolidated	Difference
Hardware	$9,901,000	$7,087,430	
Service support	$582,040	$518,080	
Cable installation	$1,276,800	$176,800	
Power	$422,603	$223,628	
Total Cost	$12,182,443	$8,005,938	34%
Inter-rack cables	6,384	384	94%
Power utilized (kw-hr)	2,116,528	1,163,538	45%