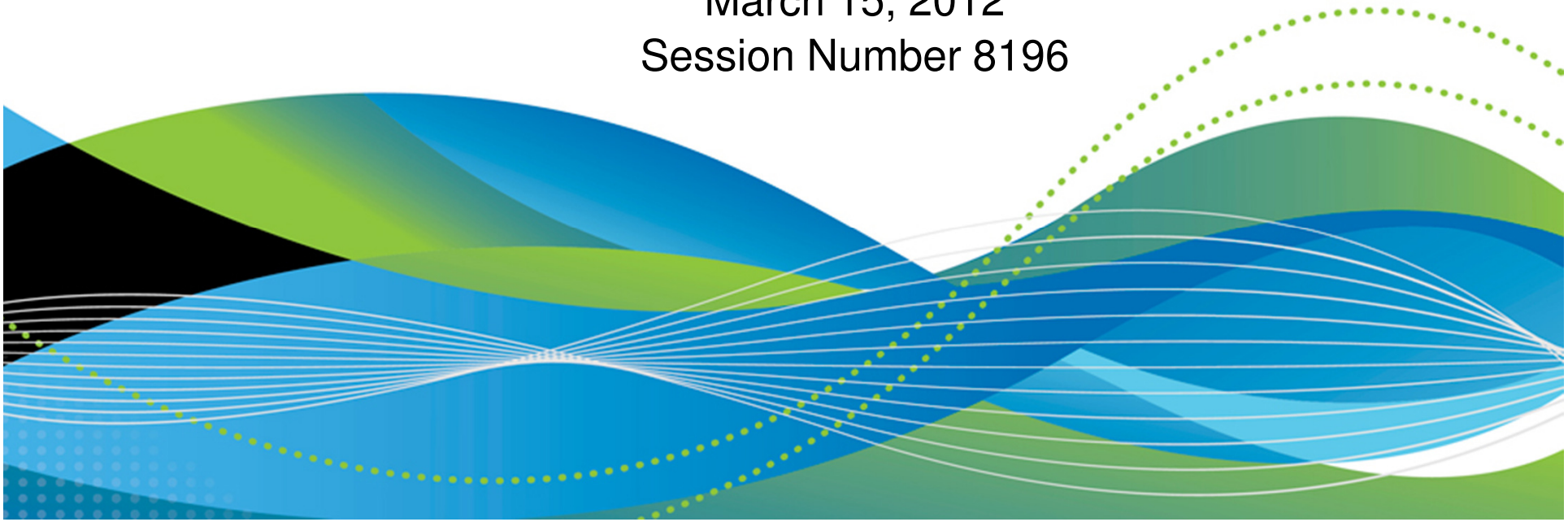
#SHAREorg

SHARE
Technology · Connections · Results

Datacenter Networking Convergence – Trends and Directions

Bob Nusbaum
Cisco Mainframe Solutions

March 15, 2012
Session Number 8196



Agenda

- What's the Problem?
- Technology and Standards for Convergence
- Solution Components
- Key Take-aways

Data Center Application Architecture

The Data Center Evolution

IT Relevance and Control

Phase 1	Phase 2	Phase 3
Mainframe	Client-Server and Distributed Computing	Service Oriented and Web 2.0 Based
Centralized	Decentralized	Virtualized

Challenges in Today's Data Centers

Customer Requirements:

- Reduce:
  - Power and Cooling Requirement
  - Cables
  - Architecture Complexity
  - Administrators
- Increase:
  - Speed
  - Utilization
  - Agility
  - Application Roll-out Speed & Efficiency

Scale of Current Problems Demands a New Approach

Infrastructure Scalability

- By 2009, 50% of large businesses will spend more on power and cooling then on new servers (Gartner, 2006)
- Quad-cores and octal-cores will drive significantly more traffic
- Storage is expected to continue to grow at a 40-70% CAGR (Gartner, 2006)

Transport Flexibility

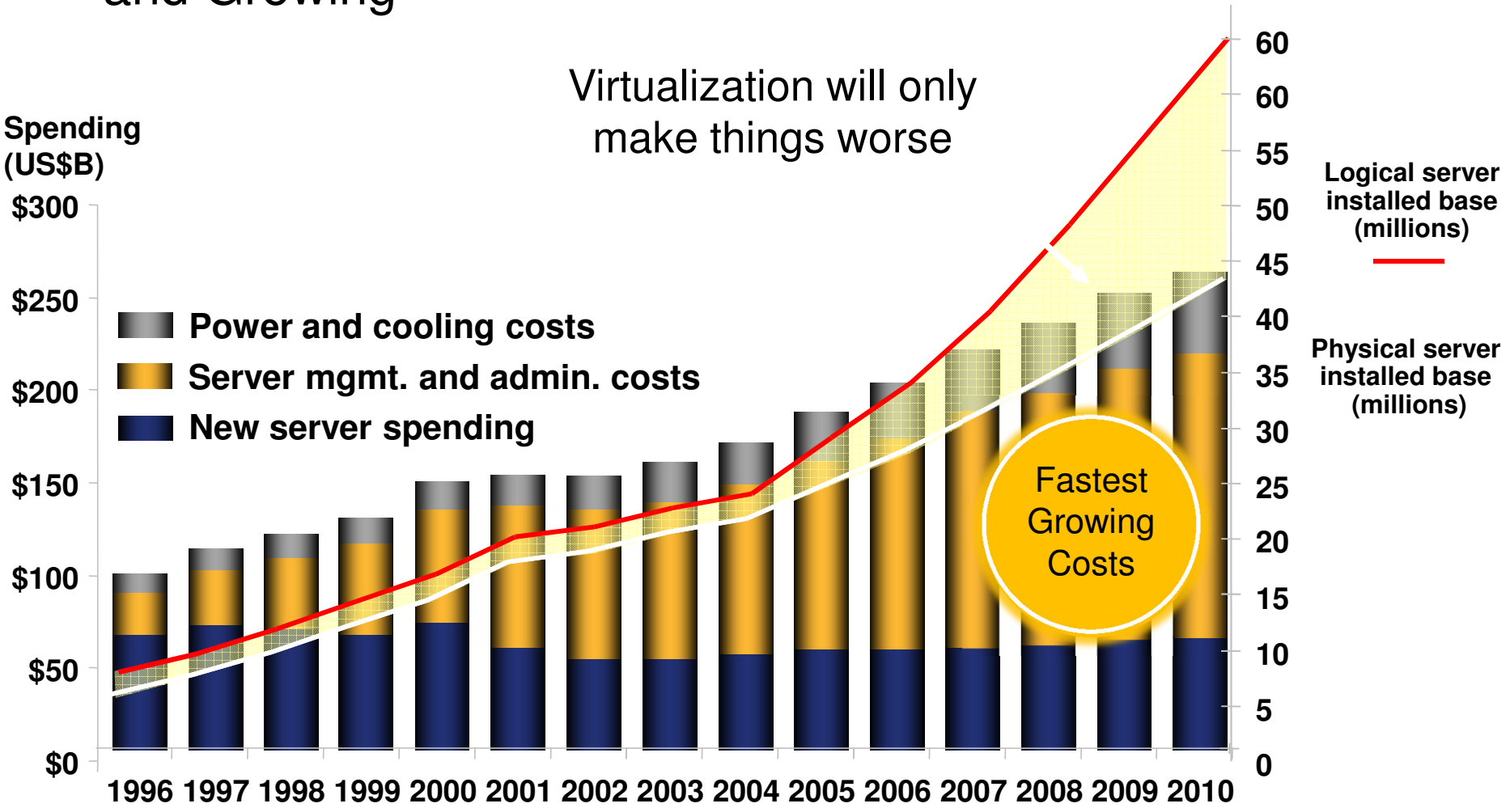
- Continued deconstruction of the server increases demands on the network
- Market transitions between transport technologies and application architectures

Operational Continuity

- Expectation of 24x7 application availability
- 54% of network downtime is caused by human error
  (Uptime Institute, 2007)

Data Center Infrastructure

Operations & Maintenance Now ~80% of IT Budgets and Growing



Source: IDC 2009

FICON: A Precursor to Convergence

- Simpler network
- Reduction in cost per port
- Asset utilization – 1.4 x performance (4:1 consolidation)
  - Cost of FICON features very small when performance enhancement is considered
- Application effectiveness
  - Up to 100% improvement
- Additional throughput
  - Roughly 6x bandwidth of ESCON
  - Shortens file maintenance windows by roughly 25%
- Competitively positions company for future
  - ESCON will have no new development
  - ESCON will not be supported on future System z models

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

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

Future Proofing - Network, Cable Plant

The Case for Converged Fabric

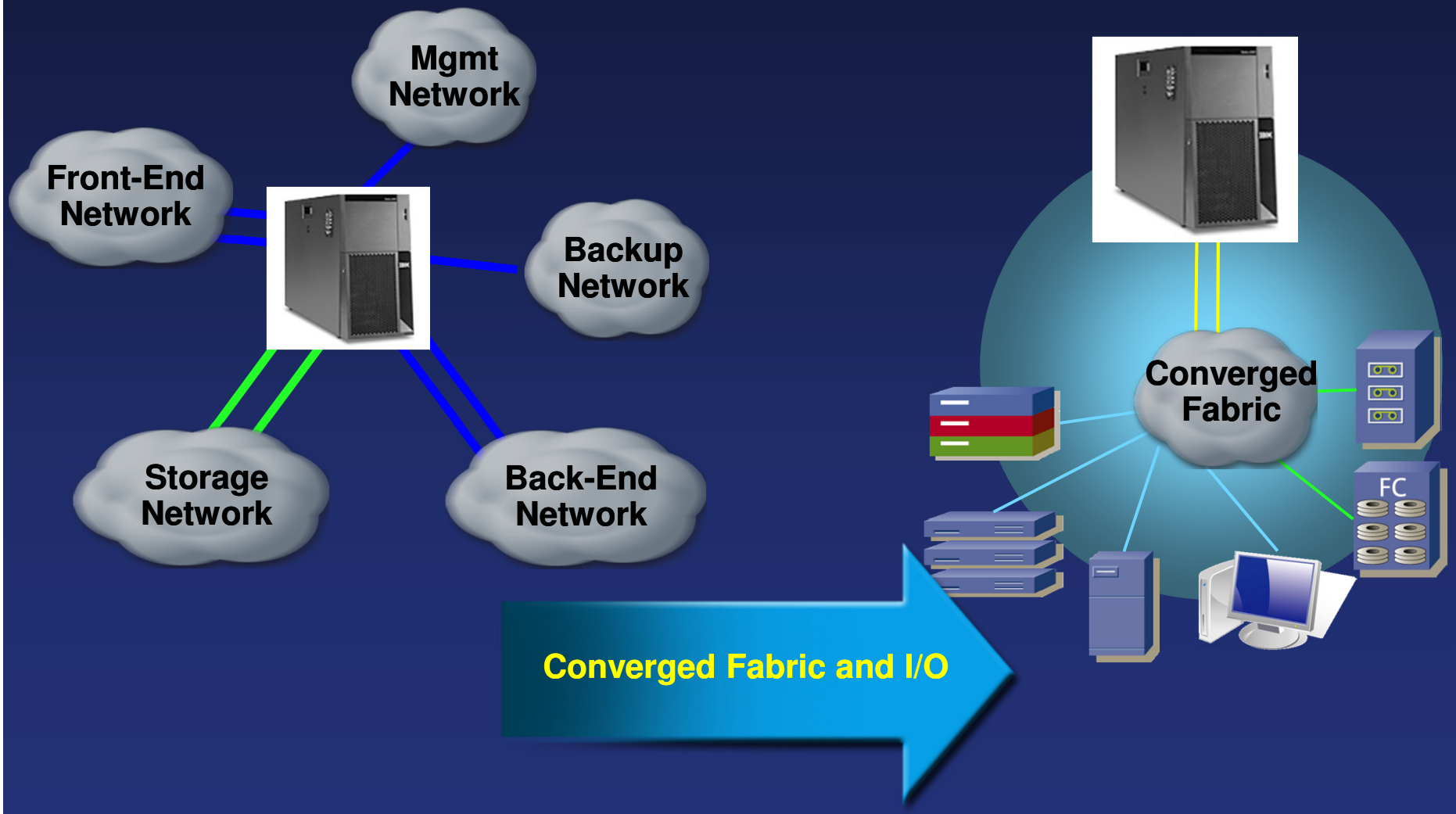
Reduce overall Data Center power consumption by up to 8%. Extend the lifecycle of current data center.

Wire hosts once to connect to any network - SAN, LAN, HPC. Faster rollout of new apps and services.

Every host is SAN-enabled. Drive storage consolidation and improve utilization.

Rack, Row, and X-Data Center VM portability become possible.

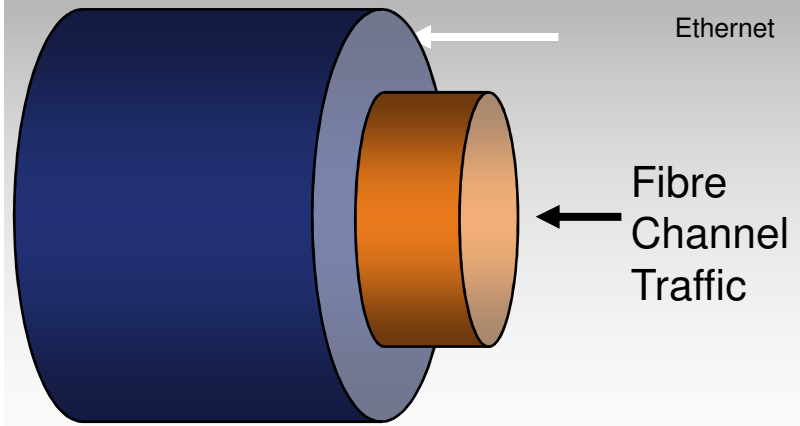
Increased Efficiency, Simpler Operations



Converged Fabric Overview
Fibre Channel over Ethernet (FCoE)

FCoE

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

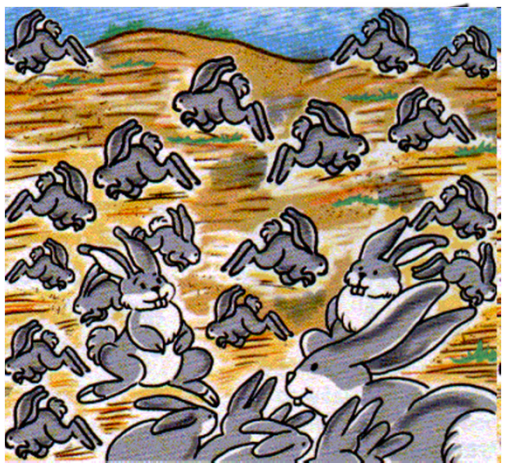


Benefits

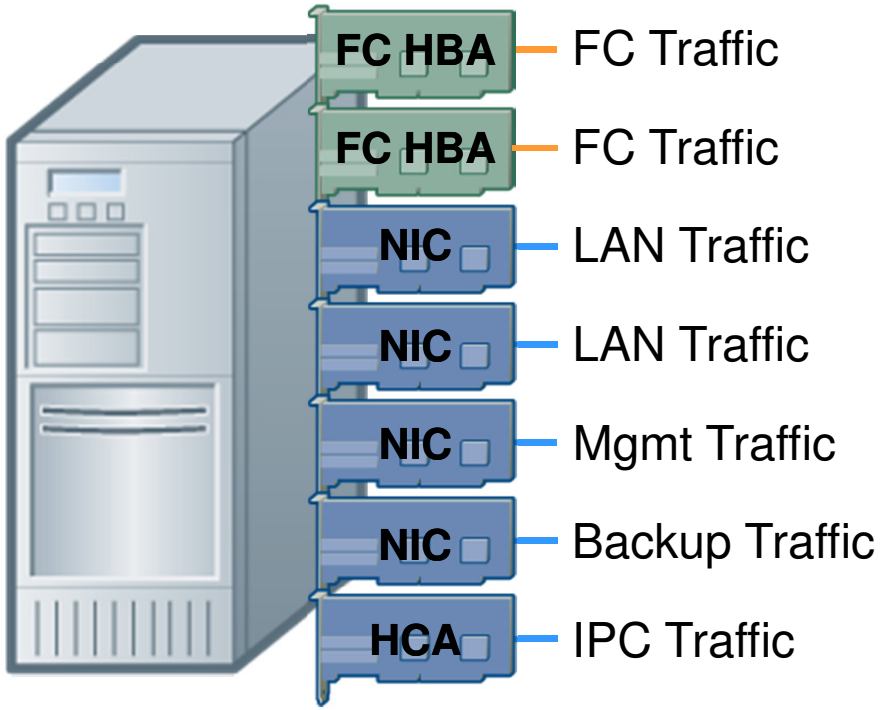
- Fewer Cables
  - Both block I/O & Ethernet traffic co-exist on same cable
- Fewer adapters needed
- Overall less power
- Interoperates with existing SAN's
  - Management SAN's remains constant
- No Gateway

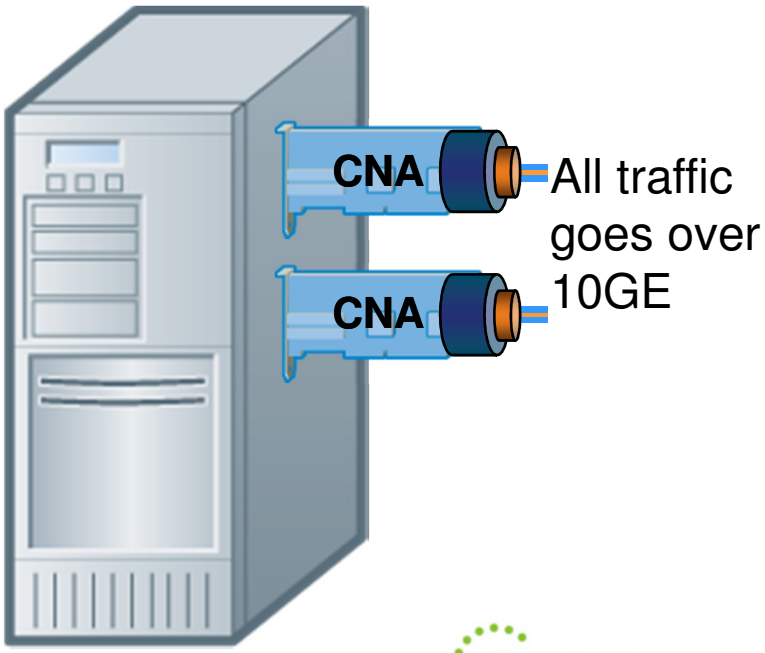
Converged Fabric

Why?



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





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

  Yet another ULP (Upper Layer Protocol) to be transported, but… a challenging one!

- And technically…

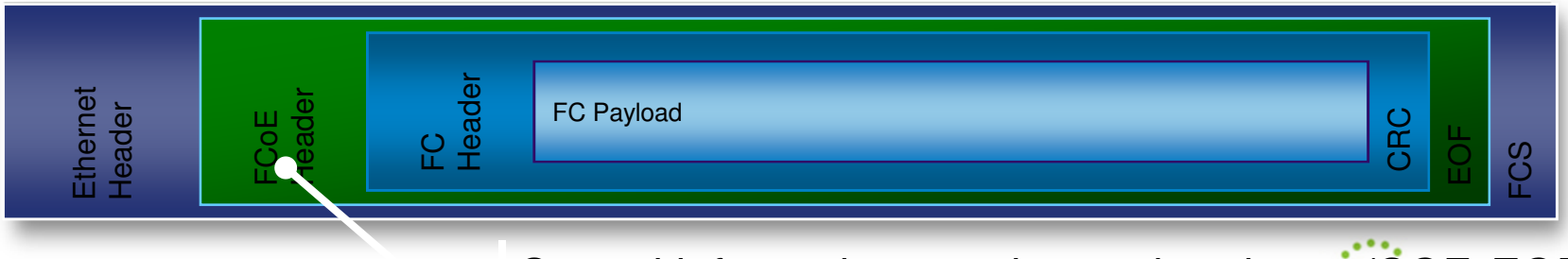
FCoE allows the extension of Fibre Channel onto a Lossless Ethernet fabric

FCoE Enablers

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

Normal Ethernet frame, ethertype = FCoE

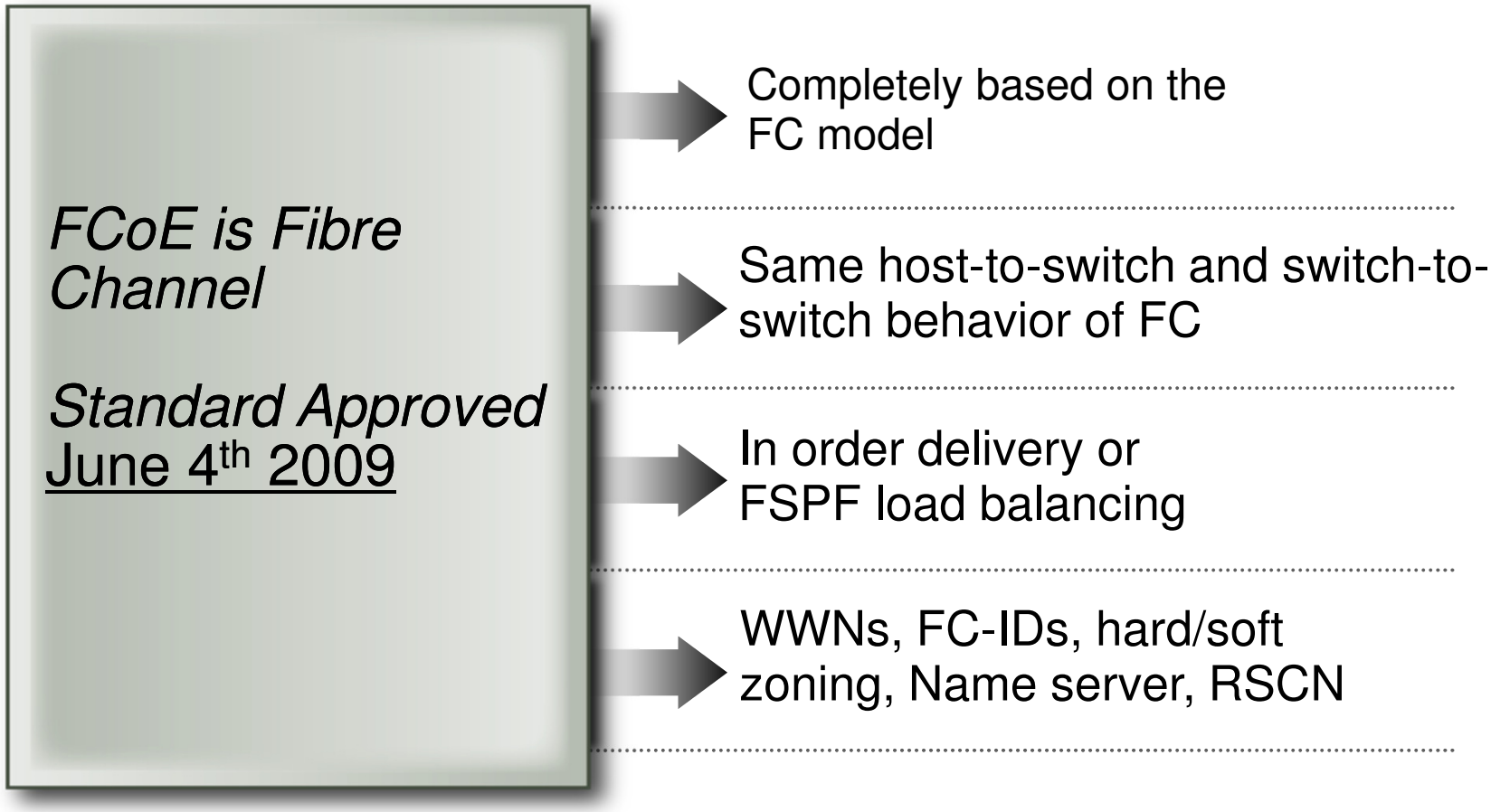
Same as a physical FC frame



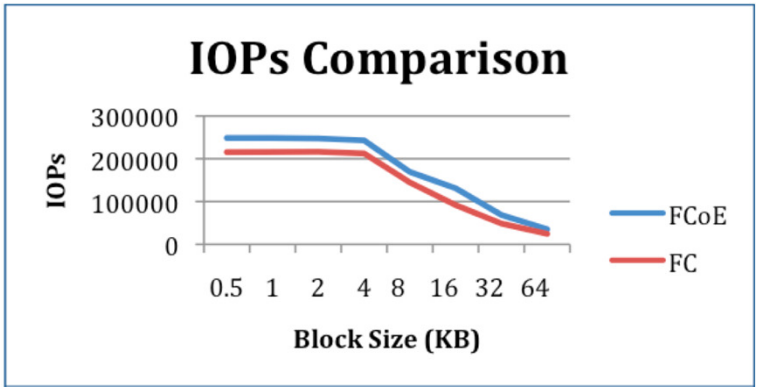
Control information: version, ordered sets (SOF, EOF)

FCoE Advantages

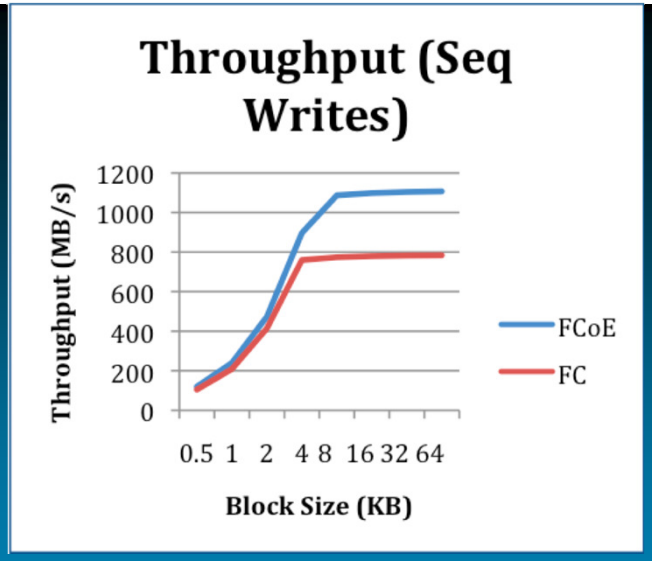
FCoE is managed like FC at initiator (CH), target (CU), and switch level

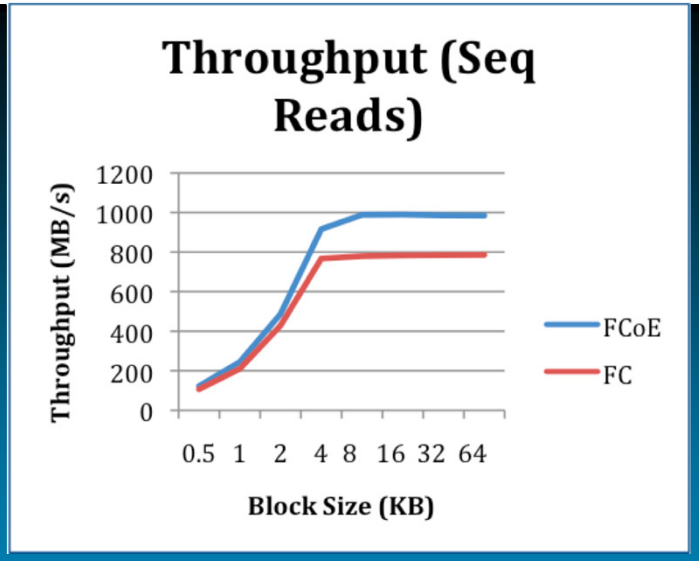


FCoE versus 8Gig Fibre Channel Comparisons -Test results



FCoE technology is capable of providing up to 250,000 IOPs





As 8G FC reaches its throughput limits; FCoE can still provide additional bandwidth, up to 25% for larger block sizes.

The FC-BB-5, FCoE, draft standard was unanimously approved as the final standard- June 4th 2009

1) FC-BB_IP: The FC-BB_IP model defines the means by which Fibre Channel networks interface with and connect across an IP network.

1) FC-BB_GFPT: The FC-BB_GFPT model defines the means by which FC physical links may be extended over any WAN. Transport infrastructure for which GFP mapping is defined.

2) FC-BB_PW: The FC-BB_PW model defines the means by which FC physical links may be extended over a wide area MPLS network

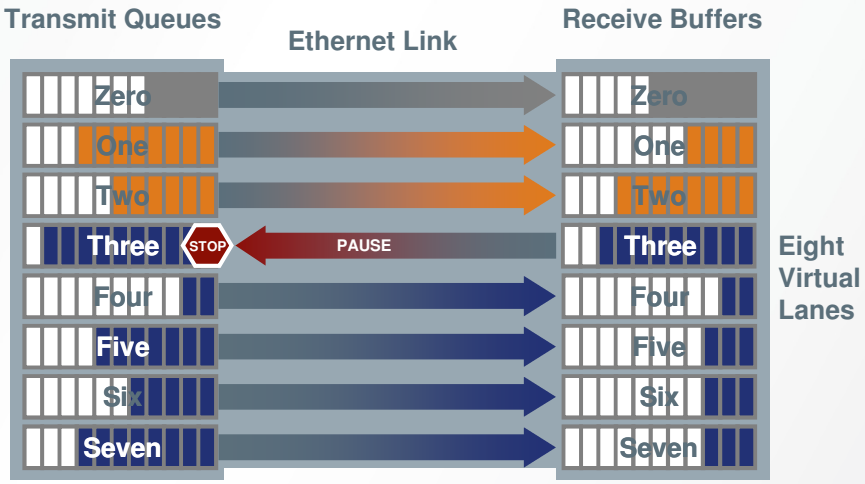
3) FC-BB_E: The FC-BB_E model defines the means by which Fibre Channel frames are transported over a Lossless Ethernet network

Nexus 5000 Enhanced Ethernet Fabric Features

Priority Flow Control IEEE 802.1Qbb	Provides class of service flow control by enabling PAUSE functionality on IEEE 802.1p lanes
Data Center Bridging Exchange **IEEE 802.1AB**	Auto-negotiation of Enhanced Ethernet capabilities DCBX (switch to NIC)
Bandwidth Management IEEE 802.1Qaz	Enhanced Transmission Selection - manage bandwidth and assign priorities to groups of IEEE 802.1p lanes based on class of traffic
Congestion Management IEEE 802.1Qau	This standard specifies protocols, procedures and managed objects that support congestion management of long-lived data flows within network domains of limited bandwidth delay product (BCN/QCN)
L2 Multipathing **"IETF TRILL"** **Transparent Interconnect of Lots of Links**	Layer-2 multipathing eliminates standby uplinks increasing the available uplink bandwidth. In the NX5K this is enabled by Ethernet Host Virtualizer Mode.

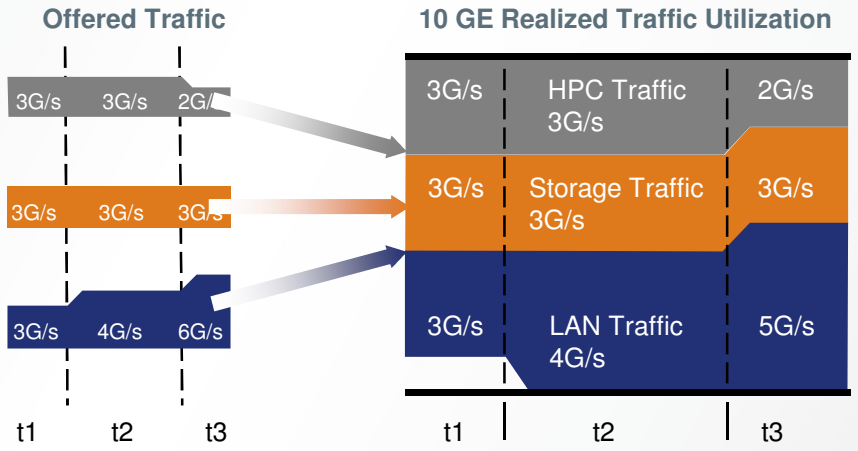
Data Center Ethernet: PFC & Bandwidth Management

Priority Flow Control



- *Enables lossless behavior for each class of service*
- *PAUSE sent per virtual lane when buffers limit exceeded*

CoS based Bandwidth Management

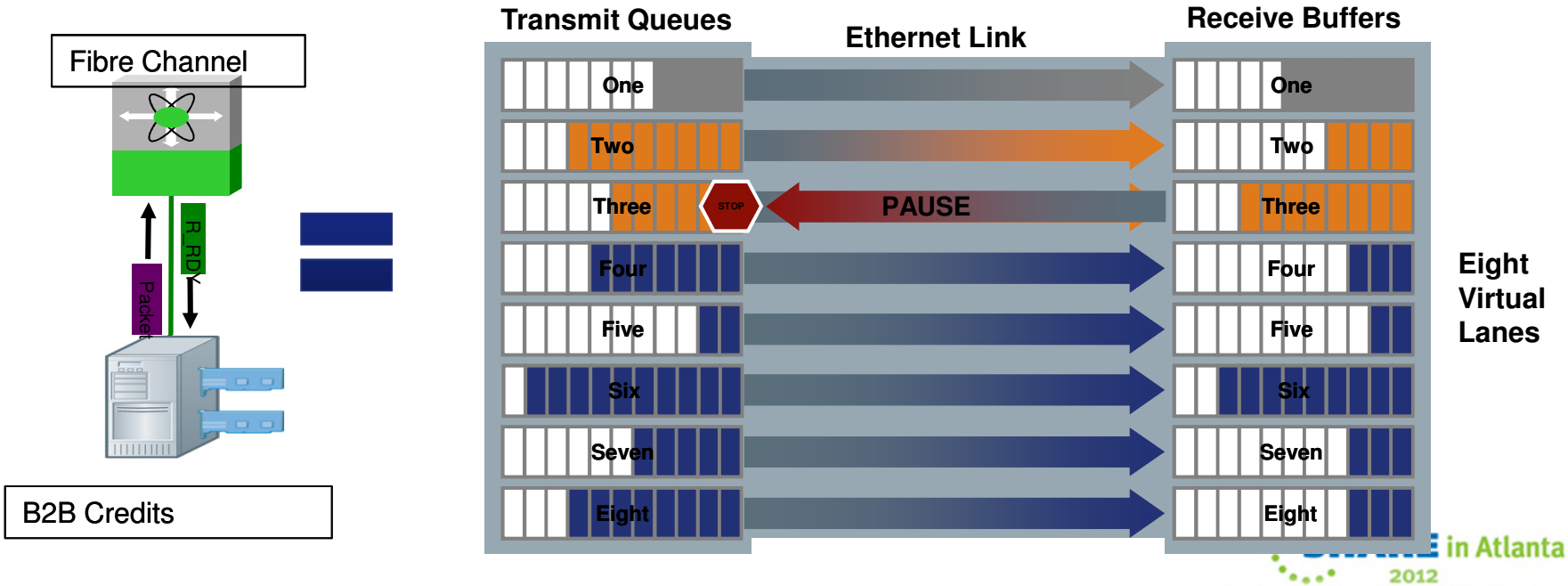


- *Enables Intelligent sharing of bandwidth between traffic classes control of bandwidth*
- *802.1Qaz Enhanced Transmission*

Priority Flow Control

Fibre Channel over Ethernet Flow Control

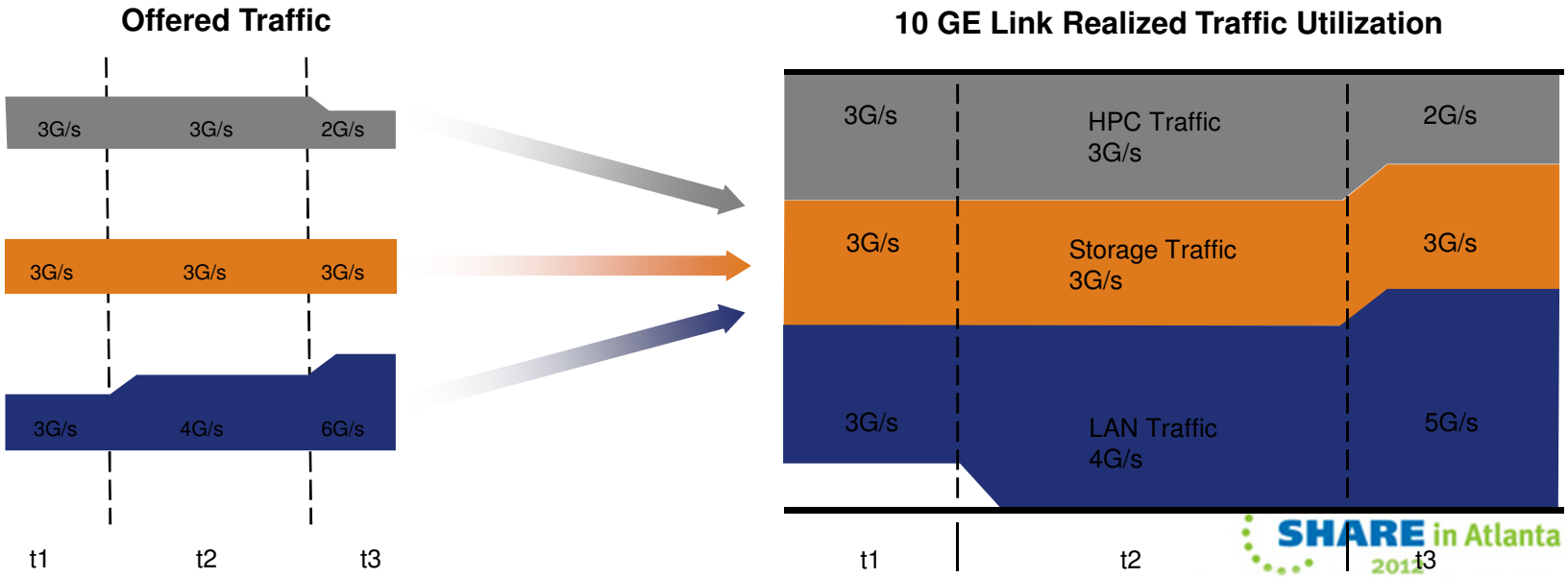
- Enables lossless Ethernet using PAUSE based on a COS as defined in 802.1p
- When link is congested, CoS assigned to FCoE will be PAUSEd so traffic will not be dropped
- Other traffic assigned to other CoS will continue to transmit and rely on upper layer protocols for retransmission



Enhanced Transmission Standard

Bandwidth Management -- IEEE 802.1Qaz

- Required when consolidating I/O – It's a QoS problem
- Prevents a single traffic class of "hogging" all the bandwidth and starving other classes
- When a given load doesn't fully utilize its allocated bandwidth, it is available to other classes
- Helps accommodate for classes of a "burtsy" nature



Will FCoE Work for FICON?

Key capabilities are there:

- **Class 2**
- **Loop Back**
- **Credit Pooling**
- **Performance and Error Statistics**
  - Number of Words Transmitted
  - Number of Words Received
  - Number of Frames Transmitted
  - Number of Frames Received
  - Number of Frames Discarded
  - Time that port is unable to receive frames from xN_Port due to zero credit
  - Time that port is unable to send frames to xN_Port due to zero credit
  - General error statistics appropriate to the port protocol:
    - Number of Code Violation Errors
    - Number of CRC/checksum errors
    - Number of times Loss of Sync was detected
    - Number of Link Failures

Standards work is in process in T11 for FICON over DCB

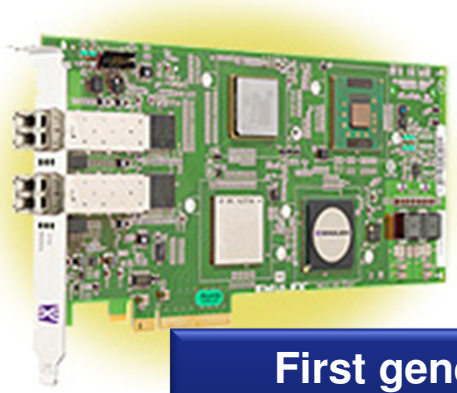
What are the necessary pieces?

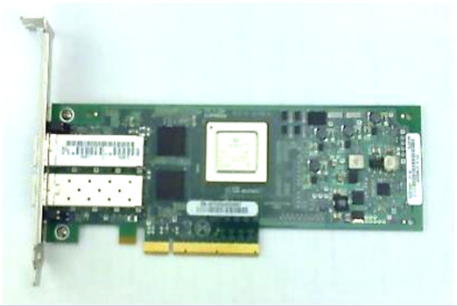
IBM System X Converged I/O Adapters and Blade Center H and HT

Hardware Converged Network Adapter Solution

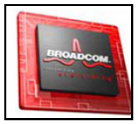
- 10 Gbps maximum throughput for high-bandwidth storage (SAN) and networking (LAN) traffic
  - Dual 10GE/FCoE ports
- Support for native drivers and utilities
  - Customer certified stacks
- Replaces multiple adapters per server/blade
- Consolidates 10GE and FC on a single interface
- Blade Center Support for Blade Open Fabric Manager for BIOS, UEFI, and Fcode

CNA Technology Rapidly Evolving





LOM



First generation
Mid-2008

QLOGIC
EMULEX

Second generation
Mid-2009

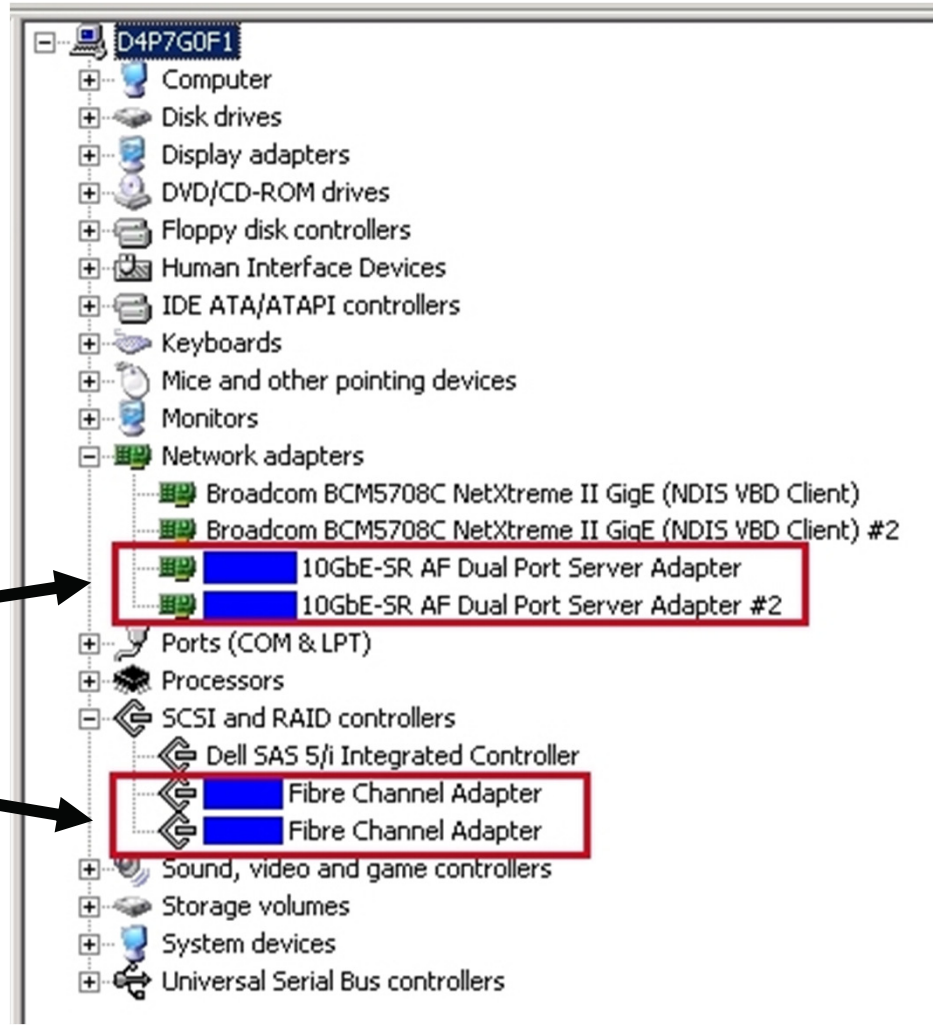
QLOGIC
EMULEX

Near future

intel
BROADCOM
And other NIC vendors

View from Operating System

- CNAs are multi-function devices
- Standard drivers
- Same management
- Operating System sees:
  - Dual port 10 Gigabit Ethernet adapter
  - Dual Port Fibre Channel HBAs



Cost Effective 10G Server Connectivity
Today

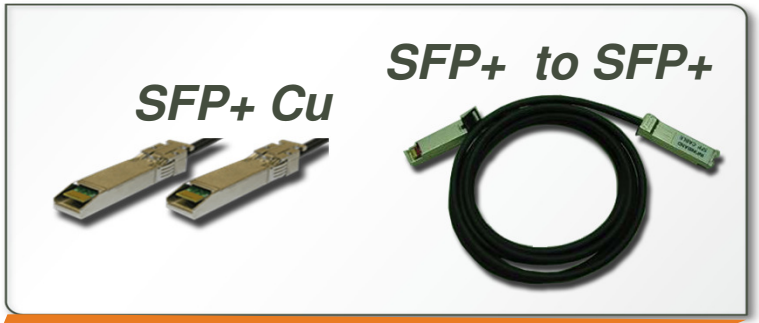
SFP+ USR – 'Ultra Short Reach'

- 100M on OM3 fiber, 30M on OM2 fiber
- Support on all Cisco Catalyst and Nexus switches

SFP+ Direct Attach

- 1, 3, 5 and 7M and 10M on Twinax active cables
- Support across all Nexus Switches

Evolution of Ethernet Physical Media
Role of Transport in Enabling 10GE Technology



Technology		Distance	Power (each side)	Transceiver Latency
SFP+ CU Copper	Twinax	5m	~0.1W	~0.25µs
SFP+ USR ultra short reach SR compatible	MM OM2 MM OM3	30m 100m	1W	~0.1µs
SFP+ SR short reach	MM OM1 MM OM3	33m 300m	1W	~0.1µs
10GBASE-T	Cat6a/7 Cat6a/7	100m 30m	~8W ~4W	2.5µs 1.5µs

Unified Fabric

What is it?
Really?

DCNM

Ethernet
DCB
Fibre Channel
FCoE only

SHARE
Technology · Connections · Results

Fabric A/B Separation

Core

NX-OS

FCoE

FCoE NAS iSCSI

FCoE

FC

FC

L3

L2

Converged Aggregation

FCoE

FCoE

Core

Fabric 'A'

FC

Fabric 'B'

FCoE

Converged Access

Edge

CNA

HBA

FCoE

FC

SHARE in Atlanta 2012

Critical Infrastructure for Data Center Networking

Converged Fabric and I/O Interfaces

Enables storage, mainframe, Ethernet, IP, and HPC traffic to converge to a single network

Lowers overall data center power draw

Cisco® Nexus Switching Platforms

Multi-Terabit platform with ops-centric design

Delivers unified fabric and I/O deployments

Designed for the most stringent availability needs

NX-OS Operating System

First multi-protocol DC-class operating system

Virtualized control plane and scalable design sets new standard for usability

Cisco MDS Storage Switching

The industry-leading features and capabilities of the Cisco MDS 9000 Family provide superior performance and efficiency for deploying Unified I/O in Data Center Networking

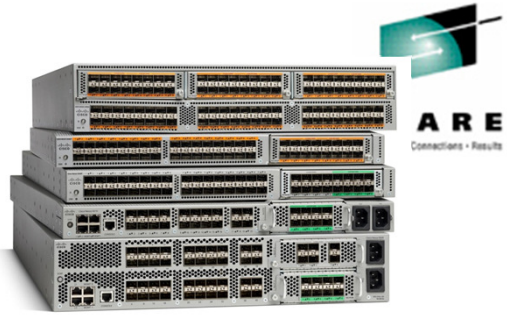
Selecting the Unified Fabric Mix

Add MDS 9000 for

- New or existing servers with HBAs
- New or existing FC or FCoE-only storage ports
- Replacing DCX, SilkWorm or McDATA switches
- SAN Extension over MAN or WAN links
- Advanced storage services (migration, encryption, etc.)

Add Nexus 7000 for

- SAN core for multi-protocol Ethernet storage devices:
  - FCoE and
  - iSCSI / NAS
- Highly available access layer for mission critical servers with CNAs
- Aggregation layer for fully converged fabrics



Add Nexus 5000 for

- Cost-effective server access using CNAs
- Small Unified Fabrics containing FC and FCoE (SAN & LAN in one box)
- Server access where the mix of HBAs and CNAs will change over time

Director Class Checklist

Attribute	MDS 9500	Nexus 7000
Physical Redundancy	Supervisors, Power Supplies, Fans	
Logical Redundancy	Port Channels, Load Balancing	
Non Disruptive Upgrades	Full ISSU	
Segmentation/Isolation	VSANs, VDCs	
Maximum Scale	528 ports	768 ports
Performance	256-Gbps/Slot	550-Gbps/Slot
FCoE modules	8-port	32-port ➔ 48

8-Port 10G FCoE Module for MDS 9500

8-Port 10G FCoE Module

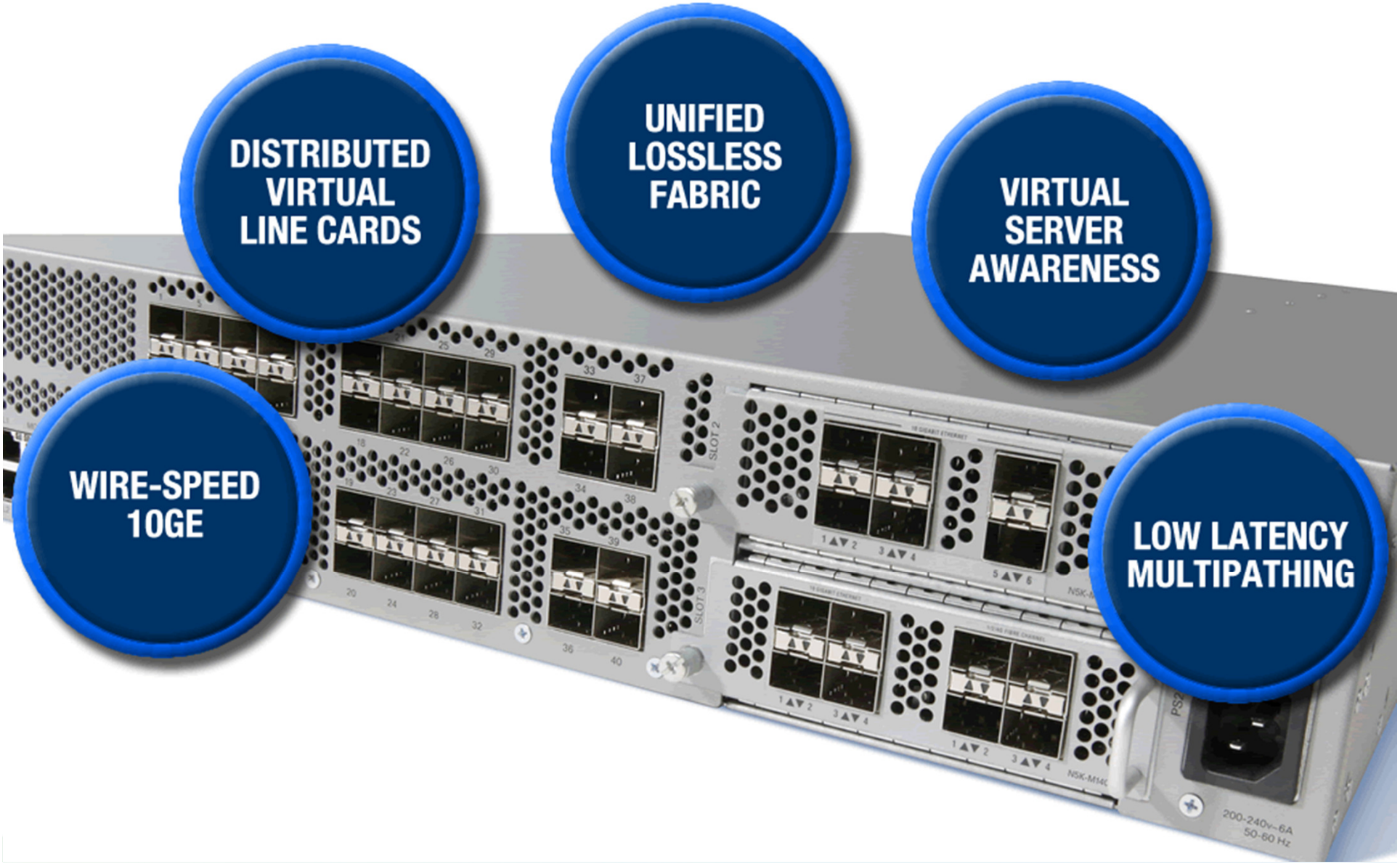
Enables integration of existing FC infrastructure into Unified Fabric

- 8 FCoE ports at 10GE full rate
- 80-Gbps front panel bandwidth
- SFP+ SR, LR, CX-1 optics support

FCoE connectivity from MDS 9500 Directors to:

- Nexus 5000 and Nexus 7000
- FCoE Storage Arrays

Cisco Nexus 5000
Now available through IBM!

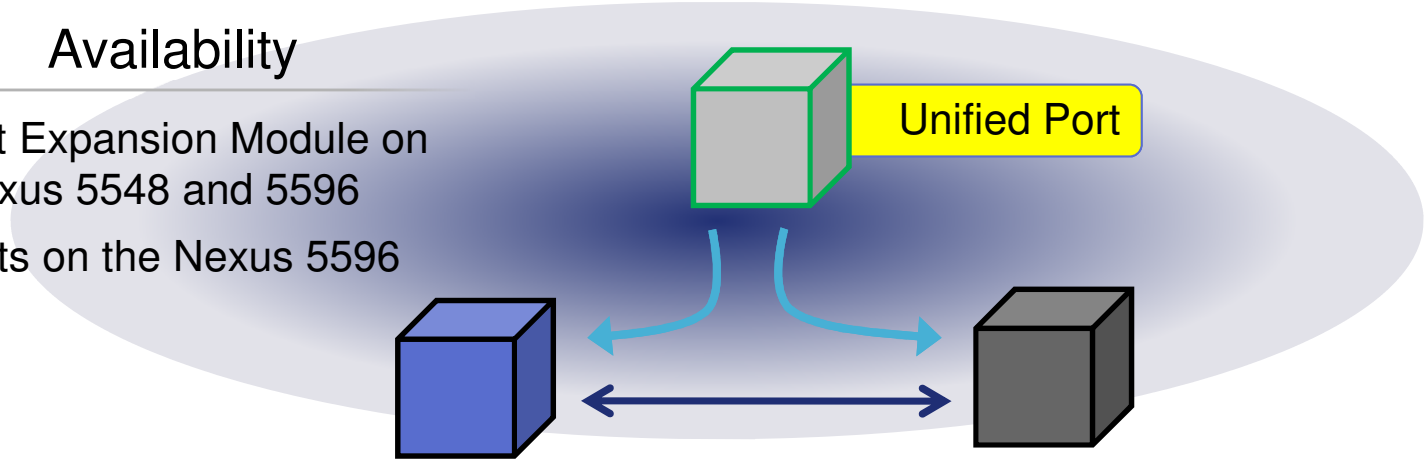


Unified Ports

Dynamic and Efficient Port Allocation

Availability

- 16-port Expansion Module on the Nexus 5548 and 5596
- All Ports on the Nexus 5596



Benefits

- Simplify switch purchase - remove ports ratio guess work
- Increase design flexibility
- Remove specific protocol bandwidth bottlenecks

Use-cases

- Flexible LAN & storage convergence based on business needs
- Service can be adjusted based on the demand for specific traffic

Summary & Take-Aways

- Fully standardized
- Major cost savings
- Products are there
- Deploy now for distributed systems
- FICON – “watch this space”

The bottom line….

“It’s One Network!”

Thank you.