



## The Next Generation Datacenter



Fabio Di Niro  
Systems Engineer  
[fabio.diniro@cisco.com](mailto:fabio.diniro@cisco.com)

## Agenda

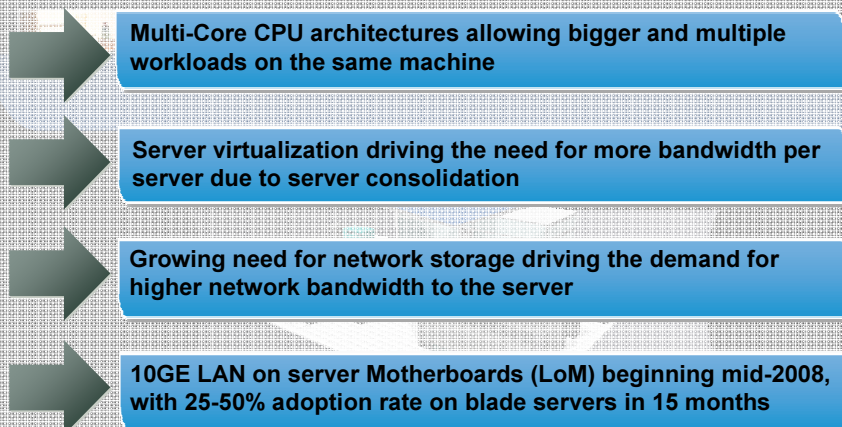


- Datacenter 3.0
- A new consolidation opportunity
- Cisco I/O consolidation solutions
- Unified Fabric Evolution
- Summary
- Q&A

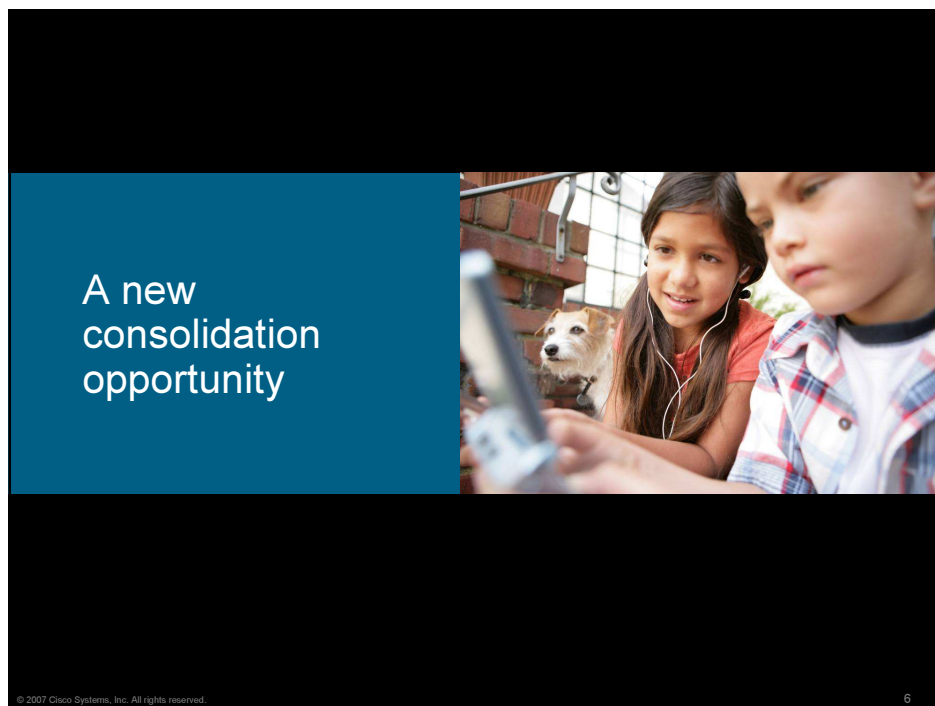
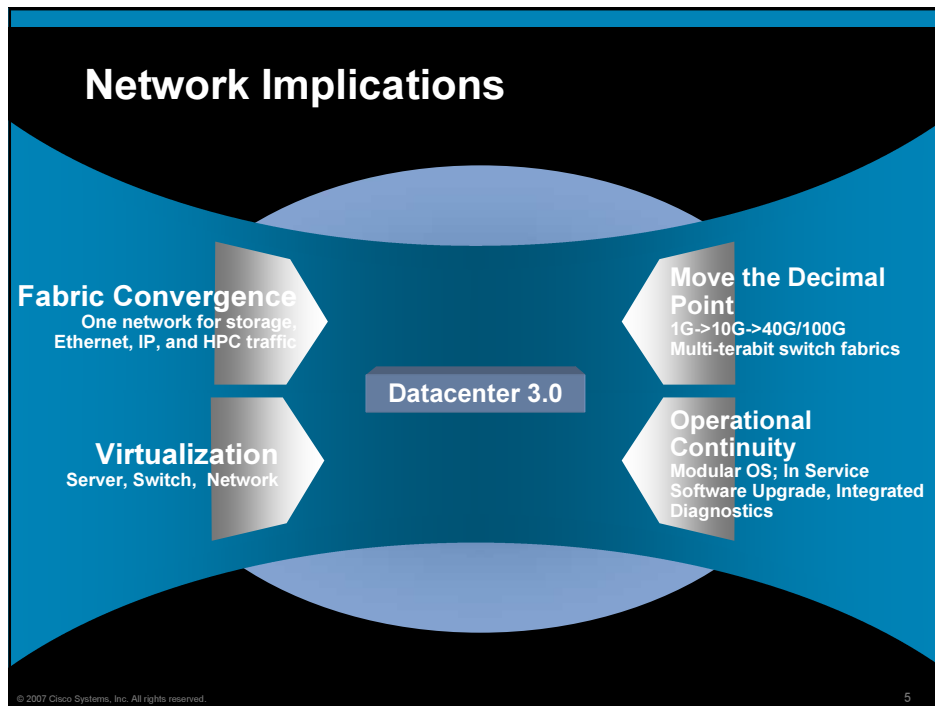
## Datacenter 3.0



## Datacenter pressures



***10GE LOMs, Multi-Core CPUs and Server Virtualization driving the adoption of 10GE network connections***



## Cabling and I/O Consolidation

*I/O Consolidation Goal is:*

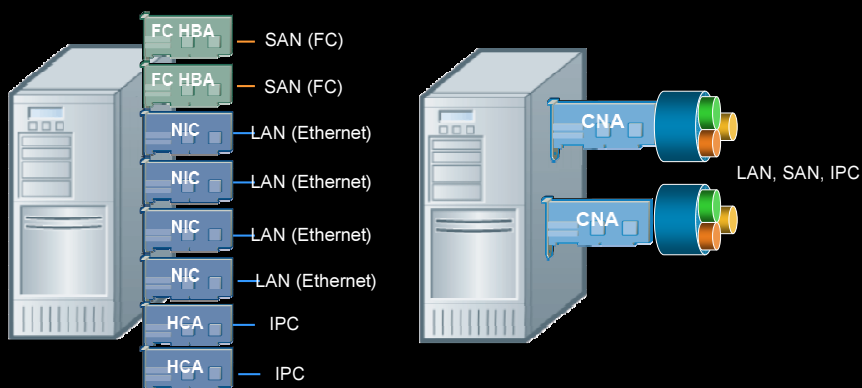
**Reduce  
Cabling**

*Less cables, less  
costs*

**Standardize  
Cabling**

*Cables are protocol  
independent*

## The Vision – Unified Data Center I/O: *Fewer HBA/HCA/NIC's per Server*



*CNA = Converged Network Adapter*

## Merging networks requirements

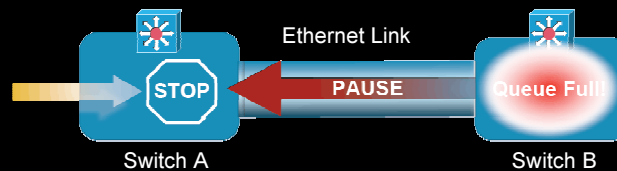
LAN/IP	Storage	IPC
	<ul style="list-style-type: none"> <li>Must follow the Fibre Channel model</li> <li>Loosing frame is not an option</li> </ul>	<ul style="list-style-type: none"> <li>Doesn't care of the underlying network, provided that                             <ul style="list-style-type: none"> <li>It is cheap</li> <li>It is low latency</li> <li>It supports APIs like OFED, RDS, MPI, sockets</li> </ul> </li> </ul>

## Why Ethernet loose frames?

Collision	Transmission Error	Congestion
<ul style="list-style-type: none"> <li>No longer present in Full Duplex Ethernet</li> </ul>	<ul style="list-style-type: none"> <li>Very rare in the Data Center</li> </ul>	<ul style="list-style-type: none"> <li>Most common cause</li> <li>Congestion is a switch issue, not a link issue                             <ul style="list-style-type: none"> <li>A full duplex IEEE 802.3 link does not lose frames</li> </ul> </li> <li>It must be dealt with in the bridge/switch                             <ul style="list-style-type: none"> <li>By IEEE 802.1</li> </ul> </li> </ul>

## Can Ethernet be lossless?

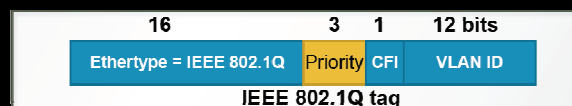
- Yes, with Ethernet PAUSE Frame



- Defined in IEEE 802.3 – Annex 31B
- The PAUSE operation is used to inhibit transmission of data frames for a specified period of time
- Ethernet PAUSE transforms Ethernet into a lossless fabric

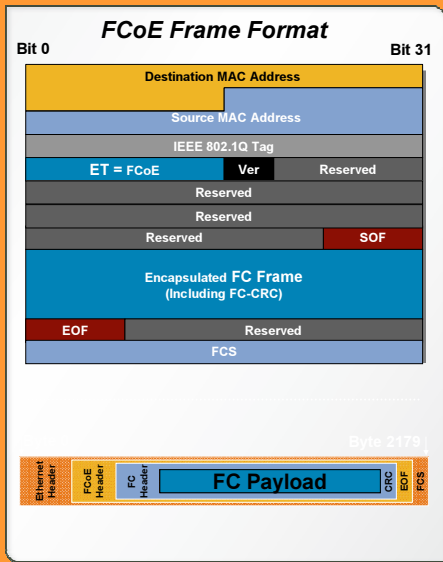
## Priority Flow Control (PFC)

- aka PPP (Per Priority Pause)
- PFC enables PAUSE functionality per Ethernet priority
  - IEEE 802.1q defines 8 priorities
  - Traffic classes are mapped to different priorities
  - No traffic interference
  - IP traffic may be paused while Storage traffic is being forwarded or vice-versa
  - Requires independent resources per priority (buffers)
- High level of industry support
- Cisco distributed proposal
- Standard Track in IEEE 802.1Qb



# FCoE Frame Format Specification

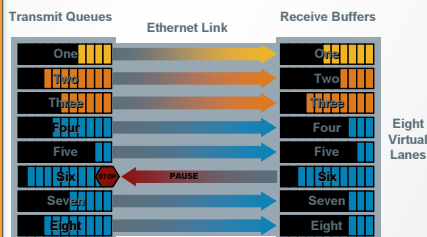
- An extension of FC over lossless Ethernet
- FCoE Specification in ANSI INCITS FC T11.3
  - Frame Format agreement Aug. 2007
  - Target completion 2H08
- Optional features being defined in IEEE
  - PAUSE enhancements 802.3x mechanisms
  - Priority Flow Control (PFC)
  - Jumbo Ethernet frames: FC encaps of 2112



© 2007 Cisco Systems, Inc. All rights reserved.

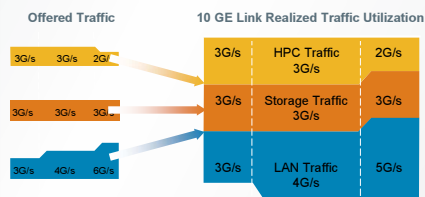
## Data Center Ethernet: PFC & Bandwidth Management

### Priority Flow Control



- Enables lossless Fabrics for each class of service
- PAUSE sent per virtual lane when buffers limit exceeded
- Standard Track in 802.1Qb

### COS based Bandwidth Management



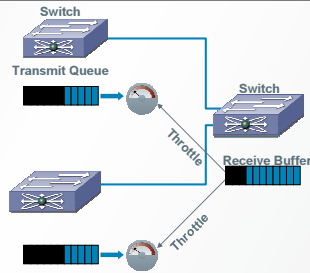
- Enables Intelligent sharing of bandwidth between traffic classes control of bandwidth
- Standard Track in 802.1Qaz

© 2007 Cisco Systems, Inc. All rights reserved.

119

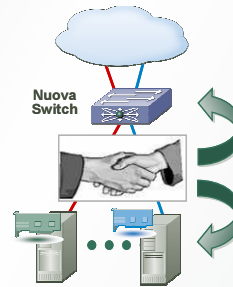
## Data Center Ethernet : Congestion Management & DCBX

### Congestion Management



- Moves congestion out of the core to avoid congestion spreading
- End-to-End congestion management
- Standards track in 802.1Qau
- N5000 Switches have capability in ASICs. Future OS release

### Data Center Bridging eXchange (DCBX) Protocol



**Handshaking Negotiation for:**

- CoS BW Management
- Priority Flow Control (PFC)
- Congestion Management (BCN/QCN)
- Application (user\_priority usage)
- Logical Link Down
- Based on LLDP (Link Level Discovery Protocol)

Cisco I/O  
consolidation solutions





## Cisco Nexus 7000 Series Data Center Class Switches

NEW



Zero Service Disruption design  
Graceful systems operations  
Integrated lights-out management

Lossless fabric architecture  
Dense 40GbE/100GbE ready  
Unified fabric

Virtualized control and data plane  
15Tb+ switching capacity  
Efficient physical and power design

Operational  
Continuity

Transport  
Flexibility

Infrastructure  
Scalability

## Extending the Cisco Nexus Family Data Center Class Switches



Simpler More Stable Layer 2 Network  
Highly Available Platform  
Preserves operational best practices

FCoE based Unified Fabric  
Virtualization Optimized Networking  
Support for CE, FCoE, DCE, and FC

Reduces power, cooling, cabling  
Up to 56 ports non-blocking 10GbE  
Up to 1.2 Tbps capacity

Operational  
Continuity

Transport  
Flexibility

Infrastructure  
Scalability

## Cisco Nexus 5000 Series

Industry First Data Center Class Switch To Deliver Unified Fabric Today



### 56-Port L2 Switch

- 40 ports 10 Gigabit Ethernet/FCoE/DCE, fixed
- Two expansion module slots



### Fiber Channel

- 8 ports 1/2/4G FC



### FC + Ethernet

- 4 Ports 10 Gigabit Ethernet/FCoE/DCE
- 4 ports 1/2/4G FC



### Ethernet

- 6 ports 10 Gigabit Ethernet/FCoE/DCE

NX-OS

DC-NM and Fabric Manager

## Evolution of Ethernet Physical Media

Role of Transport in Enabling 10GE Technology

Mid 1980's

Mid 1990's

Early 2000's

Late 2000's

10Mb

100Mb

1Gb

10Gb

UTP Cat 3

UTP Cat 5

UTP Cat 5  
SFP Fiber

X2  
SFP+ Cu (BER better than  $10^{-18}$ )  
SFP+ Fiber  
Cat 6/7

Technology	Cable	Distance	Power (each side)	Transceiver Latency
SFP+ CU Copper	Twinax	10m	~0.1W	~0.25µs
SFP+ USR ultra short reach	MM OM2 MM OM3	10m 100m	1W	~0.1µs
SFP+ SR short reach	MM OM1 MM OM3	33m 300m	1W	~0.1µs
10GBASE-T	Cat6 Cat6a/7 Cat6a/7	55m 100m 30m	~8W ~8W ~4W	2.5µs 2.5µs 1.5µs

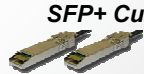
# Evolution of Ethernet Physical Media

## Role of Transport in Enabling 10GE Technology

SFP+ to SFP+



SFP+ Cu



### 10GE Copper Solution

- Low cost
- Low power and latency
- Up to 10 meters (in-rack and adjacent rack cabling)

Technology	Media	Distance	Power (each side)	Transceiver Latency
SFP+ CU Copper		10m	~0.1W	~0.25µs
SFP+ USR ultra short reach	MM OM2	10m	1W	~0.1µs
	MM OM3	100m		
SFP+ SR short reach	MM OM1	33m	1W	~0.1µs
	MM OM3	300m		
10GBASE-T	Cat6	55m	~8W	2.5µs
	Cat6a/7	100m	~8W	2.5µs
	Cat6a/7	30m	~4W	1.5µs

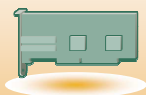
© 2007 Cisco Systems, Inc. All rights reserved.

21

## Server Adapter

### Cost

- Access for Any Ethernet Equipped Host



- Product: Software FCoE

### Compatibility

- Minimal Configuration Using Existing Driver Stacks

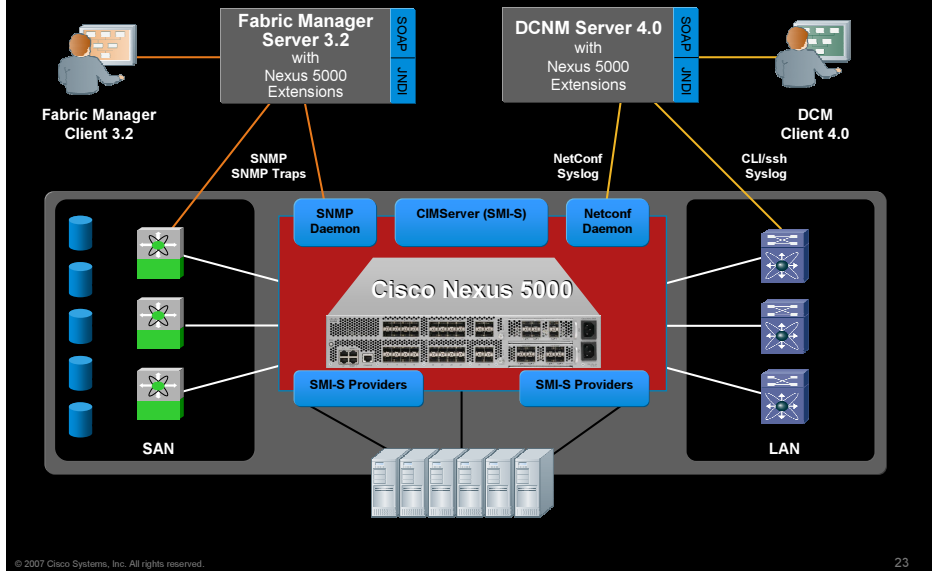


- Product: CNA (Converged Network Adapter)

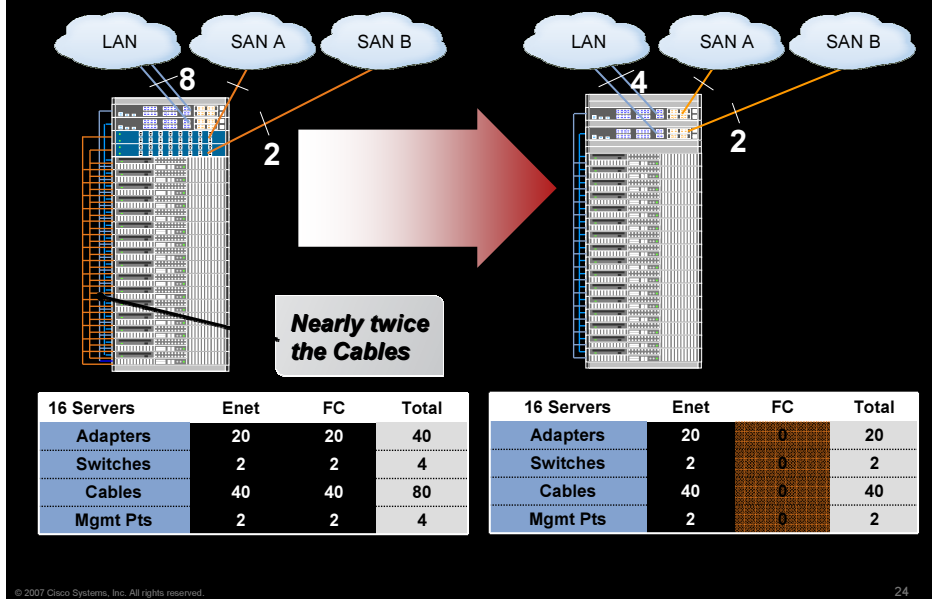
© 2007 Cisco Systems, Inc. All rights reserved.

22

## Fibre Channel and Ethernet Management



## Use Case



## Extending the Reach of Unified Fabrics

### Cross-Platform Lifecycle Investment Protection



Catalyst® 6500s are designed to interconnect with the Cisco Nexus Unified Fabric, providing increased bandwidth and traffic management



The new Nexus Family of data-center-class switches provides a Unified Fabric and I/O for hosts, targets, and networks



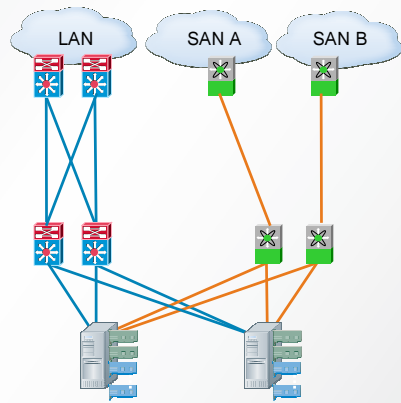
MDS 9500s interconnect with the Cisco Nexus Unified Fabric allowing storage targets to connect to every host

## Unified Fabric Evolution

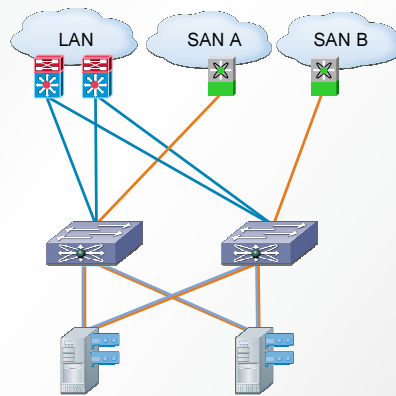


# I/O Consolidation

*Today*



*I/O Consolidation with FCoE*



Enhanced Ethernet and FCoE Ethernet FC

© 2007 Cisco Systems, Inc. All rights reserved.

27

## Summary



© 2007 Cisco Systems, Inc. All rights reserved.

28

## Summary



Data Centers are evolving  
requiring new architecture

Cisco is delivering on Data Center  
3.0 through continuing innovation in  
both the Cisco Nexus, Cisco  
Catalyst and Cisco MDS families

Cisco has a broad portfolio of  
products to meet data center  
requirements for consolidation,  
virtualization, and automation

## Additional Informations

Cisco Data Center Solutions web page

<http://www.cisco.com/go/datacenter>

Cisco Nexus family web page

<http://www.cisco.com/go/nexus>

Cisco Nexus 5000 web page

<http://www.cisco.com/go/nexus5000>

Cisco Nexus 7000 web page

<http://www.cisco.com/go/nexus7000>

Cisco Data Center Ethernet web page

<http://www.cisco.com/go/dce>

## Additional Informations

- PCI Express  
[http://en.wikipedia.org/wiki/Pci\\_express](http://en.wikipedia.org/wiki/Pci_express)
- IEEE 802.3  
<http://standards.ieee.org/getieee802/802.3.html>
- Improvements to Ethernet  
<http://www.nuovasystems.com/EthernetEnhancements-Final.pdf>
- IEEE 802.1 activities  
<http://www.ieee802.org/1/files/public/docs2007/new-cm-barrass-pause-proposal.pdf>  
<http://www.ieee802.org/1/files/public/docs2007/new-cm-pelissier-enabling-block-storage-0705-v01.pdf>  
<http://www.ieee802.org/1/files/public/docs2007/au-ko-fabric-convergence-0507.pdf>  
<http://www.ieee802.org/1/pages/802.1au.html>  
<http://www.ieee802.org/1/files/public/docs2008/az-wadekar-dcbcxp-overview-rev0.2.pdf>
- FCoE  
<http://www.fcoe.com/>  
<http://www.t11.org/>  
<http://www.open-fcoe.org/>  
[http://www.fibrechannel.org/OVERVIEW/FCIA\\_SNW\\_FCoE\\_WP\\_Final.pdf](http://www.fibrechannel.org/OVERVIEW/FCIA_SNW_FCoE_WP_Final.pdf)
- TRILL  
<http://www.ietf.org/html.charters/trill-charter.html>



CISCO

Thank You!