

SNIA

STORAGE NETWORKING INDUSTRY ASSOCIATION

EDUCATION

Unification of Data Center Fabrics with Ethernet

“A High Speed Peek into the Future!”

Matthew Brisse, Dell Inc.

SNIA Legal Notice

- The material contained in this tutorial is copyrighted by the SNIA.
- Member companies and individuals may use this material in presentations and literature under the following conditions:
 - Any slide or slides used must be reproduced without modification
 - The SNIA must be acknowledged as source of any material used in the body of any document containing material from these presentations.
- This presentation is a project of the SNIA Education Committee.

Unification of Data Center Fabrics with Ethernet: A High Speed Peek into the Future!

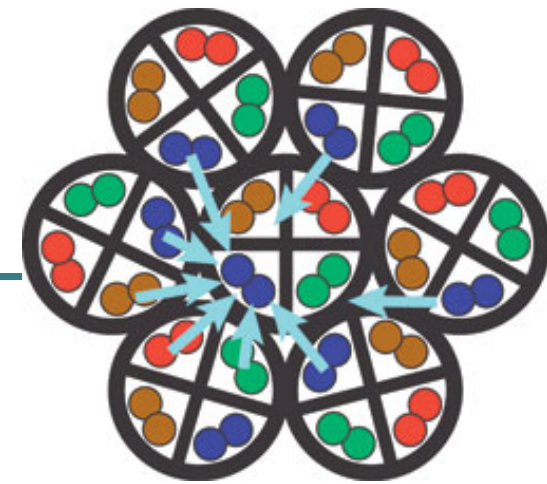
- A key challenge that the network and storage industries are now facing is the abundance of new high speed interconnect protocols proposed for future data center applications. In this presentation we take a peek into what the future may hold for high speed fabrics and investigate the potential for their unification. We will provide a market and technical overview of the competitive landscape for next generation 10Gb technologies with particular focus on the operational characteristics and implementation aspects of Ethernet.

10GE Technology

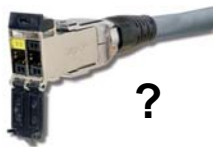
- **10GE is 10 Gbps Ethernet**
 - It has four different types of physical layer interfaces
 - 1m: Backplane (802.3ap),
XAUI (KX4) – 2.65GB/lane, 4 lane = 10GB
KR (Serial) – 10 GB 1 lane
 - 100m to 10km: Optical (SR, LR, LRM, ER)
 - 15m: CX-4 (Infiniband Type Cable-XAUI)
 - 55-100m 10GBase-T Copper (802.3an), over CAT6, CAT6a or CAT7
 - Offload **MUST** be a part of any 10GE solution
 - A “dumb-only” 10GE NIC is of limited value
 - 10GE readily consumes compute and memory resources (unless offloaded)
 - Less latency and greater bandwidth (both are important)
- **10GE is more than just another, interchangeable physical interface**
 - Offload and system characteristics are key to plumbing 10GE properly

Copper Cable Options

Category	Type	Spectral B/W	Length	Applications	Notes
Cat5	UTP	100MHz	100m	100Base-T	Common
Cat5e	UTP	100MHz	100m	1000Base-T	Common
Cat6	UTP	250MHz	55m	10GBase-T	Emerging
Cat6a	UTP	400MHz	100m	10GBase-T	Emerging
Cat7	ScTP	600MHz	100m	10GBase-T	Emerging

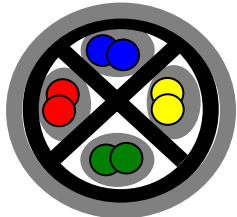


Alien Crosstalk



?

\$1.25/ft



CAT7

Longitudinal spacer

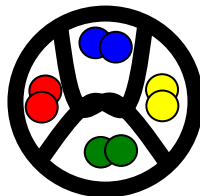
Wrapped pairs, Shielded

SHIELDED



RJ45

\$0.65/ft



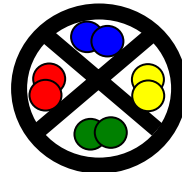
CAT6a

Longitudinal spiral spacer



RJ45

\$0.15/ft



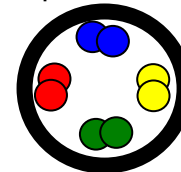
CAT6

Longitudinal spacer



RJ45

\$0.10/ft



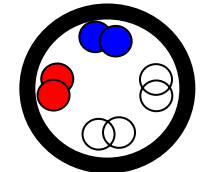
CAT5e

4 twisted pairs



RJ45

\$0.05/ft



CAT5

2 twisted pairs

UNSHIELDED

(100m)

1GE

(100m)

Center Fabrics with E

(100m)

Speed Peek into the Fu

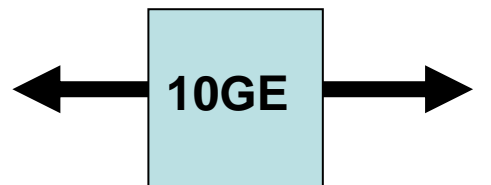
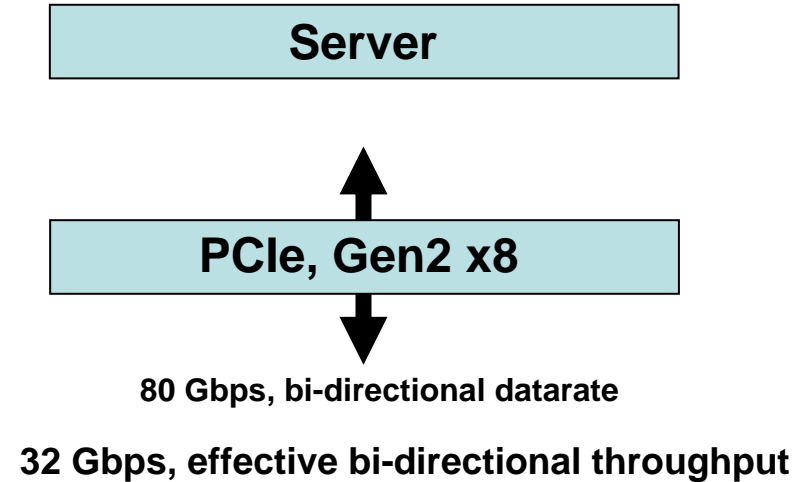
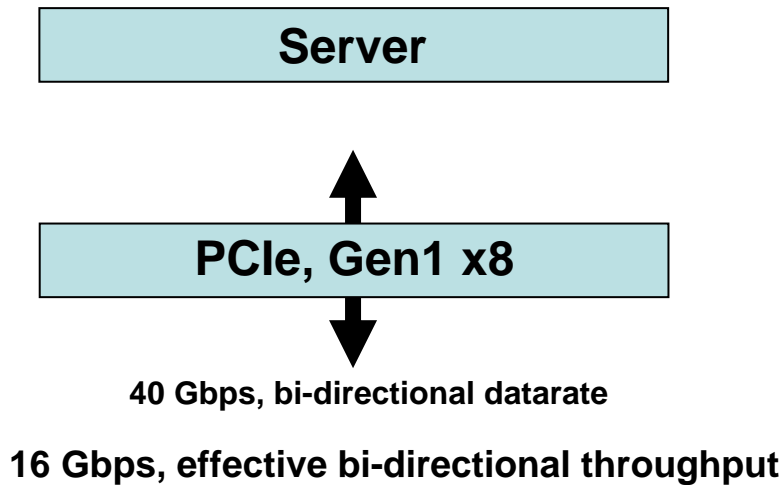
(55m)

10GE

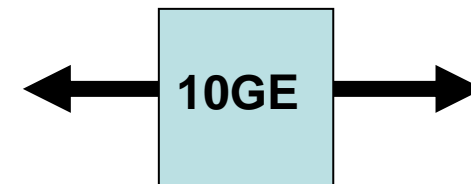
RJ45 Challenge!

PCIe and 10GE

PCI Express Generation 2 will be required to support full 10GE with x8
(Current 10GE NICs can't realize full throughput due to PCIe Gen 1 bus speed limitations but PCIe Gen 2 will realize the maximum throughput)



20Gbps bi-directional

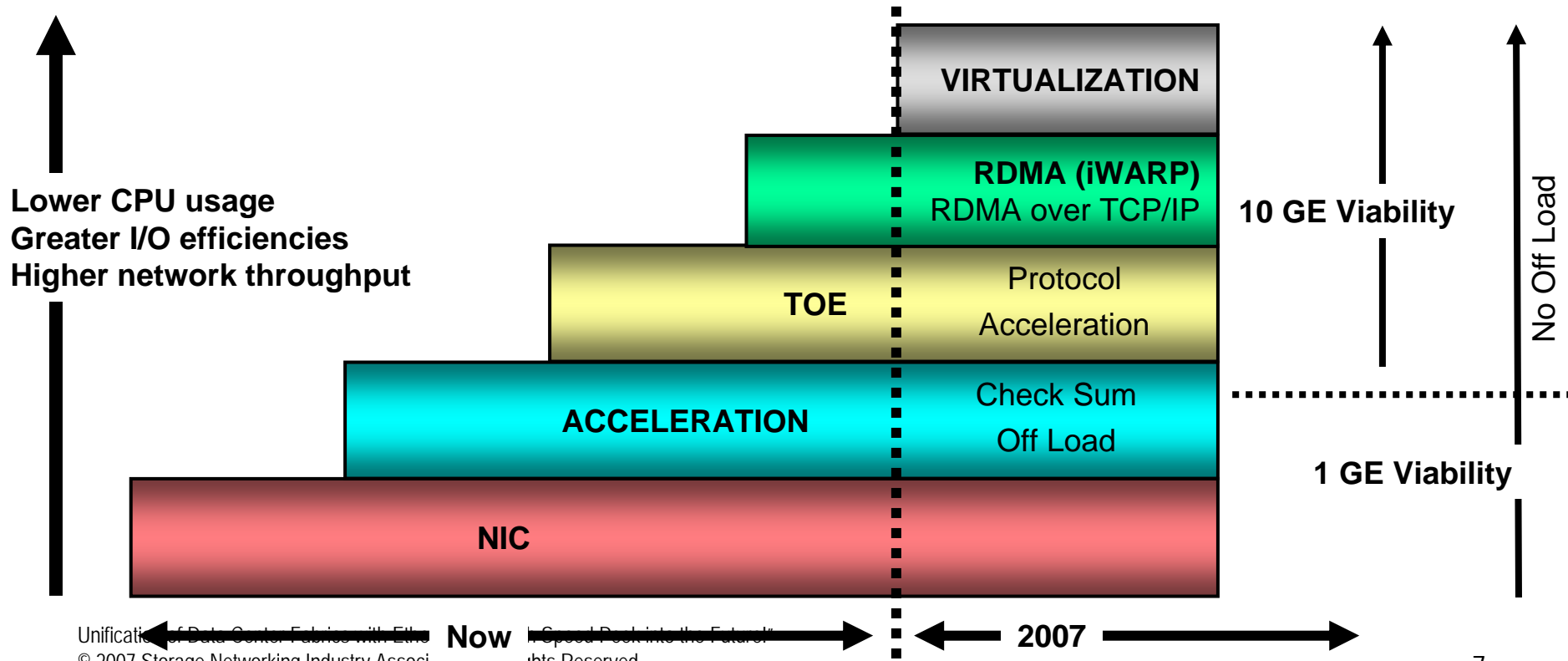


20Gbps bi-directional



10 GE Requires Multi-Function Offload Capabilities

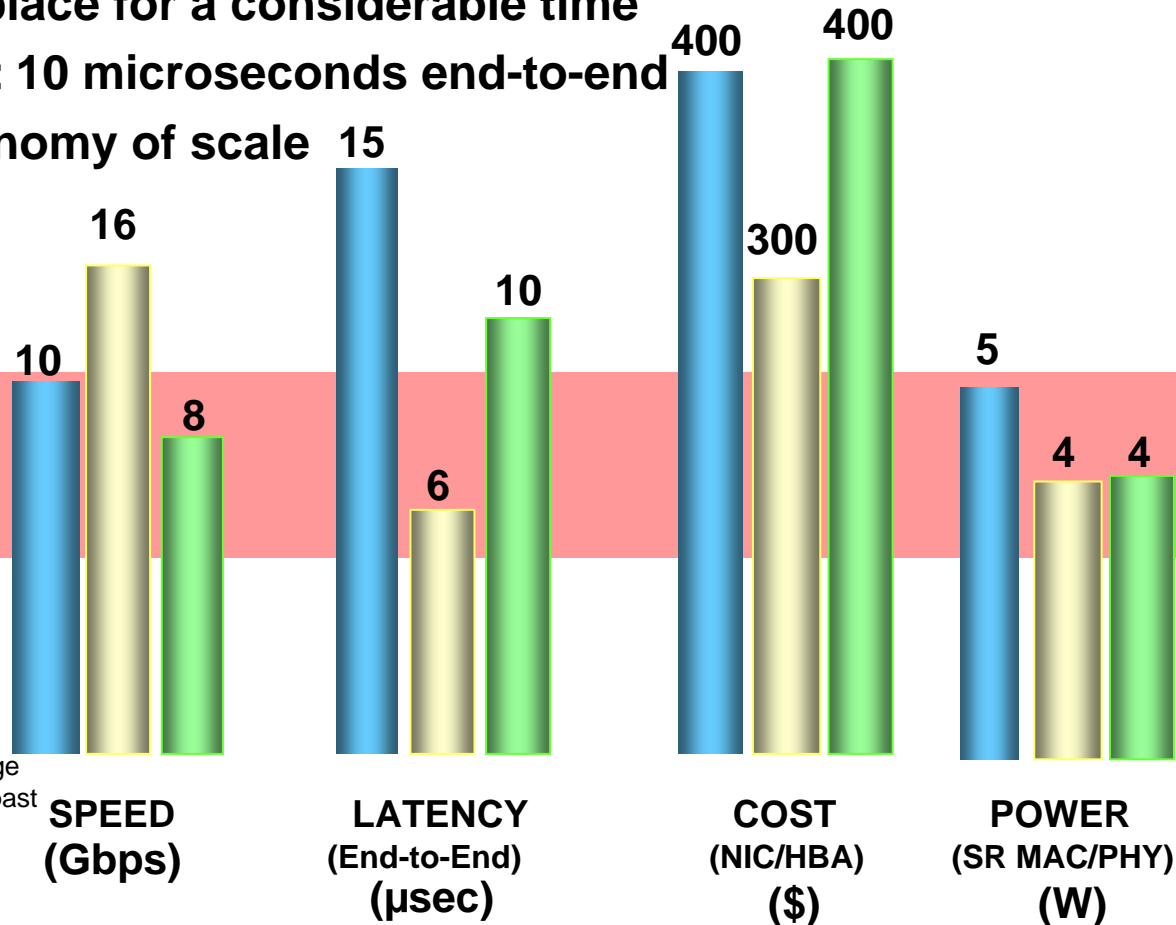
- Simple NICs offload nothing
- Stateless offload accelerates packet processing (RSS, LSO, TSO, Checksum)
- Stateful offload provides protocol and packet processing (TOE, Chimney, RDMA, iWARP)
- RDMA abstracts the underlying transport and achieves direct memory to memory data placement
- Virtualization mediates access to multiple instantiations of limited hardware resources



10GE Overall Technology Comparison

- 10GE is an acceptable technology for fabric convergence
- IB and 10GE bandwidth both have visibility to 100GE
- FC will continue to have a place for a considerable time
- 10GE latency will settle at < 10 microseconds end-to-end
- Cost will come down – Economy of scale

UNIFIED FABRIC
SWEET SPOT

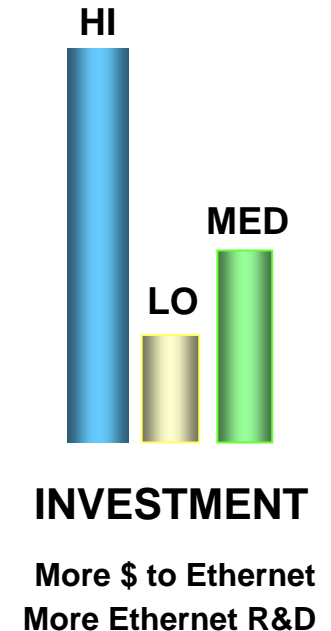
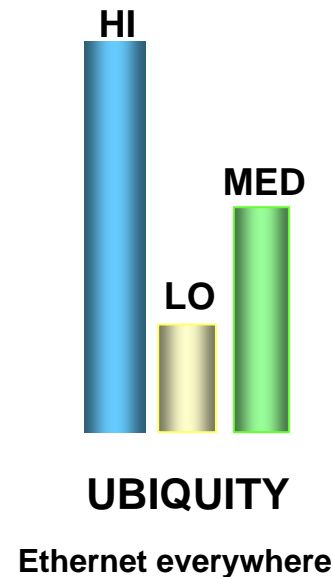
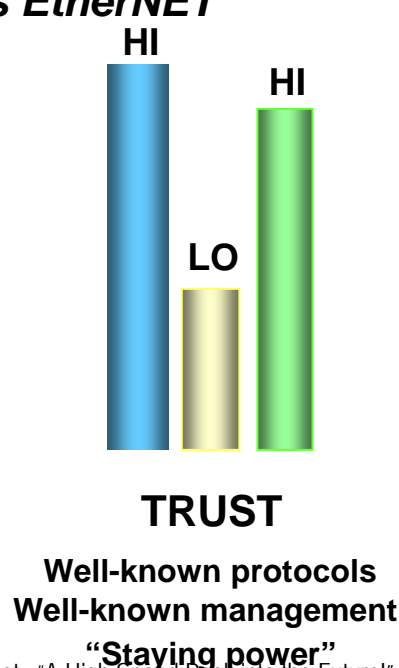
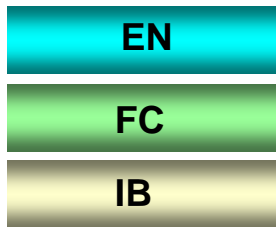


EN	Fast, some latency good enough
FC	Faster, less latency, storage
IB	Fastest, less latency, cluster + storage "was" the fabric convergence in the past

10GE non-Technology Comparison



- FC and IB will both perform “better” than 10GE in cluster and storage fabrics, respectively
- 10GE will, however, be “good enough” to potentially gain market share from IB and FC
- Non technical factors will drive 10 GE adoption and continued evolution
 - *Ethernet trust, ubiquity and robust investment climate are important driving factors*
- Ethernet always incorporates the best from competing technologies
 - *EtherNOT becomes EtherNET*



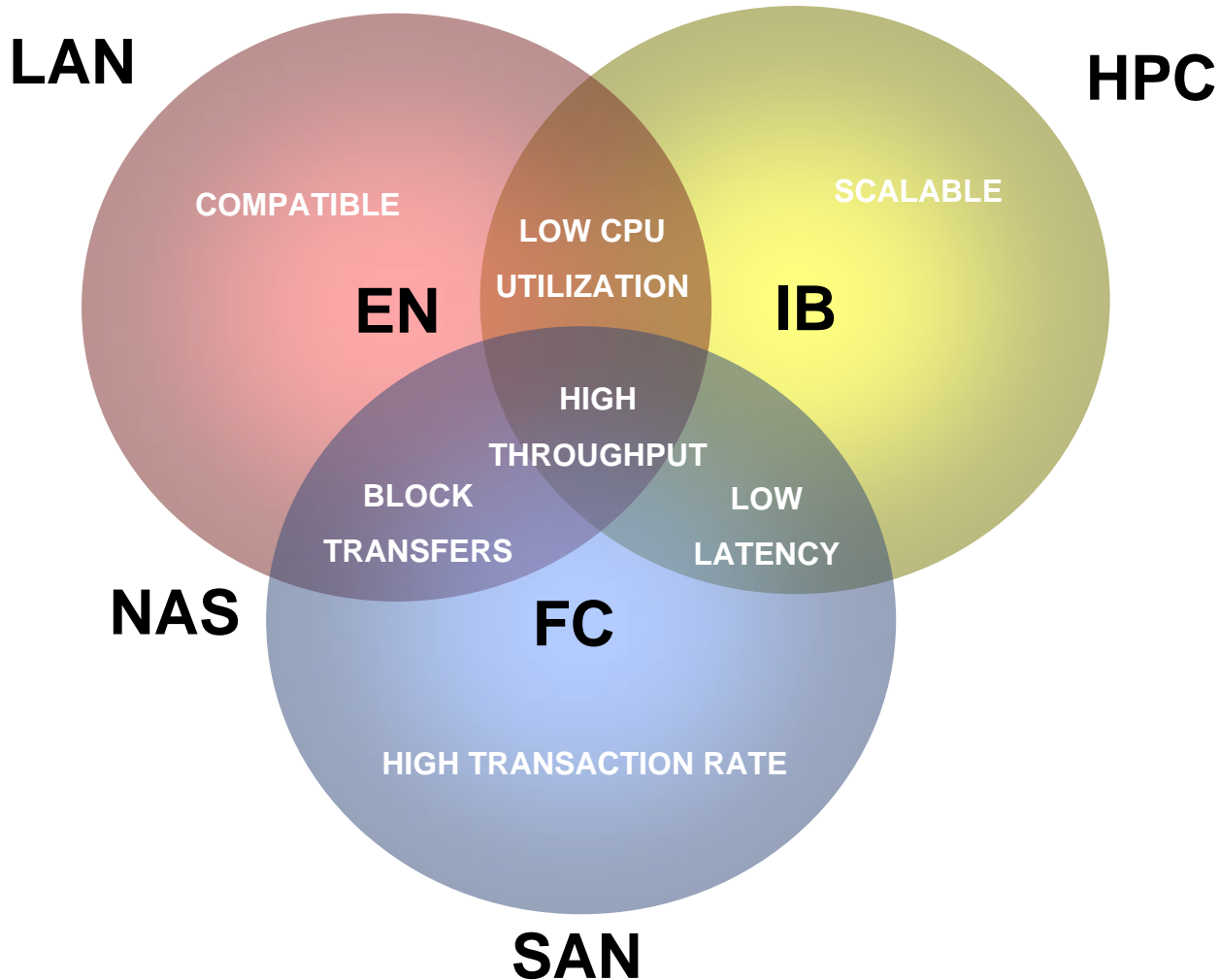
Unified Fabric

“A peek in the future”

- **Reduce** the number of **disparate networks** and Data Center **complexity**
- Provide a **seamless model** for networking, clustering and storage applications
 - Not to mention thin client, KVM and other usages
- **Consolidate Management** protocols over a common IP framework
- Leverage **economies of scale** in terms of silicon integration, interoperability and product development
- Provide support for legacy network architectures in the transition to a unified fabric architecture

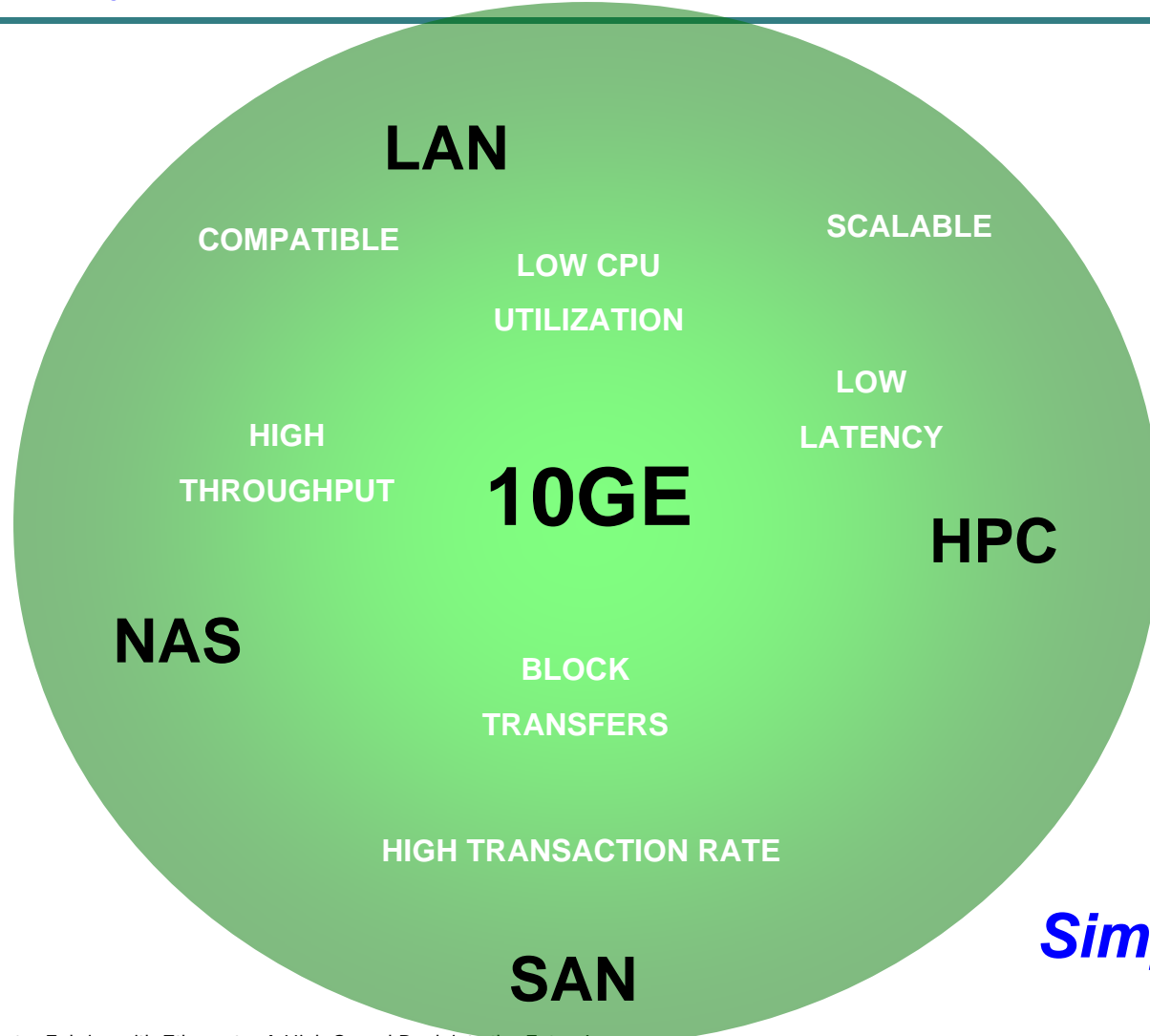
Unified Fabric Vision

Making Many Fabrics.....



Unified Fabric Vision

Making Many Fabrics..... One



Simplifying

Emerging Trends

- **Clustering**
 - Improves cost-performance over large scale cache coherent systems
 - Brings flexibility to the data center
 - Provides better fault tolerance
- **Grid**
 - Academic processing Grids tried and tested
 - Commercial deployment in niche areas only
 - Storage Grids now in discussion
- **Object Storage**
 - Completing standardisation
 - New paradigm of intelligent storage
 - Requires converged compute & storage platform
- **Systems Management**
 - Keyboard Video & Mouse (KVM) already available and other tools are coming
- **System Level Virtualization**
- **Ethernet is rapidly becoming the key to enterprise systems convergence**

Industry Initiatives

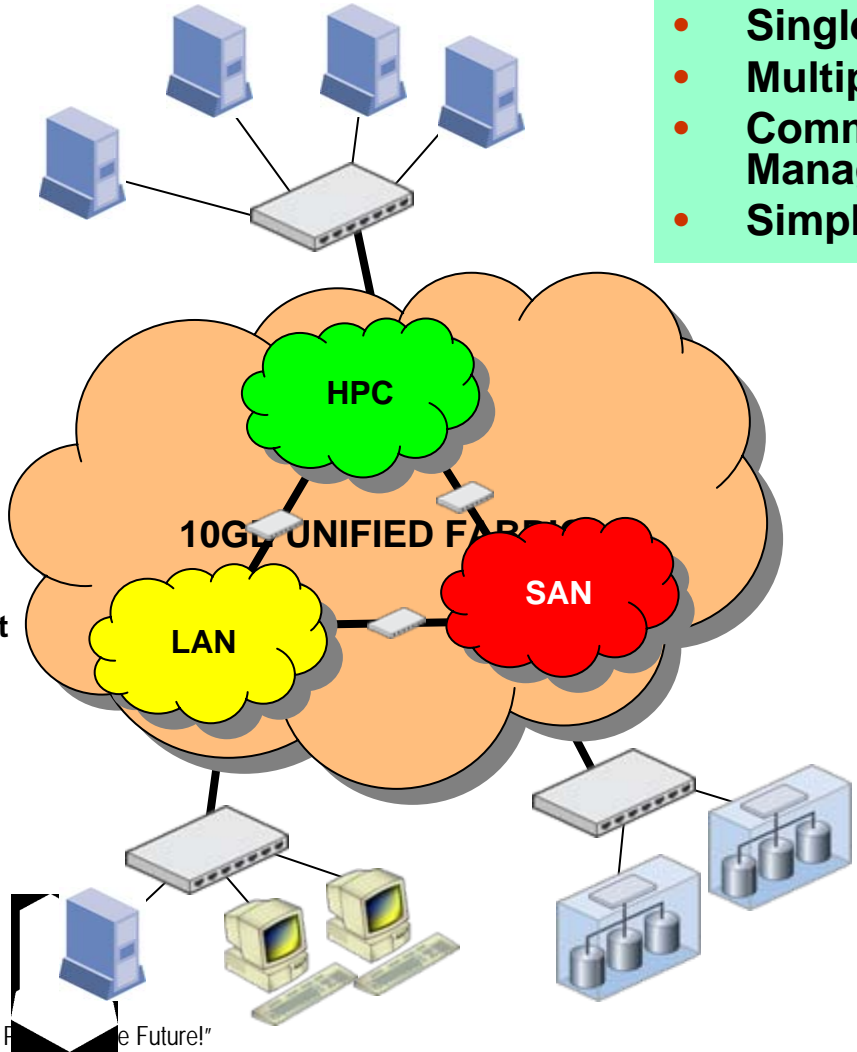
- Blurring of compute and storage is now happening
 - Traditional view of totally discrete functions is getting old
 - New applications require dynamic configurations
 - Lightweight blades clustered with shared access to I/O resources are enabling convergence
- Systems Virtualization
 - Protected virtual memory architectures return to give scale up virtualisation
 - Scale out virtualization still requires major developments in the industry but everyone expects it to happen
- Server re-purposing paradigm maturing
 - The 3 tier Data center model is now passing its sell by date
- Evolving Ethernet enables system wide convergence

Target 10GE Applications

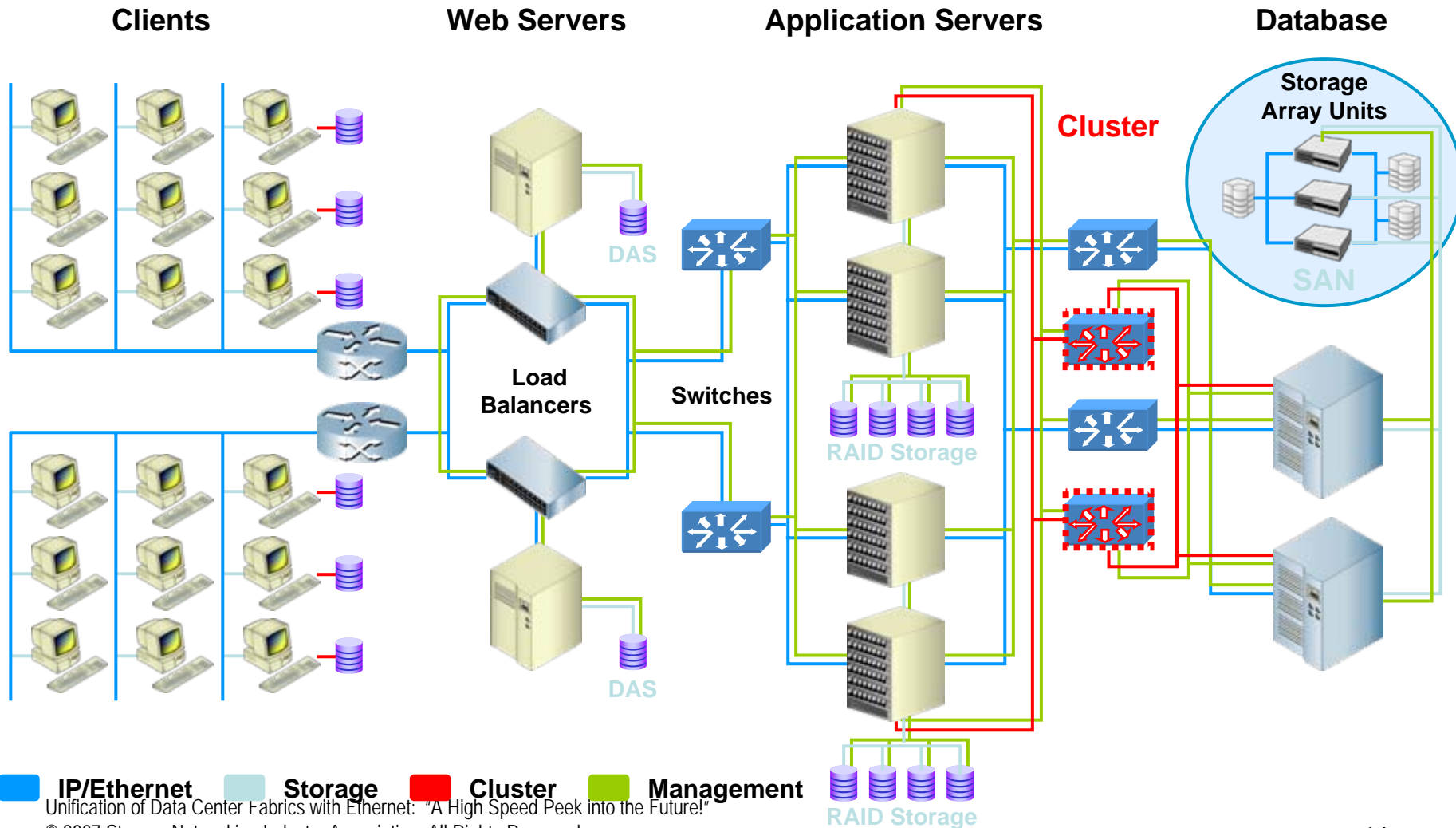
10GE provides the basis for a unified datacenter fabric

- **Servers**
 - HPC
 - server-to-server interconnect
 - Virtualization
 - increased bandwidth
 - High-Performance Transactions
 - financial applications
 - IPTV streaming media applications
 - Digital Cinema
 - web farms
- **Switches**
 - Edge Layer
 - server-to-switch uplink
 - iSCSI storage array to switch
 - Aggregation & Core Layer
 - switch-to-switch interconnect
- **iSCSI Storage Arrays**
 - General Purpose
 - array-to-switch connectivity

- **Single Adapter**
- **Multiple Offloads**
- **Common Management**
- **Simplicity**



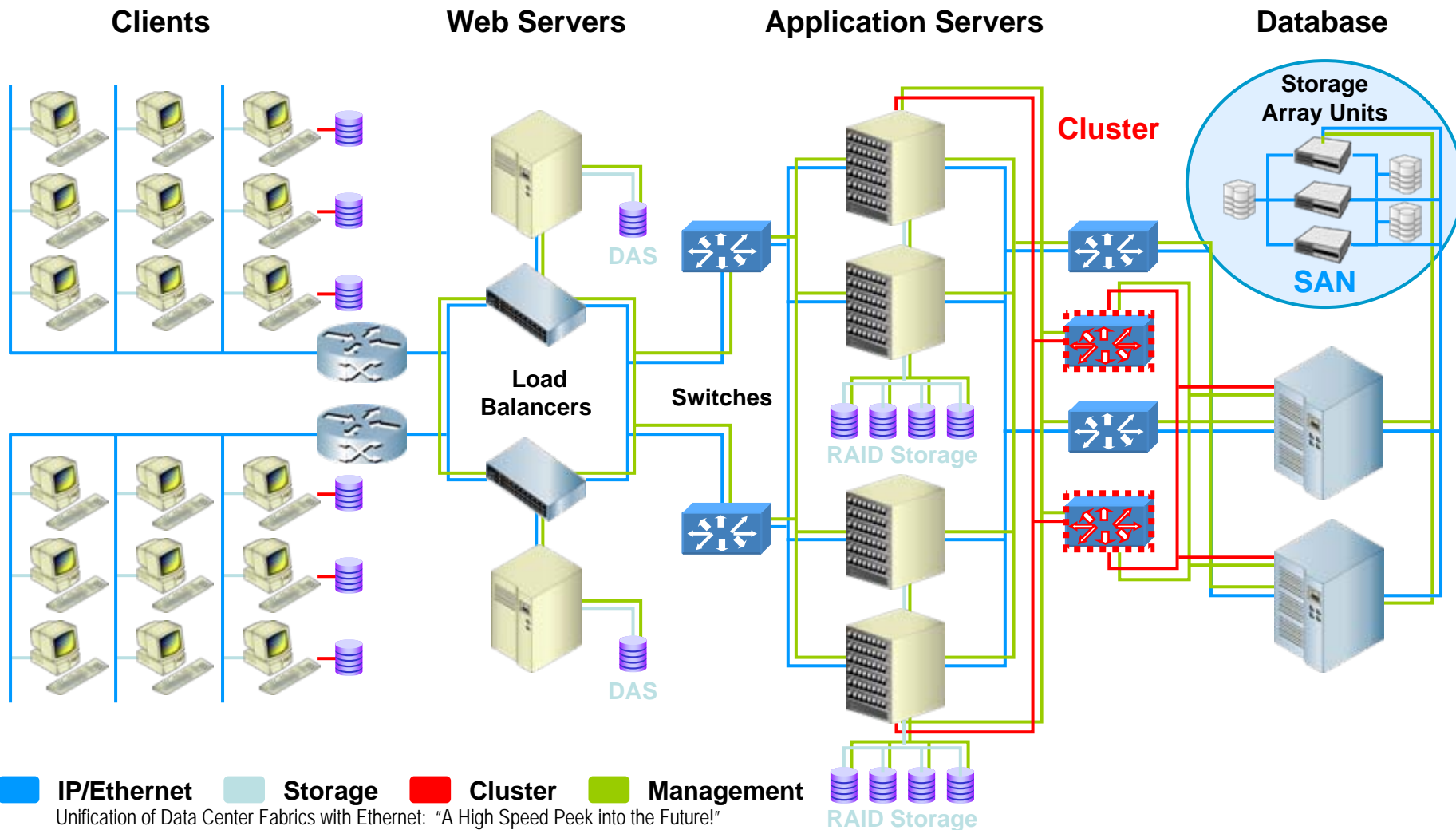
Enterprise Network Today



Unification of Data Center Fabrics with Ethernet: "A High Speed Peek into the Future!"

© 2007 Storage Networking Industry Association. All Rights Reserved.

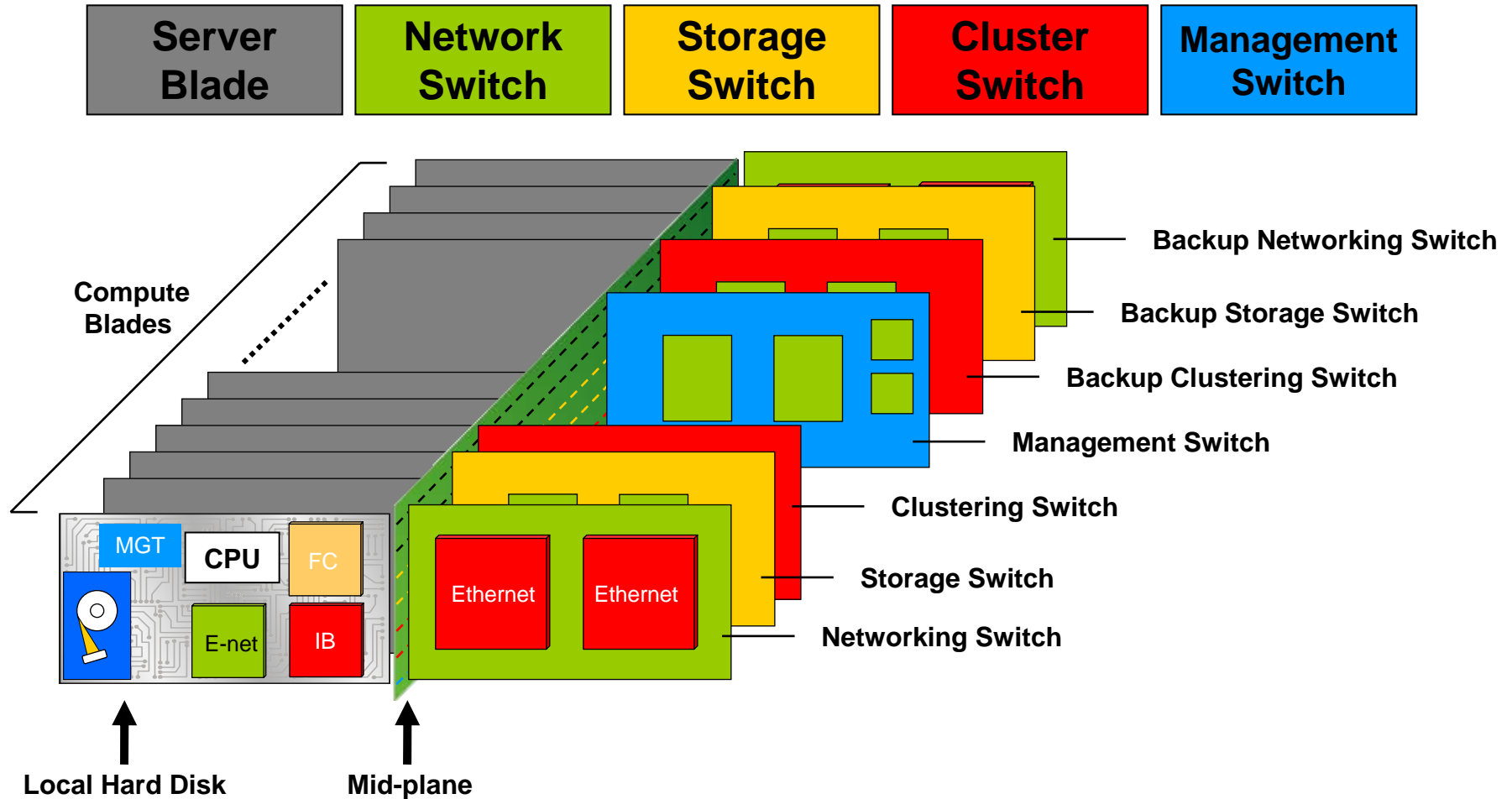
Converged Network



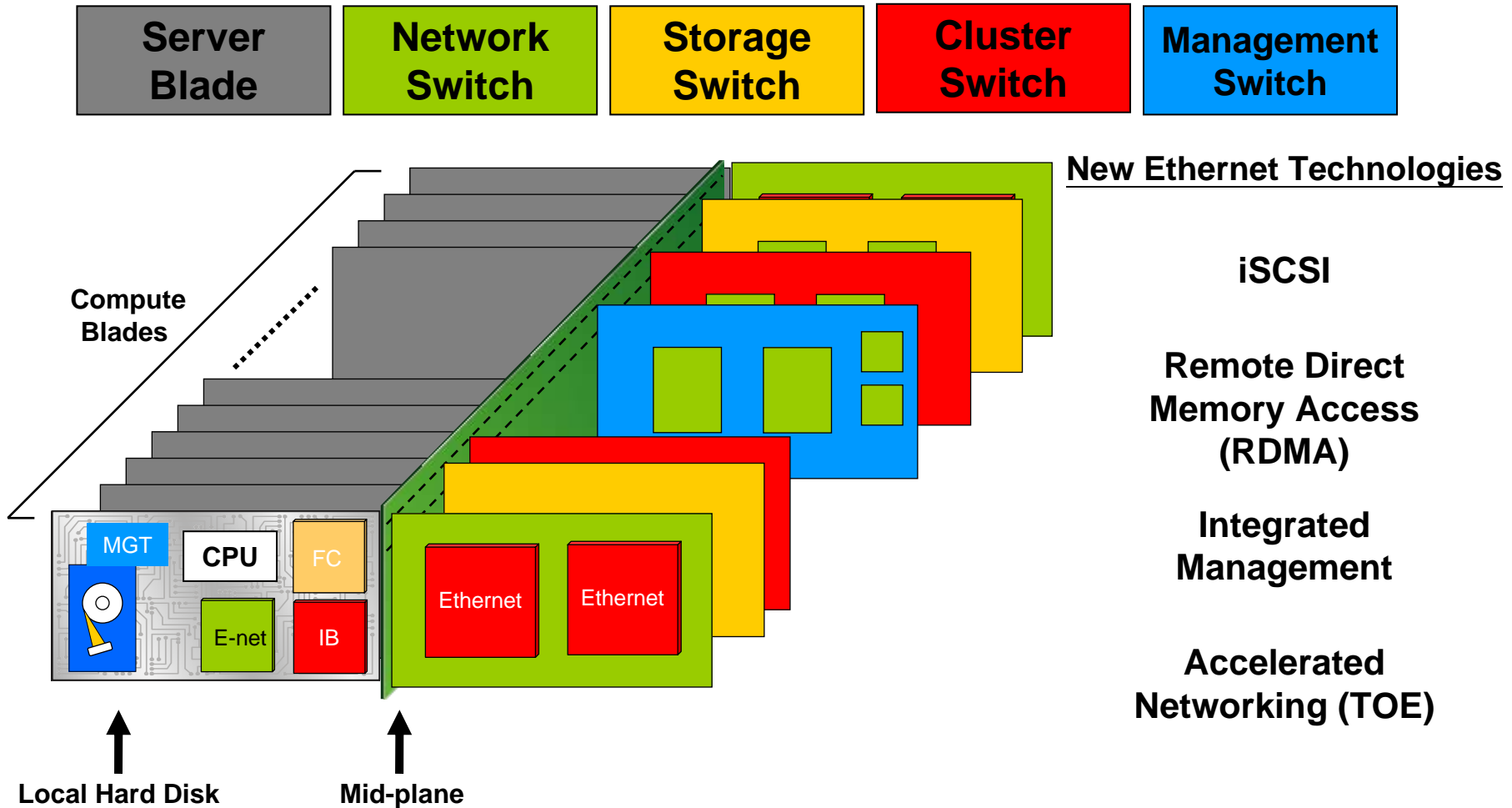
Unification of Data Center Fabrics with Ethernet: "A High Speed Peek into the Future!"

© 2007 Storage Networking Industry Association. All Rights Reserved.

Unified Fabric Example – Server Blades

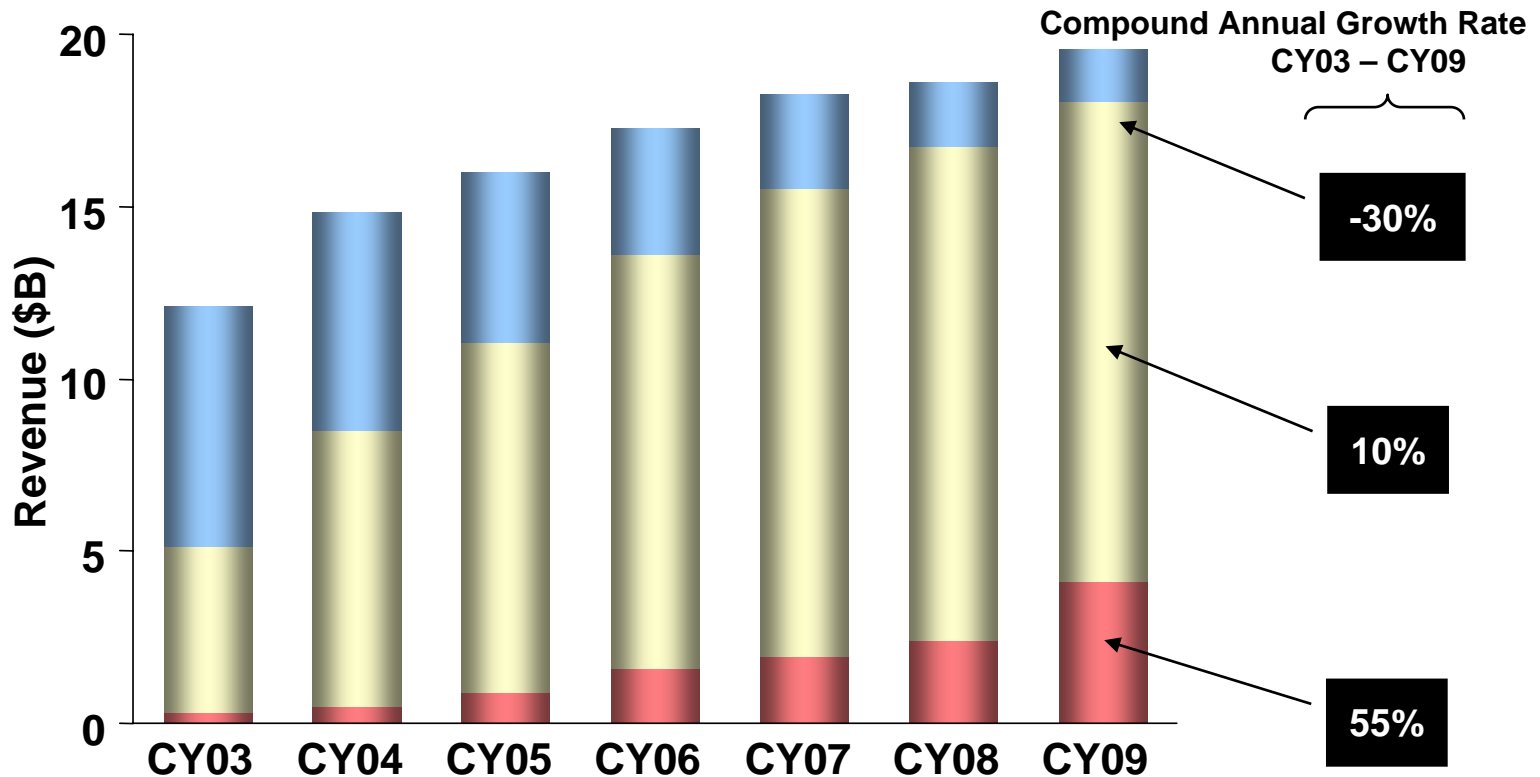


Unified Fabric Example – Server Blades



10GE Switch Market Growth

- 10GE growth (2008-09) driven by adoption of low-cost 10GBASE-T copper
- 1GE will be pushed to the edge of LANs and must be aggregated with 10GE
- 10GE will be initially positioned for inter-switch connections and clusters
- Vendors drive low-cost 10GE, accelerating adoption resulting from customer demand



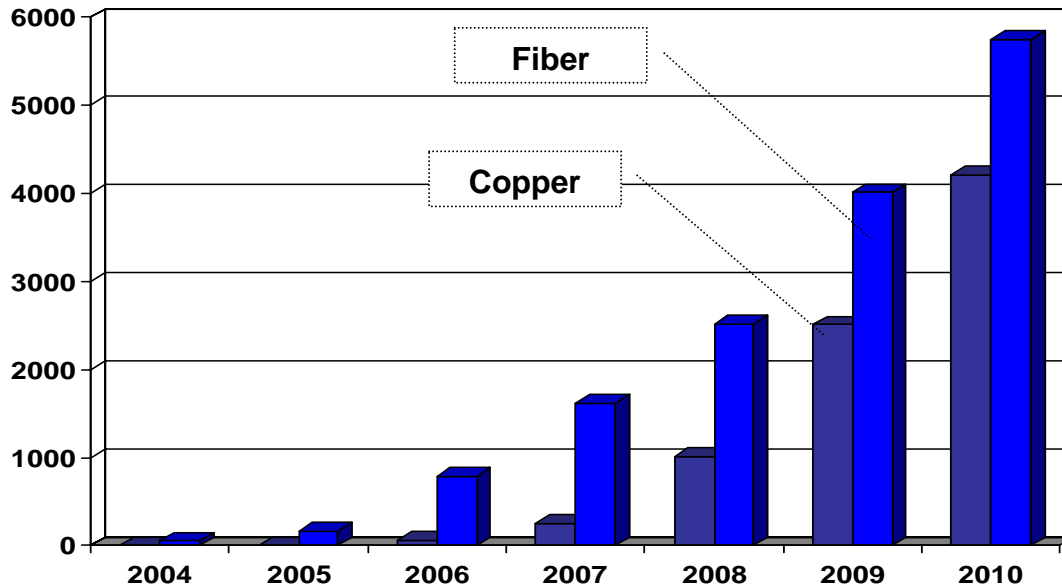
Source: IDC, 2005

Unification of Data Center Fabrics with Ethernet: "A High Speed Peek into the Future!"

© 2007 Storage Networking Industry Association. All Rights Reserved.

10GE Market

5x port growth from 2005 to 2006; **10m** ports, **\$4B** market in 2010

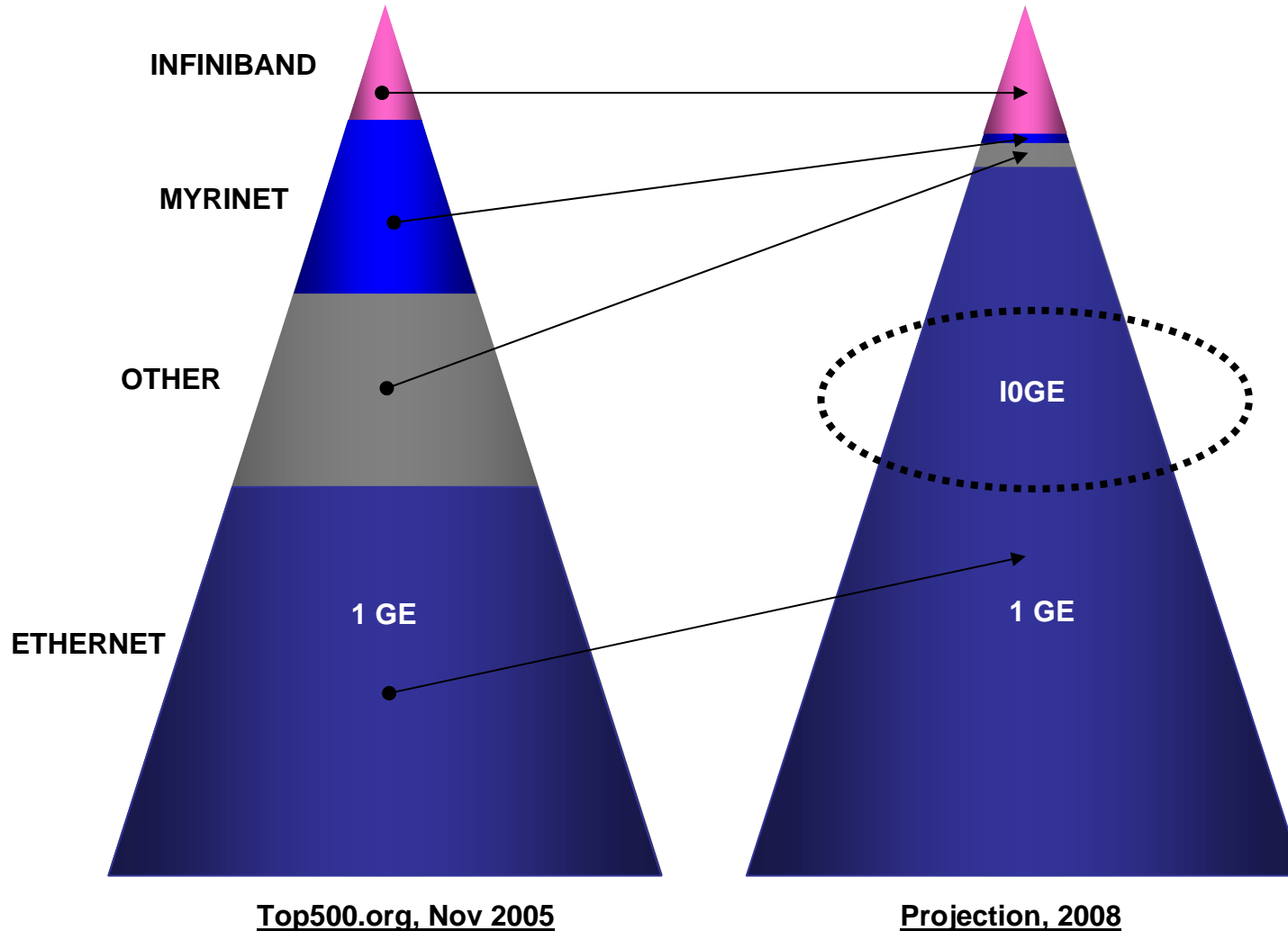


Available market data represents growth in the 10GE switch port market, not the NIC or other 10GE markets. The growth in 10GE infrastructure, however, implies growth in other 10GE markets, including server port shipments.

2008 – Year of copper
10W today, 5W power '08
2-3W down on the Motherboard

000's switch ports	2004	2005	2006	2007	2008	2009	2010
Copper	1	6	69	255	1,013	2,509	4,209
Fiber	57	163	785	1,616	2,511	4,008	5,748
Total	58	169	854	1,871	3,524	6,517	9,958
% Copper	1%	4%	8%	14%	29%	39%	42%

Cluster Fabric Trends

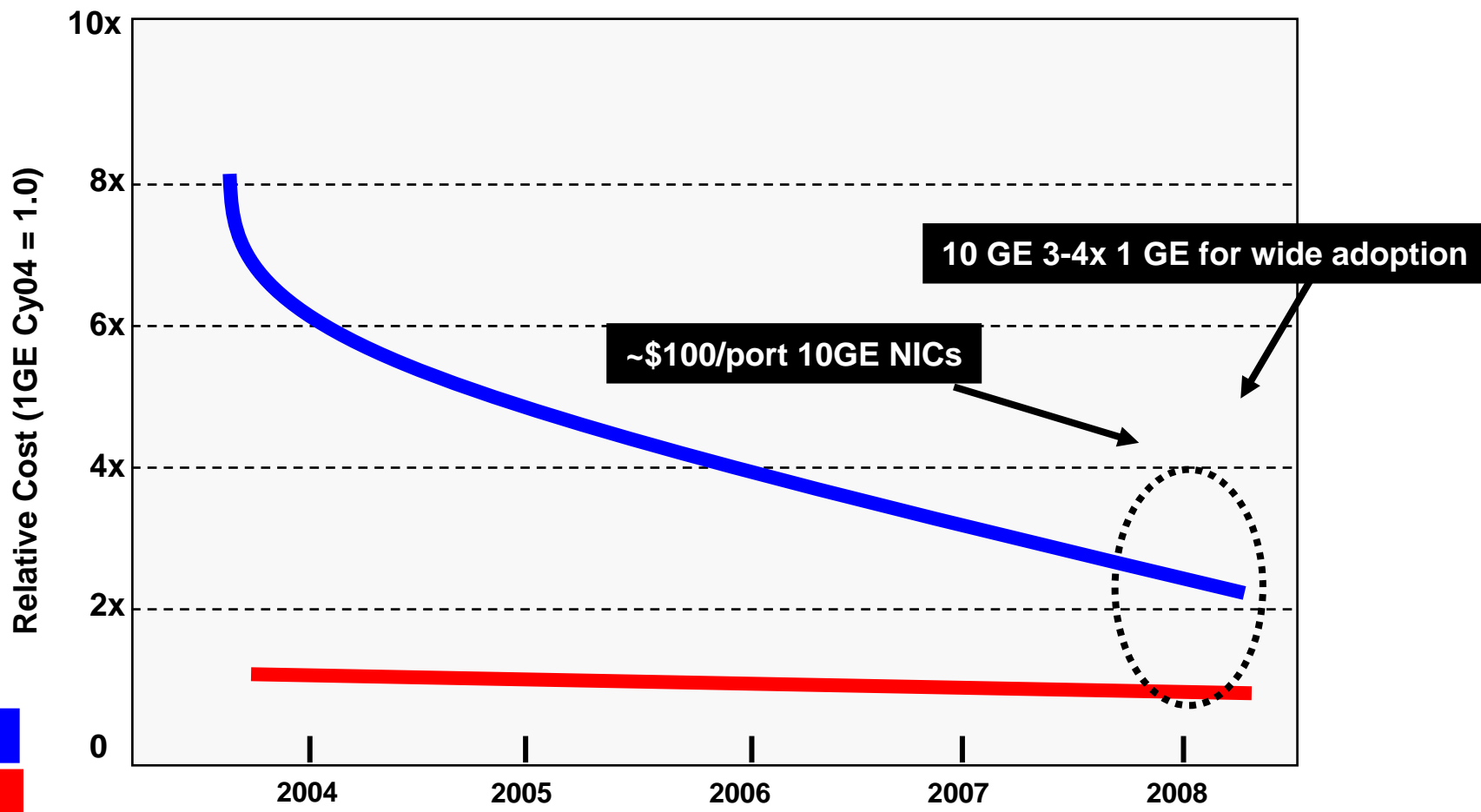


Unification of Data Center Fabrics with Ethernet: "A High Speed Peek into the Future!"
© 2007 Storage Networking Industry Association. All Rights Reserved.



NIC 10GE Cost Trends

1GE vs. 10GE RELATIVE PRICE TARGET

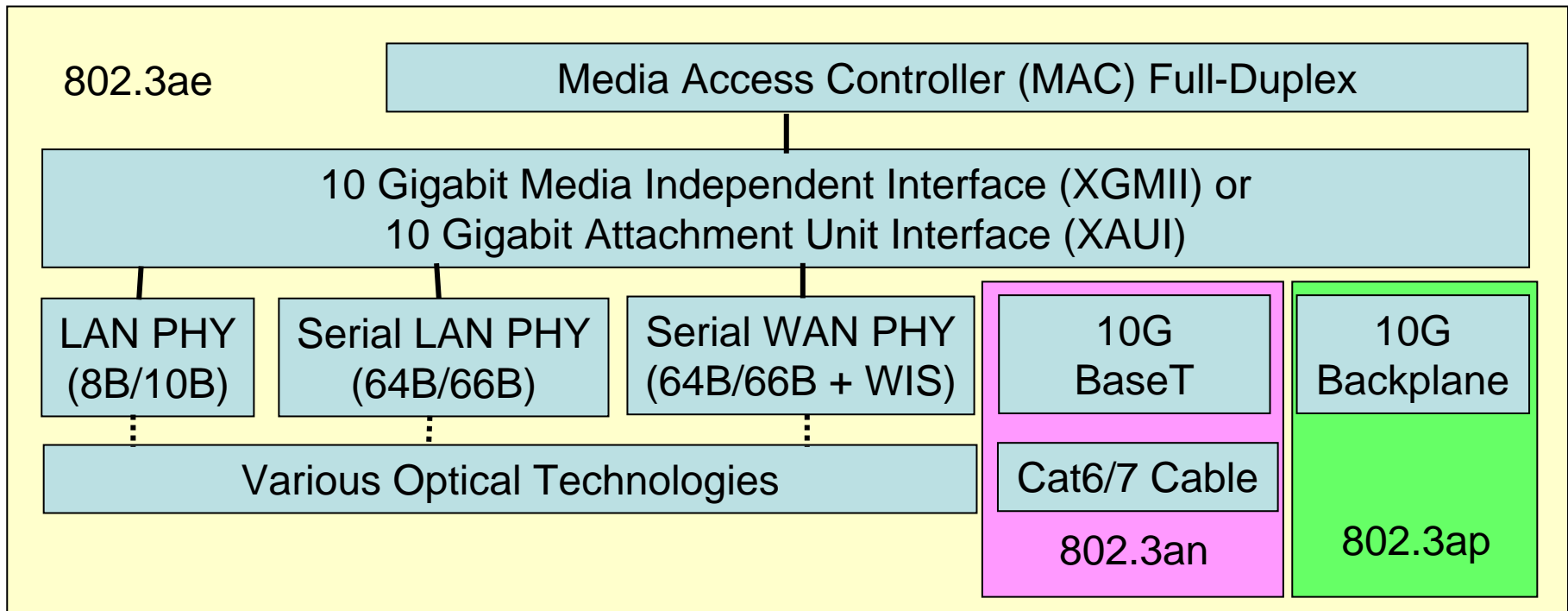


10GE

1 GE

Unification of Data Center Fabrics with Ethernet: "A High Speed Peek into the Future!"
© 2007 Storage Networking Industry Association. All Rights Reserved.

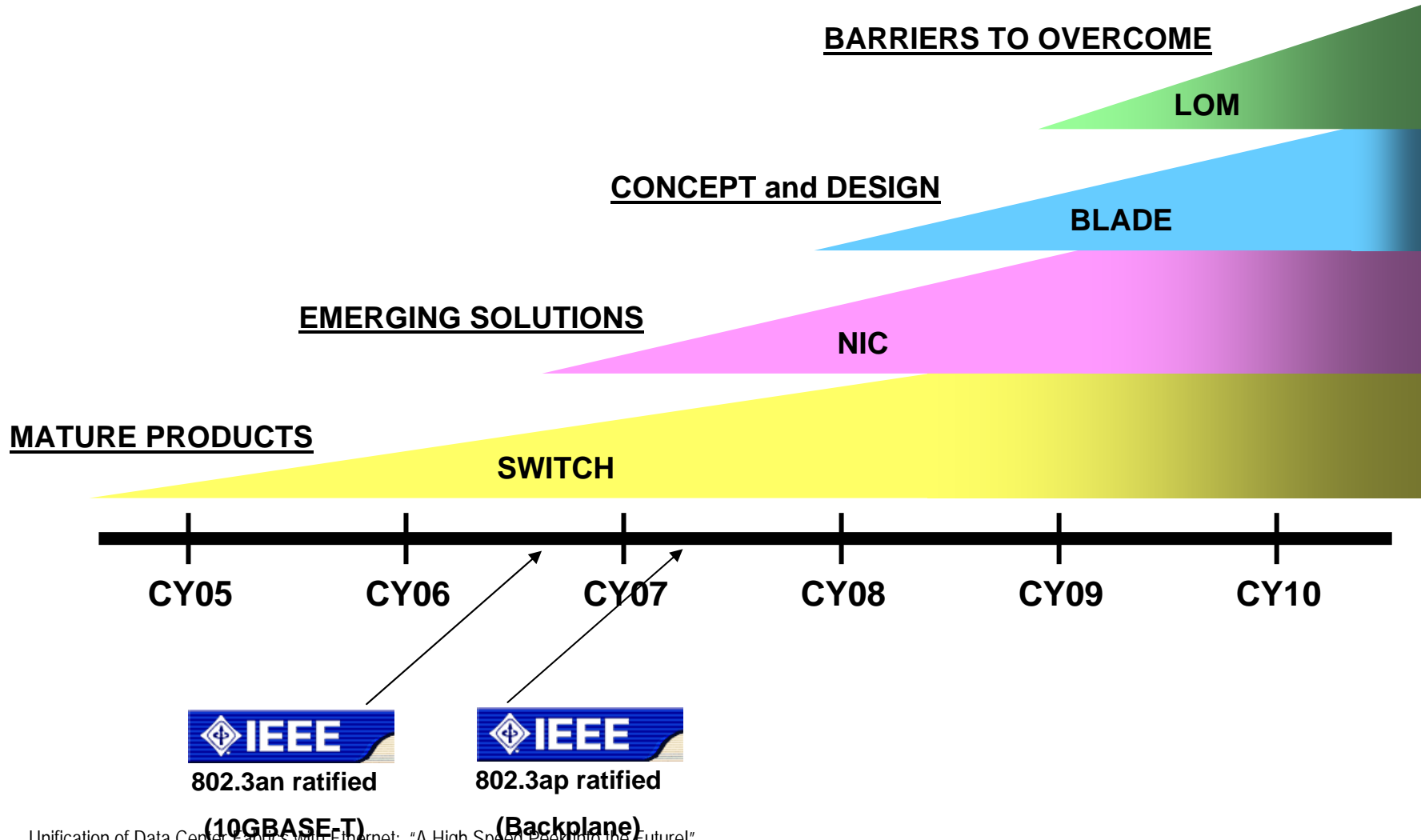
10GbE Standards



Layer 2

802.3an and 802.3ap make 10GbE even more compelling for data center applications

10GE Copper/Backplane Standards and Opportunity Timeline



Changes with 10GbE

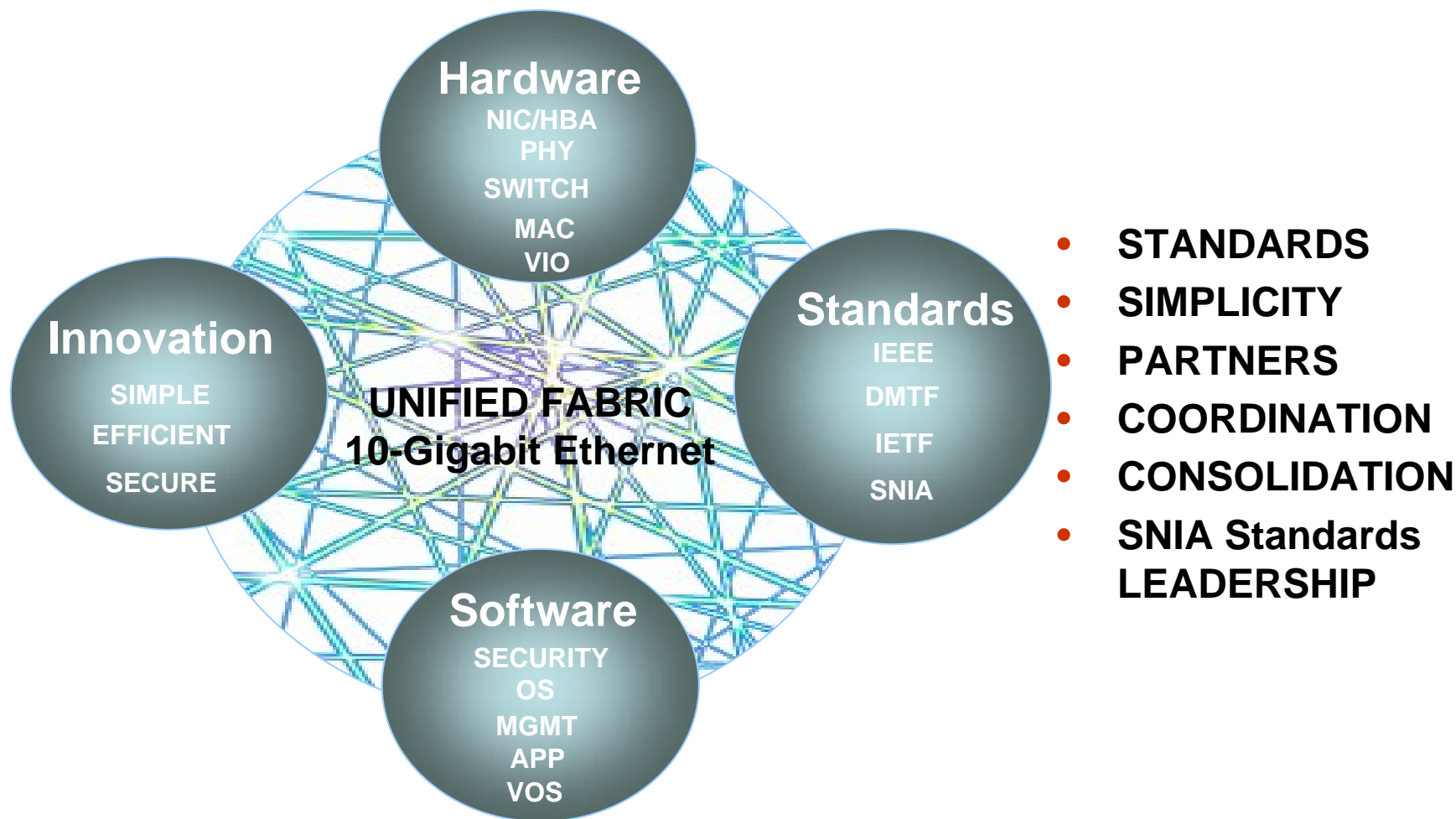
- 802.3ae defines 10GbE
 - 10Gbps Data Rates
 - Full-duplex only; no need for carrier-sensing multi-access / collision detection (CSMA/CD)
 - Optical Physical Layer
 - LAN PHY and WAN PHY options
 - WAN PHY compatible with SONET
- 802.3an adds twisted pair cabling
 - Cat 6 and Cat 7
- 802.3ap adds backplane specifications
 - Blade servers and communications equipment

IEEE 10GE Standards

- IEEE P802.3ae, 10GE Working Group
 - Definition of the 10GE MAC architecture and components
- IEEE P802.3an, 10GBASE-T Task Force.
 - Definition of 10GE encoding over CAT 6/7 (Copper) cable
- IEEE P802.3ap, Backplane Ethernet Task Force.
 - Definition of 1/10GE encoding over backplanes (KX, KX4, KR)
- IEEE P802.3aq, 10GBASE-LRM Task Force.
 - Definition of short haul MMF/1310nm LR encoding
- IEEE P802.3ar, Congestion Management Task Force.
 - Definition of L2 congestion discovery and avoidance protocols
- IEEE P802.3as, Frame Expansion Task Force.
 - Not “Jumbo” frames; expansion of frame envelope, not data size
- IEEE P802.3at, DTE Power Enhancements Task Force.
 - Definition of how to deliver power at 40-60W over 1 and 10GE



Unified Fabric End Goal



Call to Action

- Get involved!!
 - Participate in Standards
 - Drive your vendors
 - Leverage the power of the user
- Ethernet vendors are developing new protocol acceleration features and the speed is increasing from 1G to 10G, Understand the future impact
- 10Gb Ethernet with protocol acceleration could be used a 'Unified fabric' to handle SAN, HPC, management and client network needs in addition to traditional data networking requirements.
- Be prepared for the transition! Plan today!

Q&A / Feedback

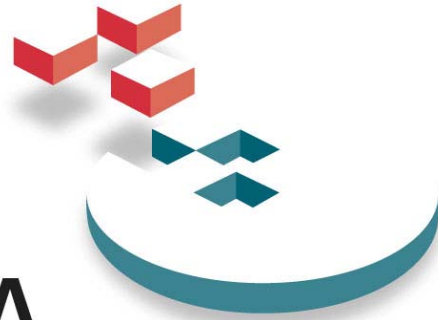
- Please send any questions or comments on this presentation to SNIA: trackapplications@snia.org

**Many thanks to the following individuals
for their contributions to this tutorial.**

SNIA Education Committee

**Allen Light
Matt Brisse**

**Robert Winter
Robert Peglar**



SNIA

STORAGE NETWORKING INDUSTRY ASSOCIATION

EDUCATION

Questions?