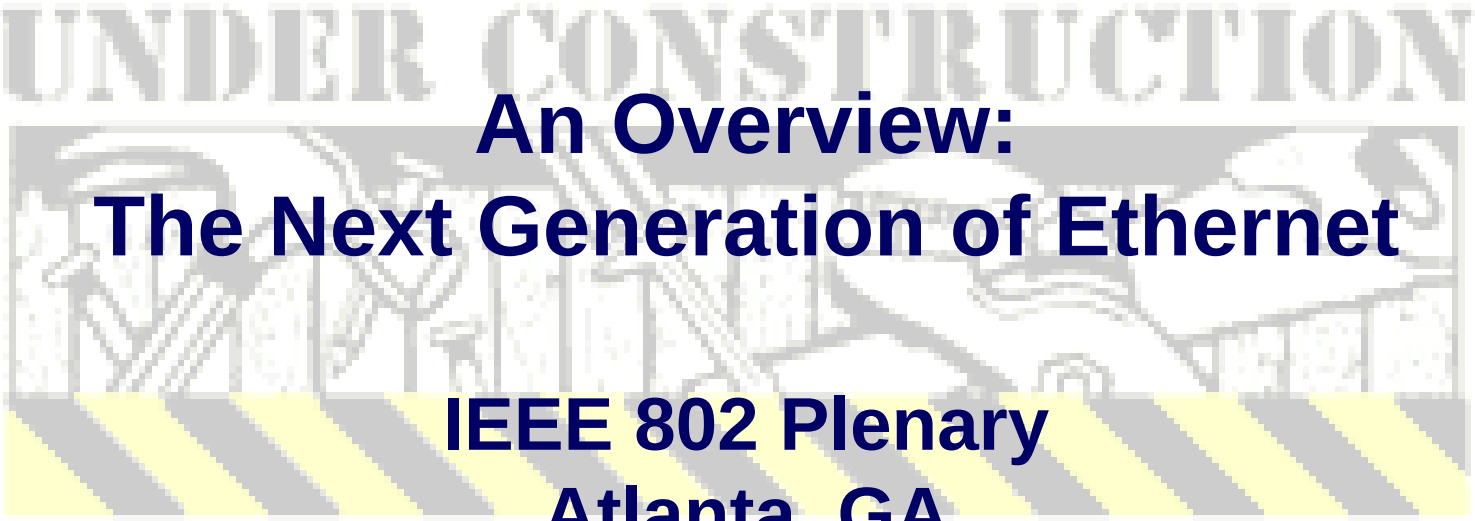


An Overview:
The Next Generation of Ethernet

IEEE 802 Plenary
Atlanta, GA
November 12, 2007

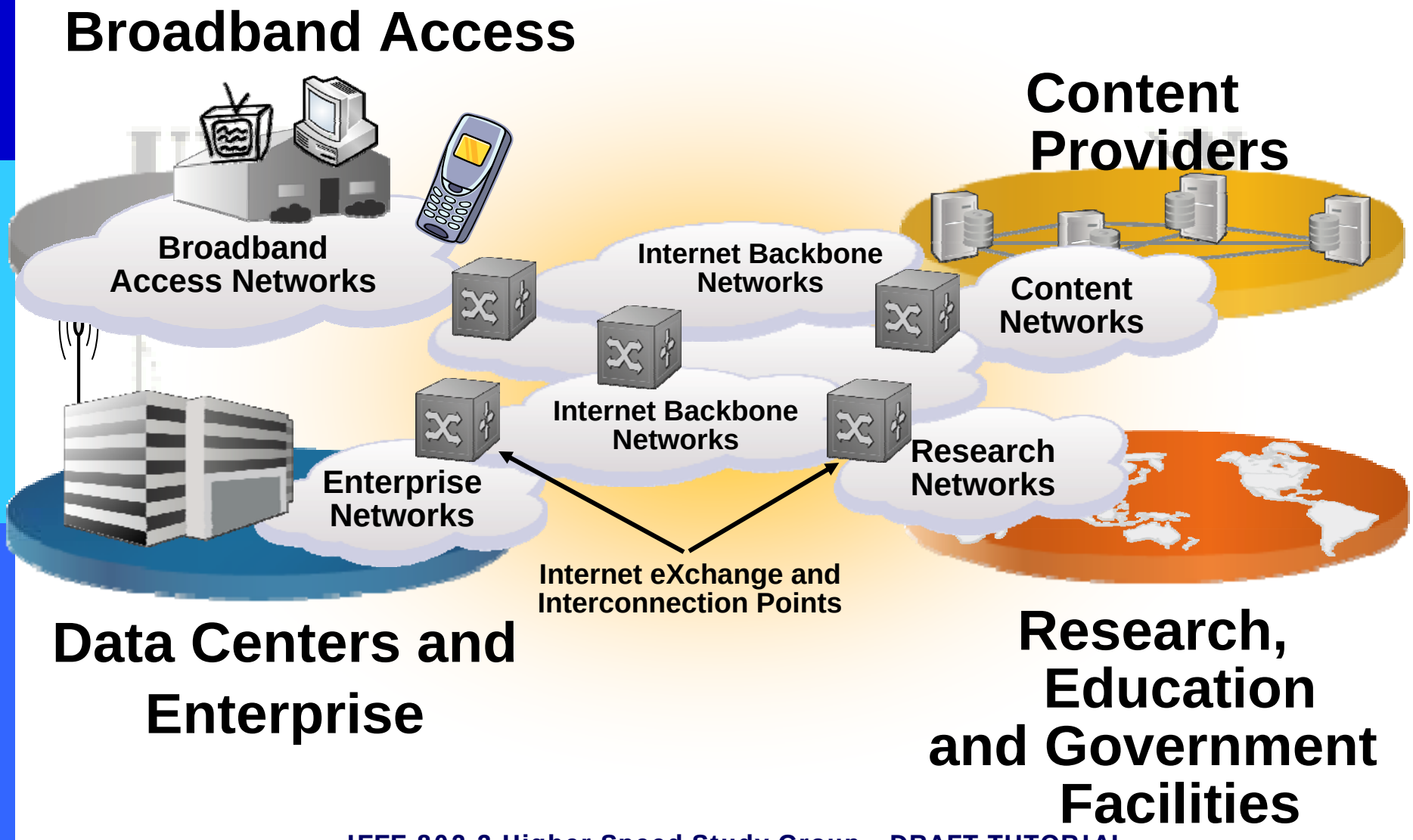
Contributors

- Howard Frazier, Broadcom
- Mark Nowell, Cisco
- Chris Cole, Finisar
- John D'Ambrosia, Force10 Networks
- John Jaeger, Infinera
- Chris DiMinico, MC Communications
- Jack Jewell, JDSU
- Adam Healey, LSI

Agenda

- Overview
- Objectives
  - Project Authorization Request
  - 5 Criteria
- The Need for a Higher Speed Ethernet
- Architecture
- Cu PMDs
  - Backplane
  - Cabling
- Fiber PMDs
  - MMF
  - SMF
- Summary

The Ethernet Ecosystem: Where is(n't) the bottleneck?



HSSG Objectives

- Support full-duplex operation only
- Preserve the 802.3 / Ethernet frame format utilizing the 802.3 MAC
- Preserve minimum and maximum FrameSize of current 802.3 standard
- Support a BER better than or equal to 10^{-12} at the MAC/PLS service interface
- Provide appropriate support for OTN
- Support a MAC data rate of 40 Gb/s
- Provide Physical Layer specifications which support 40 Gb/s operation over:
  - at least 100m on OM3 MMF
  - at least 10m over a copper cable assembly
  - at least 1m over a backplane
- Support a MAC data rate of 100 Gb/s
- Provide Physical Layer specifications which support 100 Gb/s operation over:
  - at least 40km on SMF
  - at least 10km on SMF
  - at least 100m on OM3 MMF
  - at least 10m over a copper cable assembly

Project Authorization Request Key Elements (1 of 3)

- Project is an amendment to the IEEE 802.3 standard
- **Title:** Amendment: Media Access Control Parameters, Physical Layers and Management Parameters for 40 Gb/s and 100 Gb/s Operation
- **Target Completion Date: May 2010**
- **Scope:** Define 802.3 Media Access Control (MAC) parameters, physical layer specifications, and management parameters for the transfer of 802.3 format frames at 40 Gb/s and 100 Gb/s.

HSSG Project Authorization Request Key Elements (2 of 3)

- **Purpose:** The purpose of this project is to extend the 802.3 protocol to operating speeds of 40 Gb/s and 100 Gb/s in order to provide a significant increase in bandwidth while maintaining maximum compatibility with the installed base of 802.3 interfaces, previous investment in research and development, and principles of network operation and management. The project is to provide for the interconnection of equipment satisfying the distance requirements of the intended applications.

HSSG Project Authorization Request Key Elements (3 of 3)

- **Need:** The project is necessary to provide a solution for applications that have been demonstrated to need bandwidth beyond the existing capabilities. These include data center, internet exchanges, high performance computing and video-on-demand delivery. Network aggregation and end-station bandwidth requirements are increasing at different rates, and is recognized by the definition of two distinct speeds to serve the appropriate applications.

Broad Market Potential (1 of 2)

- Broad sets of applications
- Multiple vendors and numerous users
- Balanced cost (LAN versus attached stations)

- **Bandwidth requirements for computing and core networking applications are growing at different rates, which necessitates the definition of two distinct data rates for the next generation of Ethernet networks in order to address these applications:**
  - **Servers, high performance computing clusters, blade servers, storage area networks and network attached storage all currently make use of 1G and 10G Ethernet, with significant growth of 10G projected in ’07 and ’08. I/O bandwidth projections for server and computing applications indicate that there will be a significant market potential for a 40 Gb/s Ethernet interface.**
  - **Core networking applications have demonstrated the need for bandwidth beyond existing capabilities and the projected bandwidth requirements for computing applications. Switching, routing, and aggregation in data centers, internet exchanges and service provider peering points, and high bandwidth applications, such as video on demand and high performance computing environments, have demonstrated the need for a 100 Gb/s Ethernet interface.**

Broad Market Potential (2 of 2)

- Broad sets of applications
- Multiple vendors and numerous users
- Balanced cost (LAN versus attached stations)

- There has been wide attendance and participation in the study group by end users, equipment manufacturers and component suppliers. It is anticipated that there will be sufficient participation to effectively complete the standardization process.
- Prior experience scaling IEEE 802.3 and contributions to the study group indicates:
  - 40 Gb/s Ethernet will provide approximately the same cost balance between the LAN and the attached stations as 10 Gb/s Ethernet.
  - The cost distribution between routers, switches, and the infrastructure remains acceptably balanced for 100 Gb/s Ethernet.
- Given the topologies of the networks and intended applications, early deployment will be driven by key aggregation & high-bandwidth interconnect points. This is unlike the higher volume end system application typical for 10/100/1000 Mb/s Ethernet, and as such, the initial volumes for 100 Gb/s Ethernet are anticipated to be more modest than the lower speeds. This does not imply a reduction in the need or value of 100 Gb/s Ethernet to address the stated applications.

Compatibility

- **IEEE 802 defines a family of standards. All standards shall be in conformance with the IEEE 802.1 Architecture, Management, and Interworking documents as follows: 802. Overview and Architecture, 802.1D, 802.1Q, and parts of 802.1f. If any variances in conformance emerge, they shall be thoroughly disclosed and reviewed with 802. Each standard in the IEEE 802 family of standards shall include a definition of managed objects that are compatible with systems management standards.**

- As an amendment to IEEE Std 802.3, the proposed project will remain in conformance with the IEEE 802 Overview and Architecture as well as the bridging standards IEEE Std 802.1D and IEEE Std 802.1Q.
- As an amendment to IEEE Std 802.3, the proposed project will follow the existing format and structure of IEEE 802.3 MIB definitions providing a protocol independent specification of managed objects (IEEE Std 802.1F).
- The proposed amendment will conform to the full-duplex operating mode of the IEEE 802.3 MAC.
- As was the case in previous IEEE 802.3 amendments, new physical layers specific to either 40 Gb/s or 100 Gb/s operation will be defined.
- By utilizing the existing IEEE 802.3 MAC protocol, this proposed amendment will maintain maximum compatibility with the installed base of Ethernet nodes.

Distinct Identity

- Substantially different from other IEEE 802 standards
- One unique solution per problem (not two solutions to a problem)
- Easy for the document reader to select the relevant specification

- The proposed amendment is an upgrade path for IEEE 802.3 users, based on the IEEE 802.3 MAC.
- The established benefits of the IEEE 802.3 MAC include:
  - Deterministic, highly efficient full-duplex operation mode
  - Well-characterized and understood operating behavior
  - Broad base of expertise in suppliers and customers
  - Straightforward bridging between networks at different data rates
- The Management Information Base (MIB) for IEEE 802.3 will be extended in a manner consistent with the IEEE 802.3 MIB for 10 / 100 / 1000 / 10000 Mb/s operation.
- The proposed amendment to the existing IEEE 802.3 standard will be formatted as a collection of new clauses, making it easy for the reader to select the relevant specification.
- Bandwidth requirements for computing and networking applications are growing at different rates. These applications have different cost / performance requirements, which necessitates two distinct data rates, 40 Gb/s and 100 Gb/s.

Technical Feasibility

- Demonstrated system feasibility
- Proven technology, reasonable testing
- Confidence in reliability

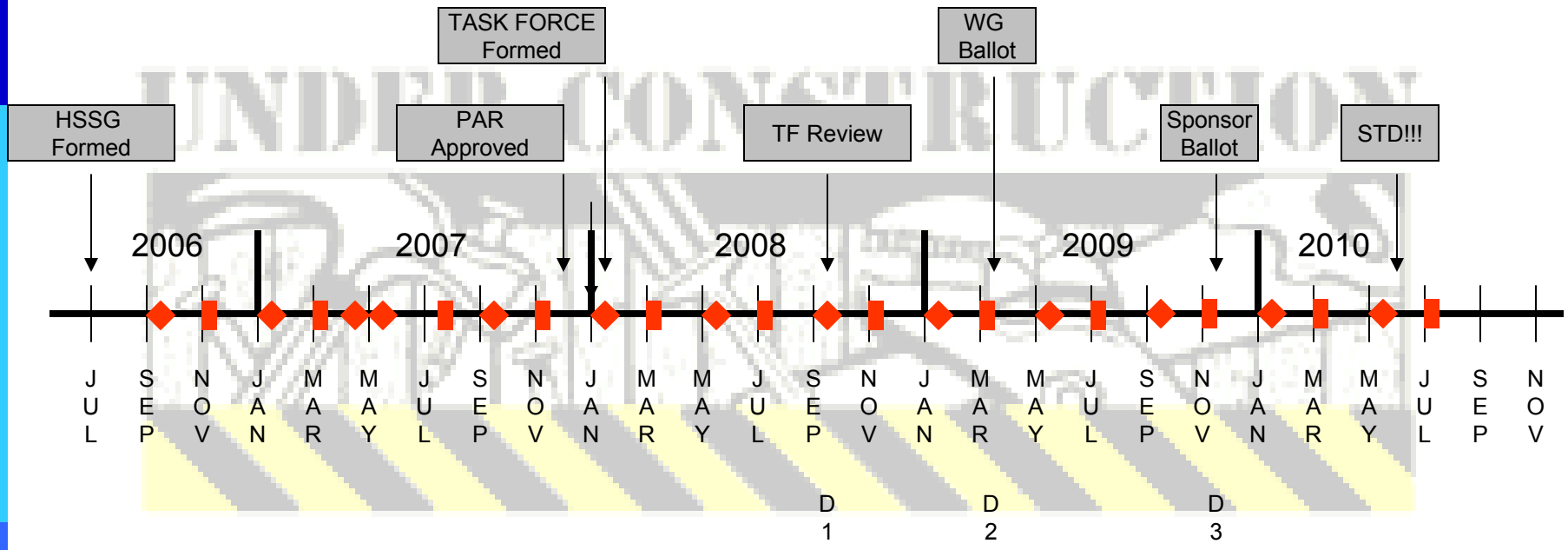
- The principle of scaling the IEEE 802.3 MAC to higher speeds has been well established by previous work within IEEE 802.3.
- The principle of building bridging equipment which performs rate adaptation between IEEE 802.3 networks operating at different speeds has been amply demonstrated by the broad set of product offerings that bridge between 10, 100, 1000, and 10000 Mb/s.
- Systems with an aggregate bandwidth of greater than or equal to 100 Gb/s have been demonstrated and deployed in operational networks.
- The proposed project will build on the array of Ethernet component and system design experience, and the broad knowledge base of Ethernet network operation.
  - The experience gained in the development and deployment of 10 Gb/s technology is applicable to the development of specifications for components at higher speeds. For example, parallel transmission techniques allow reuse of 10 Gb/s technology and testing.
  - Component vendors have presented data on the feasibility of the necessary components for higher speed solutions. Proposals, which either leverage existing technologies or employ new technologies, have been provided.
- The reliability of Ethernet components and systems can be projected in the target environments with a high degree of confidence. Presentations demonstrating this have been provided.

Economic Feasibility

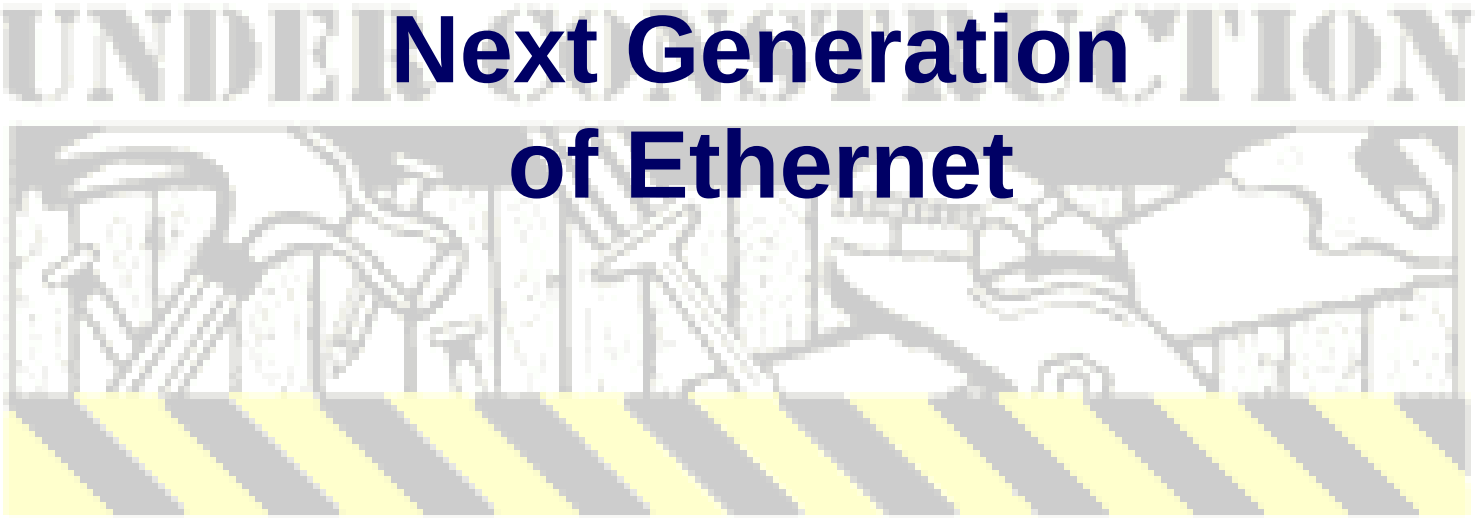
- **Known cost factors, reliable data**
- **Reasonable cost for performance**
- **Consideration of installation costs**

- The cost factors for Ethernet components and systems are well known. The proposed project may introduce new cost factors which can be quantified.
- Presentations indicate that for the server market and computing applications the optimized rate to provide the best balance of performance and cost is 40 Gb/s. For the network aggregation market and core networking applications, the optimized rate offering the best balance of performance and cost is 100 Gb/s.
- In consideration of installation costs, the project is expected to use proven and familiar media, including optical fiber, backplanes, and copper cabling technology.
- Network design, installation and maintenance costs are minimized by preserving network architecture, management, and software.

Possible Timeline



Applications for the Next Generation of Ethernet

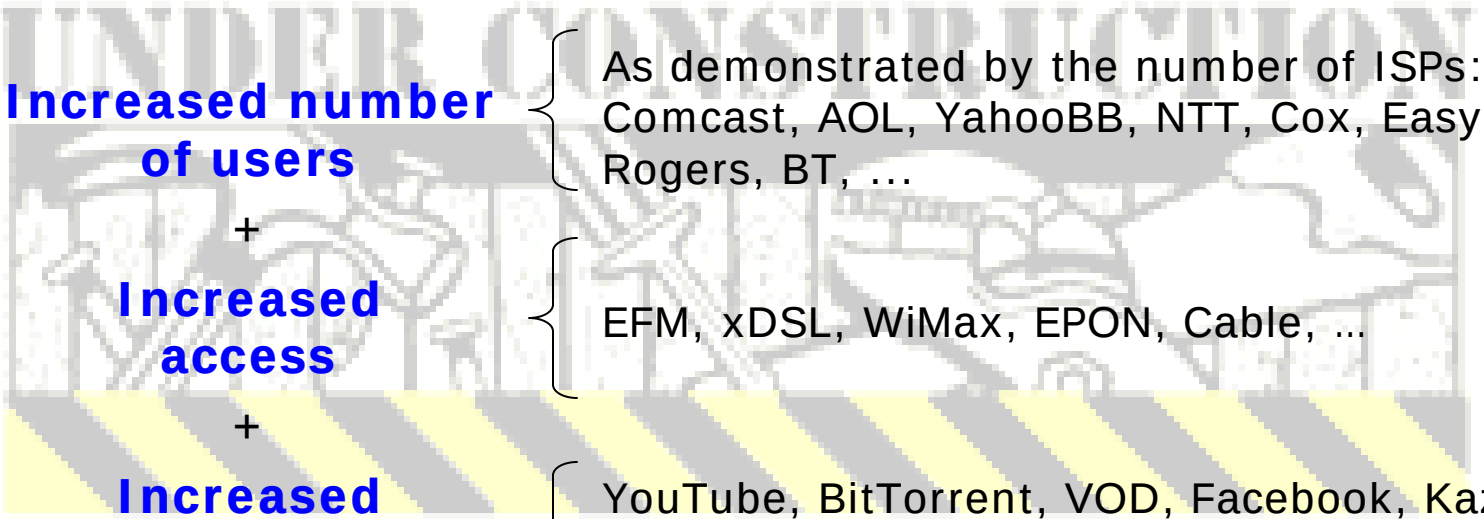


Howard Frazier, Broadcom

Mark Nowell, Cisco

Why High Speed Ethernet?

Fundamental bottleneck happening in today's networks:



Increased number of users — As demonstrated by the number of ISPs: Comcast, AOL, YahooBB, NTT, Cox, EasyNet, Rogers, BT, …

+

Increased access — EFM, xDSL, WiMax, EPON, Cable, …

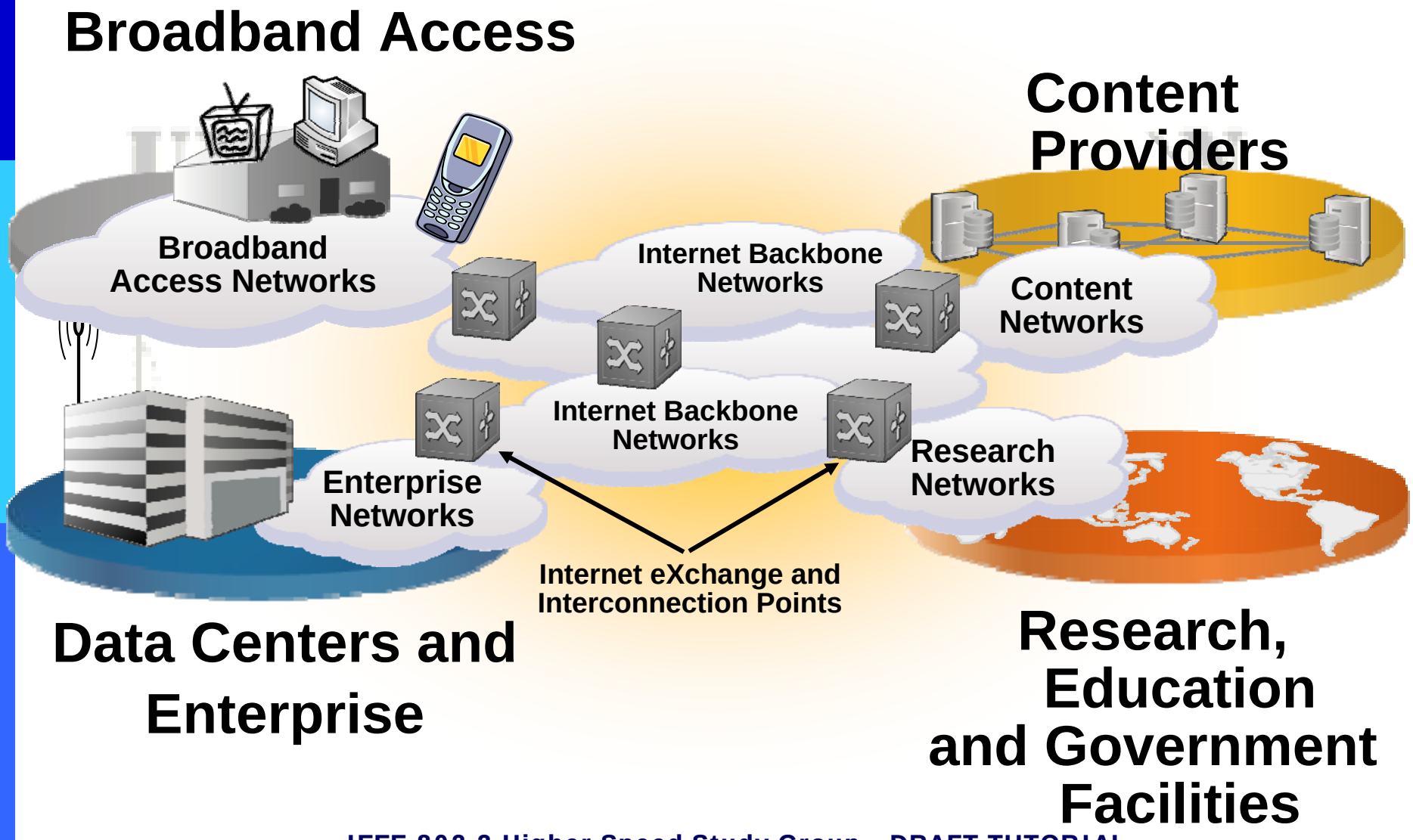
+

Increased services — YouTube, BitTorrent, VOD, Facebook, Kazaa, Netflix, iTunes, 2nd life, Gaming, gambling, …

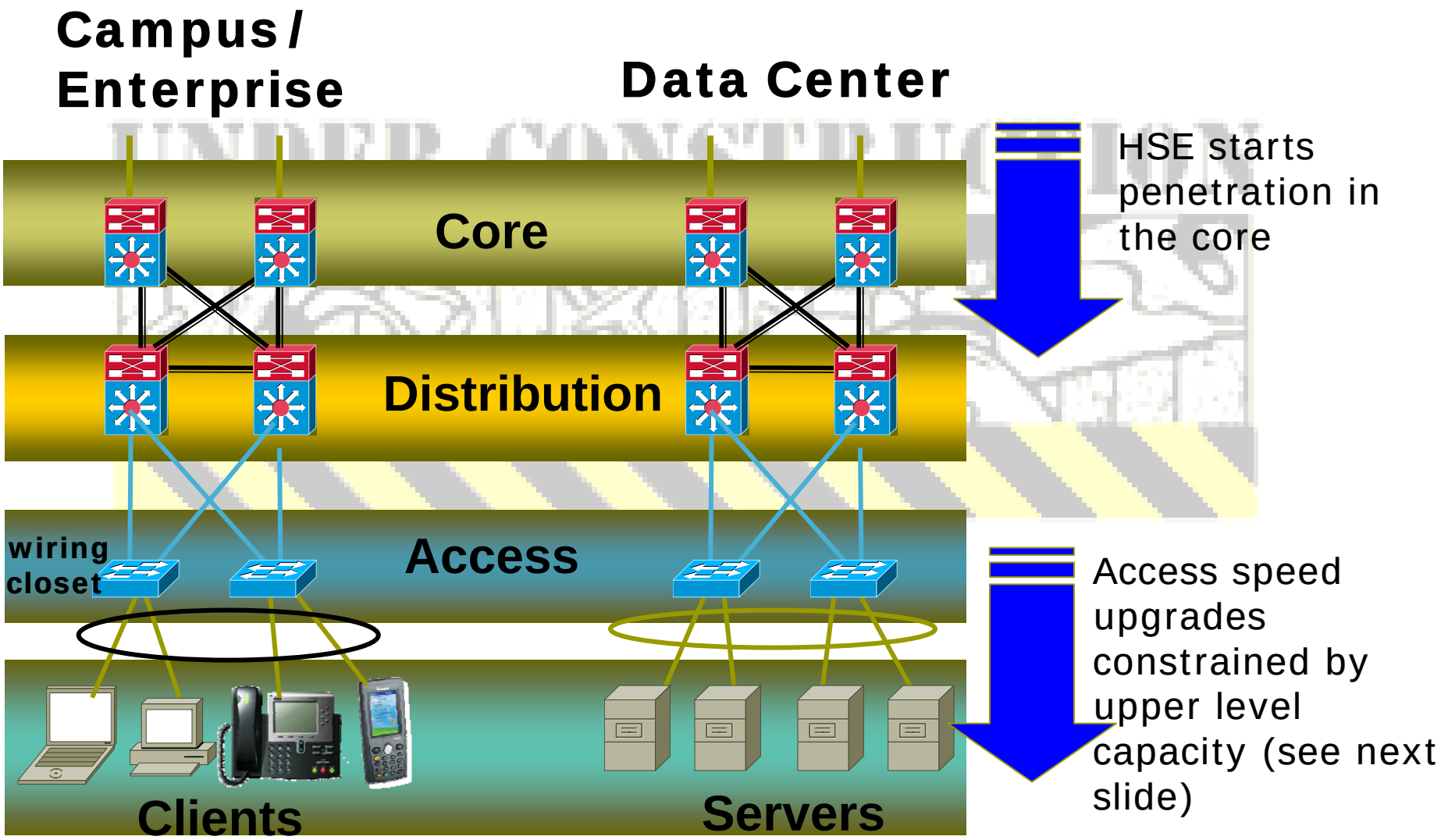
=

Bandwidth bottleneck

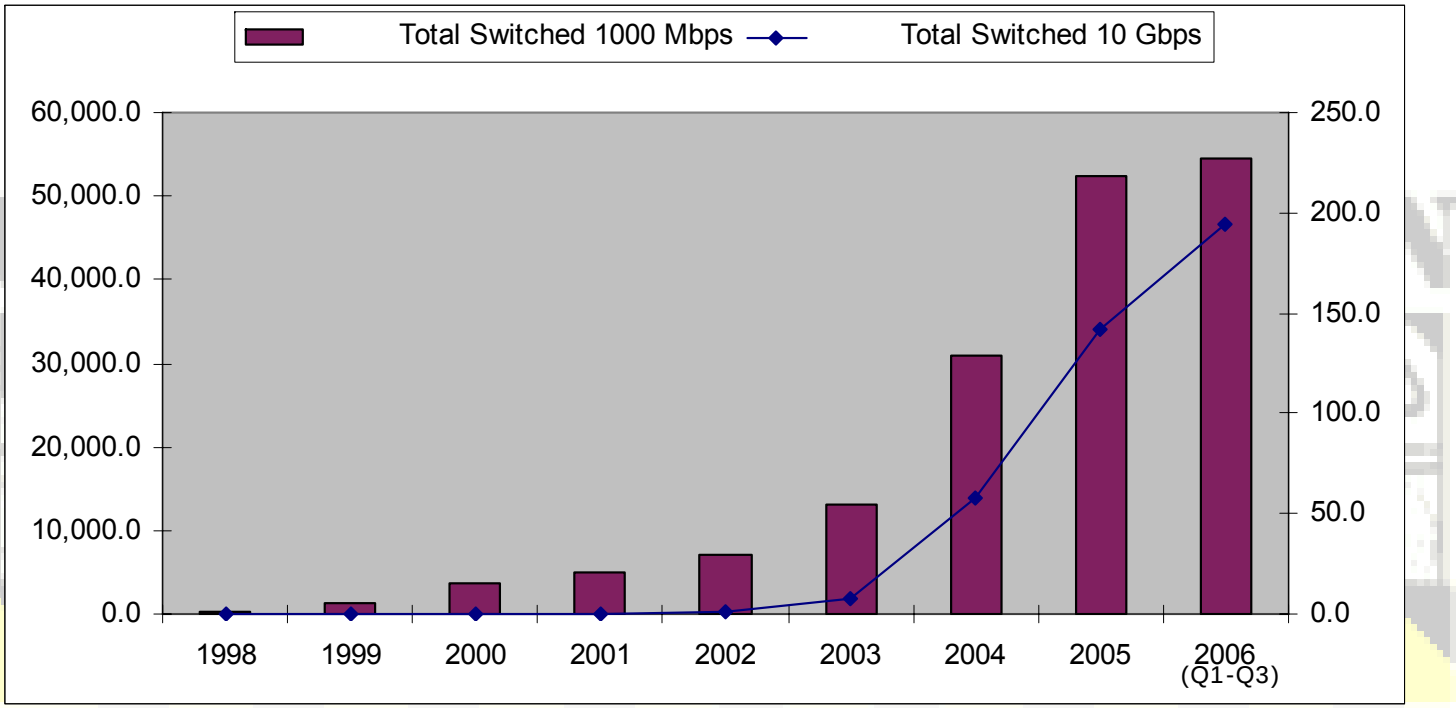
The Ethernet Ecosystem: Where is(n't) the bottleneck?



High Speed Ethernet penetration



Correlation between speeds and ports growth



Source: Cisco (barbieri_01_0107.pdf)

- **Between 2003 and 2006, GbE growth and 10GE growth were correlated. Symbiotic relationship.**

- **2007: 10GE growth being constrained by lack of higher speed interface (Source: Sprint, Yahoo, EDS, Amazon, AMS-IX, Cox, NTT, Equinox)**

High Speed Ethernet

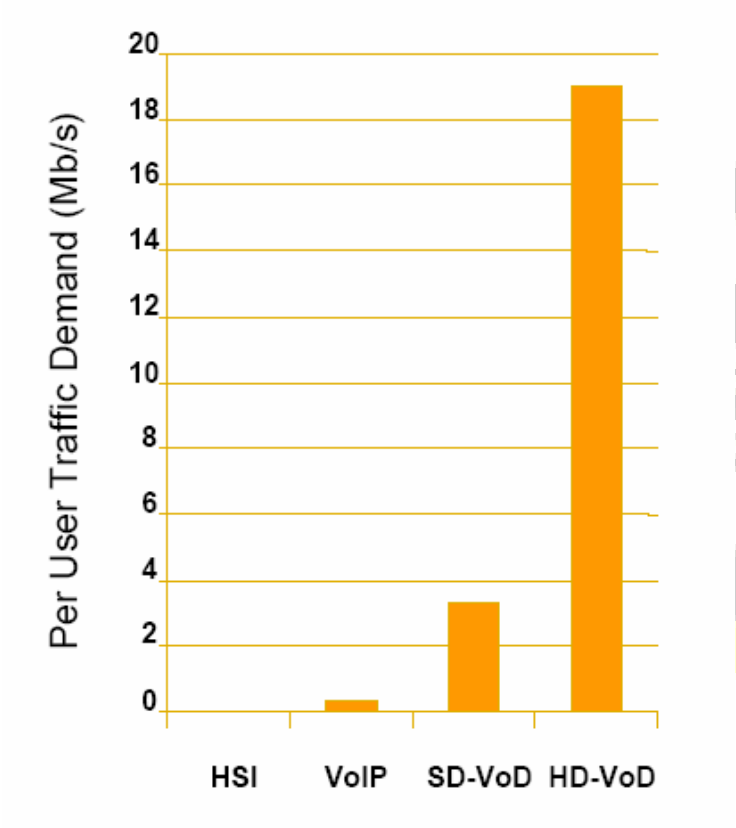
Broad Market applications

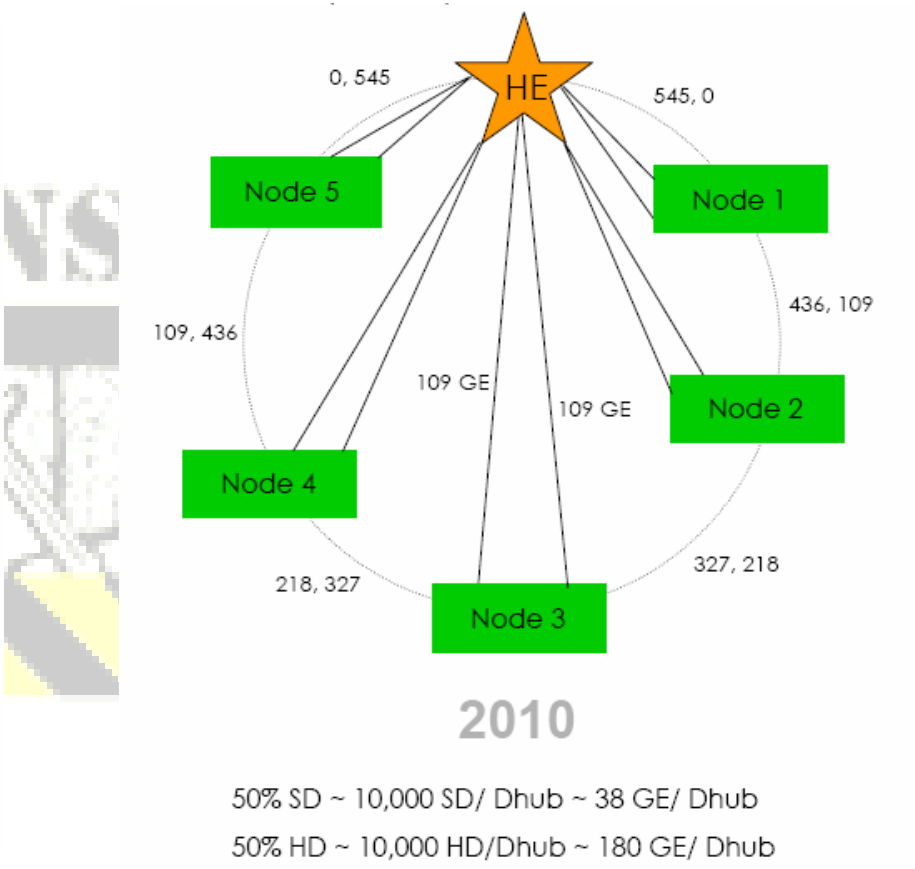
Video

VOD+IPTV is the biggest change in TV since color

- How many would stay with back to Black & White TV?
- Today's change to HDTV resolutions is minor compared to TV's change to interactivity via DVRs, iTV & VOD

Video: BW driver



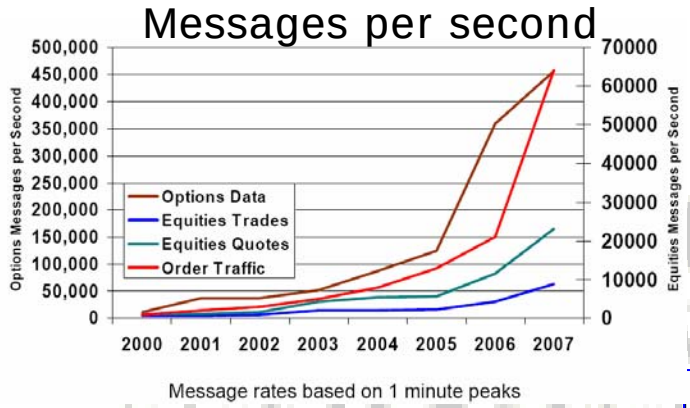


Example: 2010 regional Hub, 1M subs

Video is driving bandwidth requirements across networks segments (Backbone core, access networks, Data center)

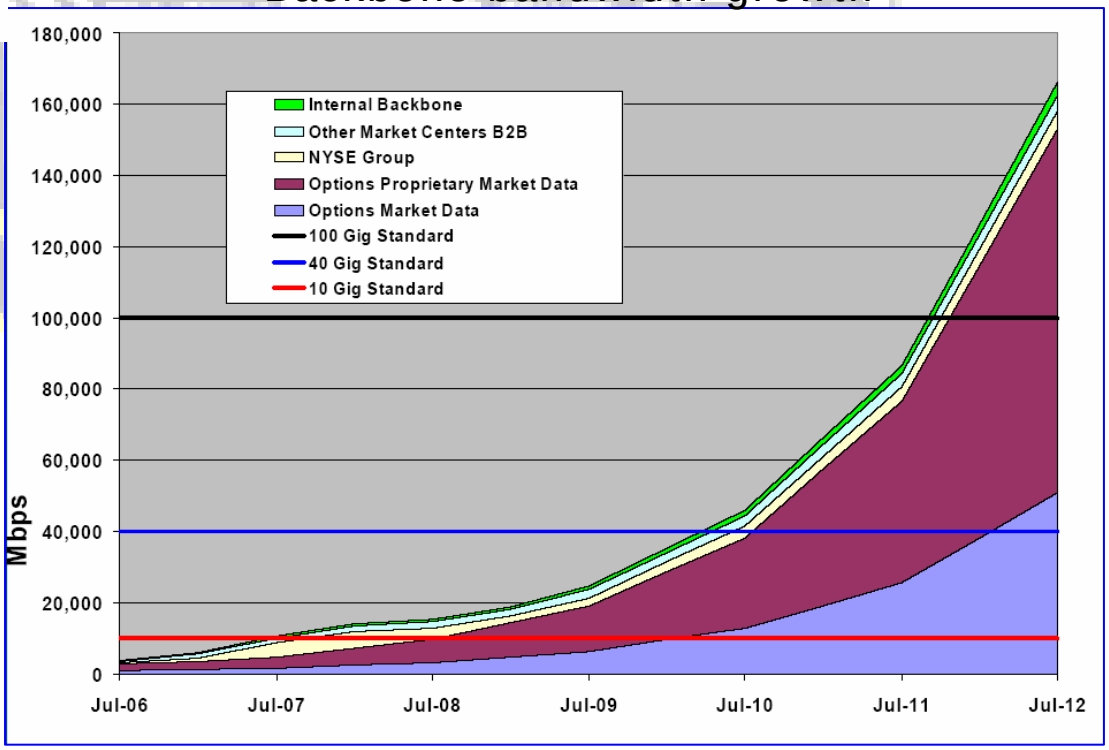
Source Comcast (Saxena_01_0107.pdf)

Financial services



Historical snapshot

Backbone bandwidth growth



Future Bandwidth growth

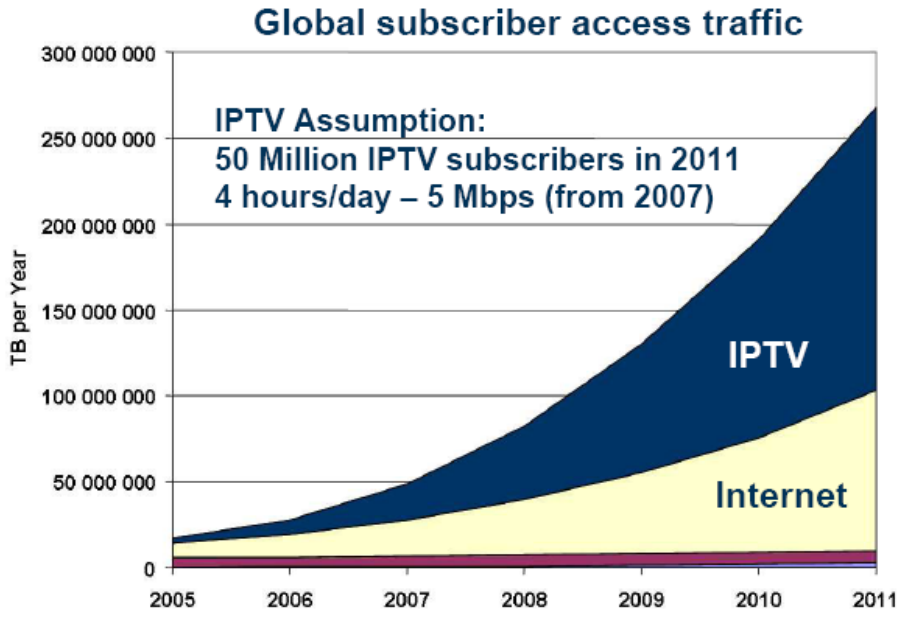
- expanded trading models
- increasing customer access choice
- increased sensitivity to latency and speed

"We need 100G now"
Andrew Bach - NYSE

Broadband Access

Broadband access

Fixed access traffic growth



Access traffic growth through increased:

- broadband penetration
- bandwidth demanding services

Service	Bandwidth
VoIP	100 kbps
IMS/Video conf	0.7 - 1.5 Mbps
Internet	0.2 - 5.0 Mbps
Gaming	0.2 - 0.5 Mbps
SDTV (MPEG-2)	6 Mbps
SDTV (MPEG-4)	3 Mbps
HDTV (MPEG-2)	20 Mbps
HDTV (MPEG-4)	10 Mbps

Typical service portfolio (2010):

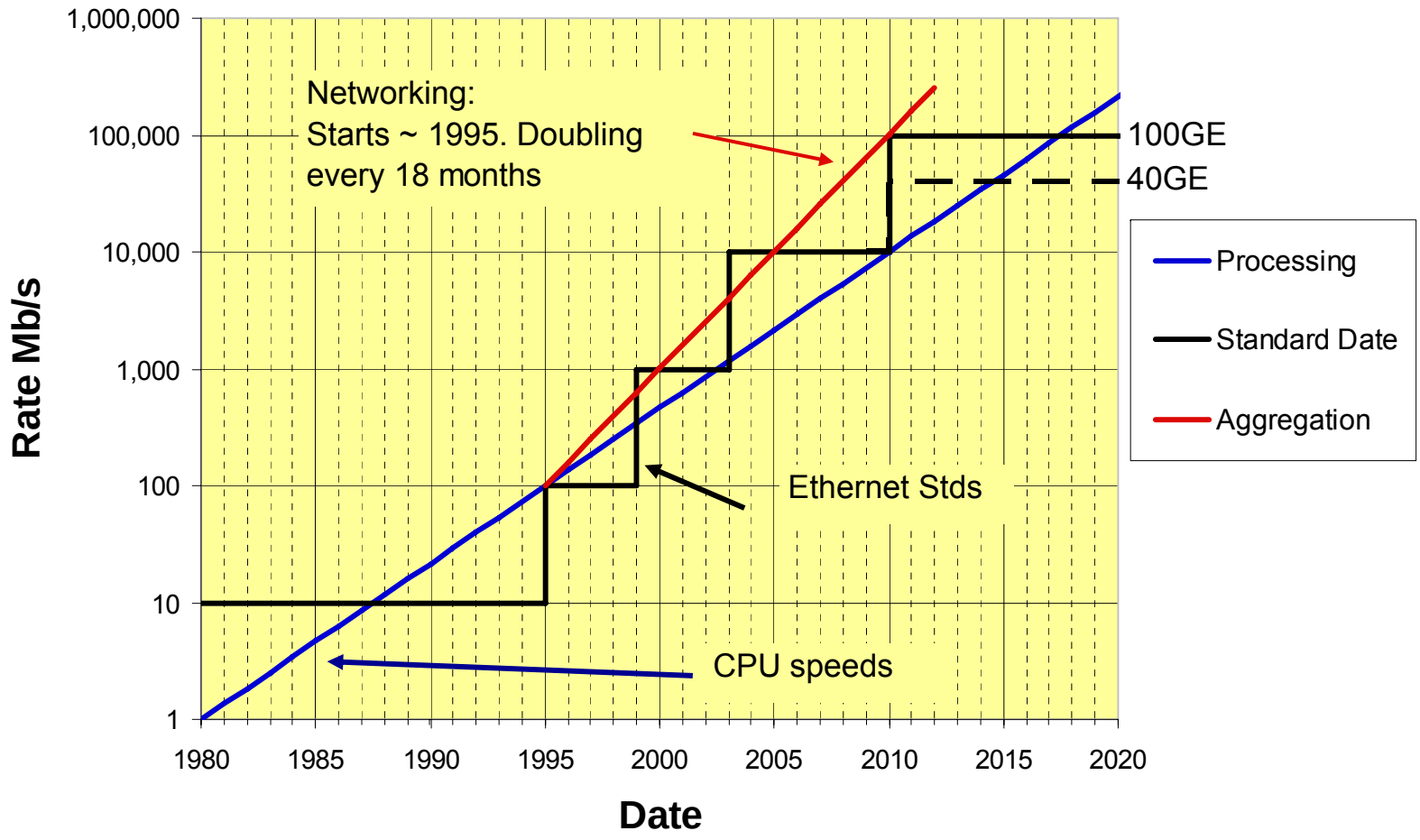
1 HDTV, 2 SDTV, gaming, voice, high-speed internet → 25 – 30 Mbps (MPEG-4)

Reference: Alping_01_1106.pdf

What is driving industry needs?

- Computing (System thoughput doubles approximately every 2 yrs)
  - Ref: www.ieee802.org/3/hssg/public/jan07/muller_01_0107.pdf
- Networking: driven by the aggregation of data from multiple computing platforms
  - Number of computing platforms growing – multiplicative effect on networking

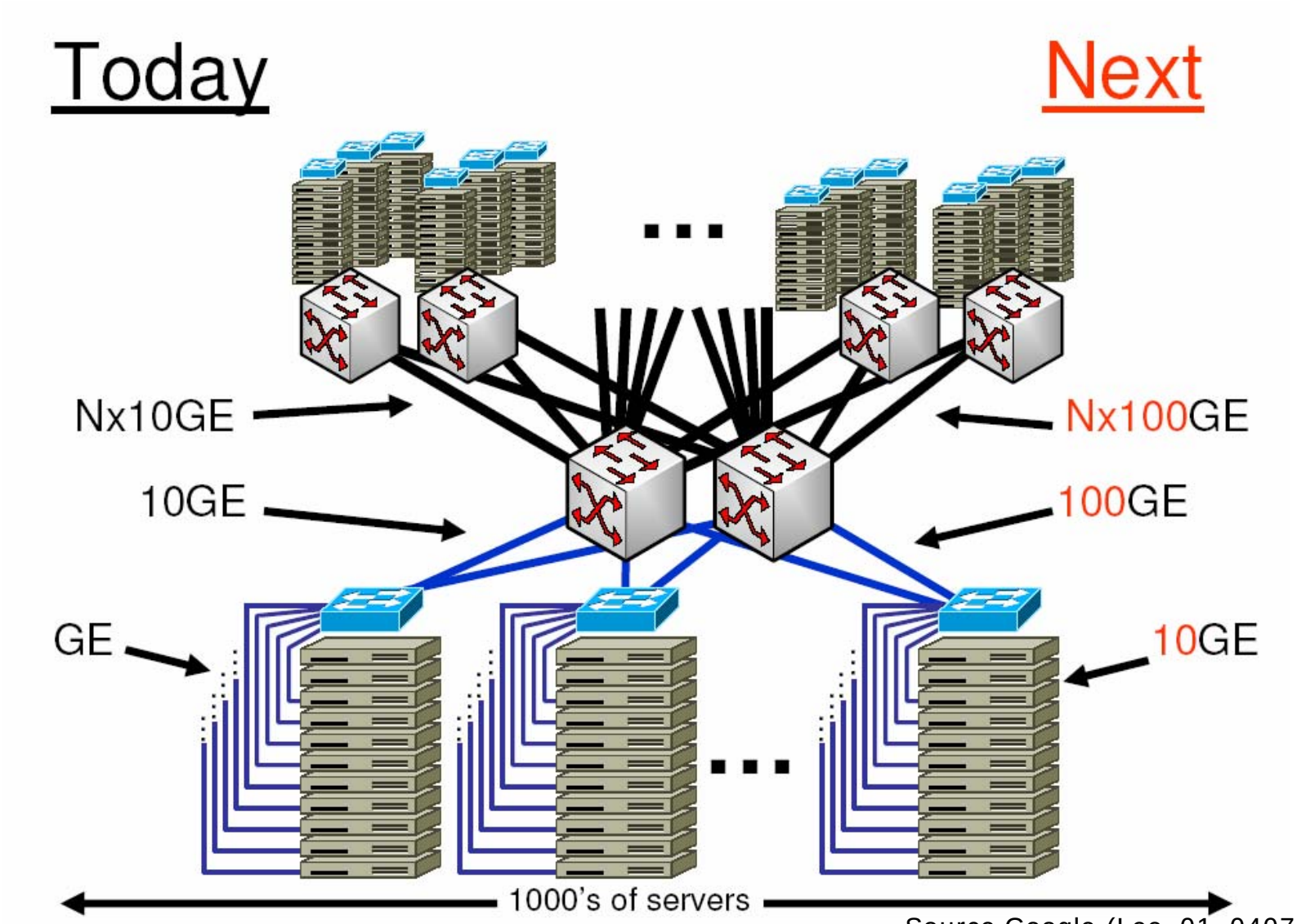
Trying to show the relationship



Observations from the graph

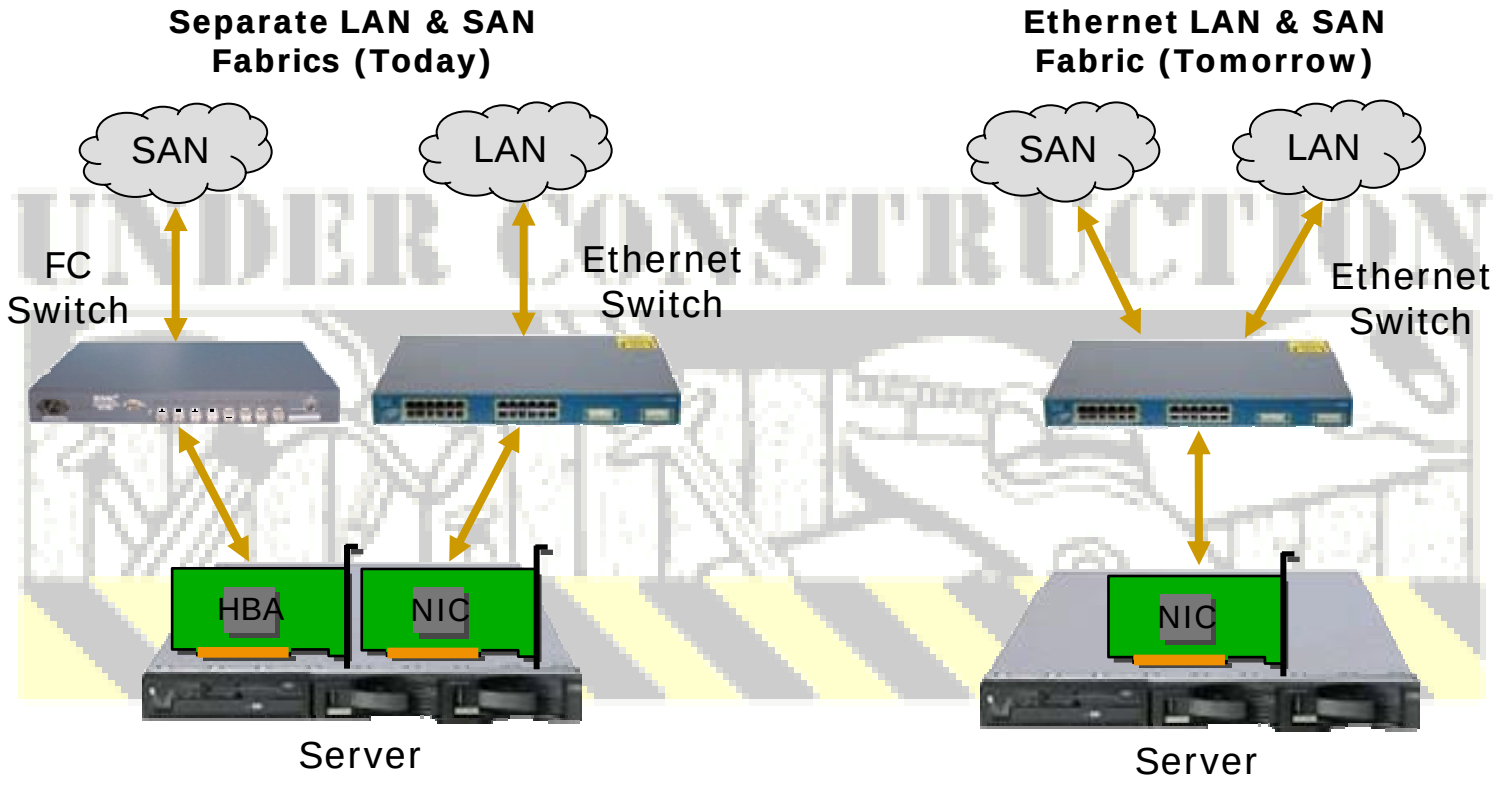
- Networking and computing interface needs are diverging
  - Makes sense. There is a multiplication factor going on here as more users, content, storage etc…
- 100M and 1GE – shows both applications needs met in similar times. Agrees with the perception of success of these standards
- For servers 10GE is just starting and sufficient for next 5+ yrs
  - Agrees with the data that says 10GE is just starting on servers (~1-2% of market).
  - Challenges the perception that 10GE was timed wrong – NOT for networking application
- For network aggregation applications, 100GE will be late
- For servers, 100GE is fairly far out. 40GE needed sooner

Massive Data Center – Server Virtualization



Source Google (Lee_01_0407.pdf)

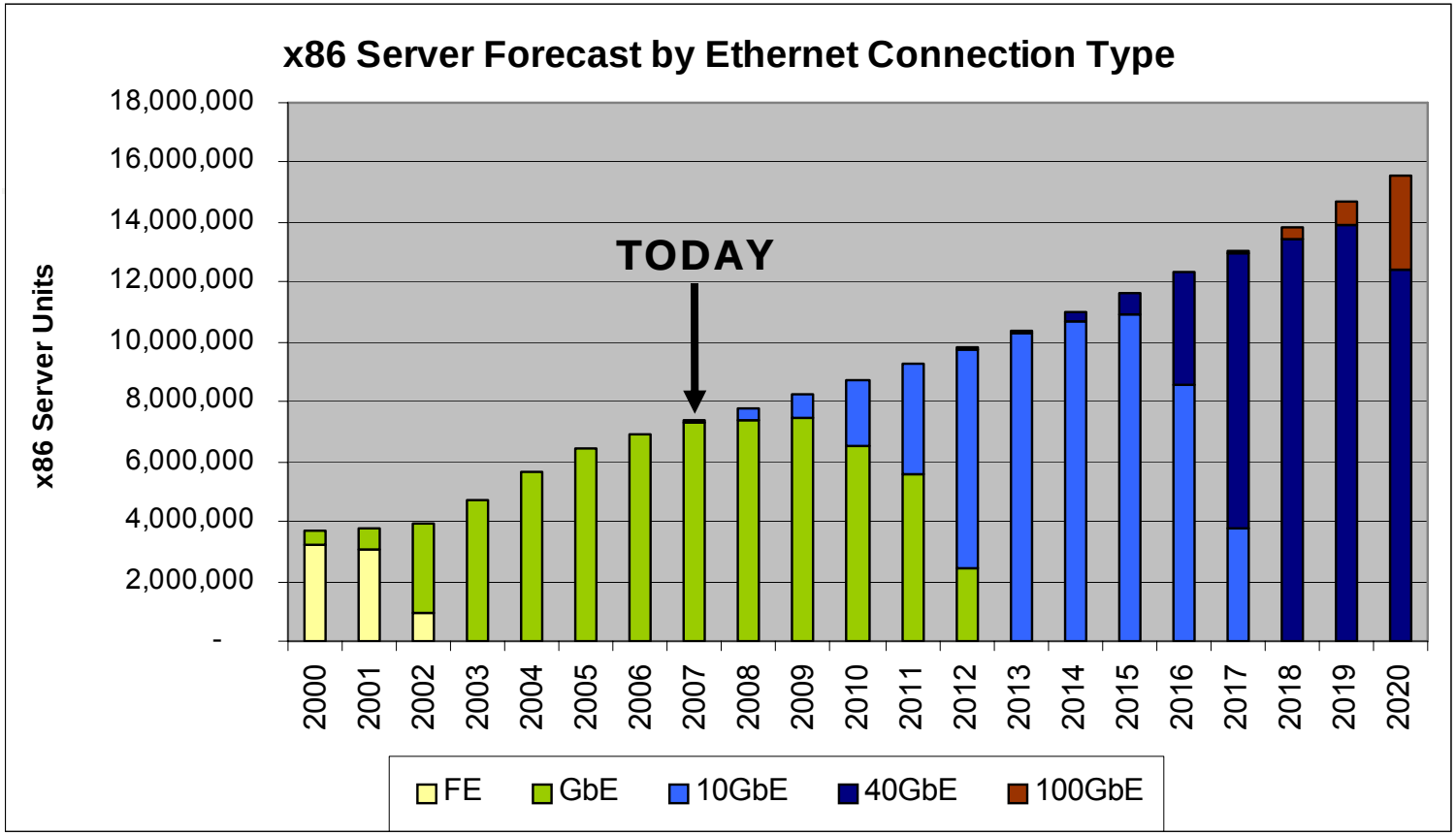
Vision: A Single, Flexible Ethernet Server I/O



- **Goal:** Reduce the cost, size, power, and complexity of servers by eliminating the need for multiple I/O connections & fabrics

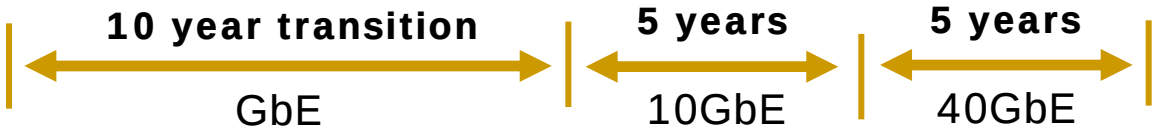
from http://www.ieee802.org/3/hssg/public/apr07/hays_01_0407.pdf

x86 Server Ethernet Connection Forecast



from http://www.ieee802.org/3/hssg/public/apr07/hays_01_0407.pdf **updated:** July 2007

Base Server Connection:



40G Market Potential for Server Networking

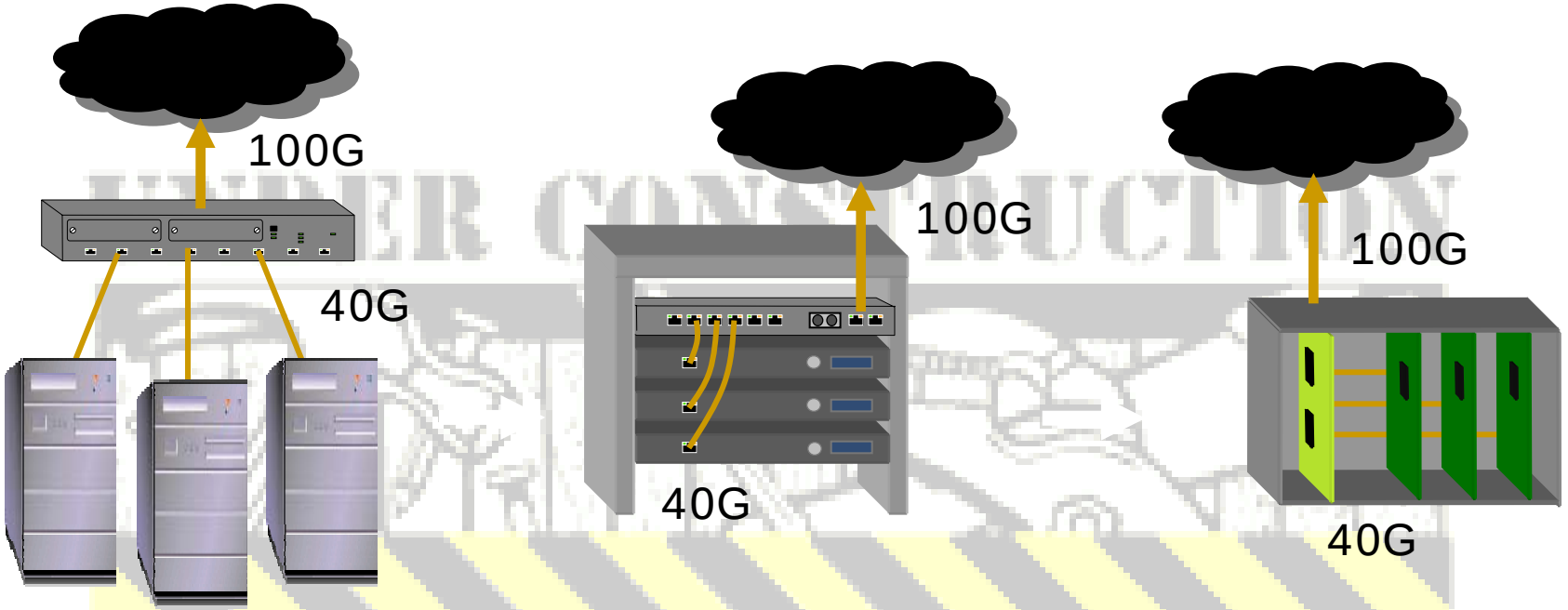
40G & 100G Data Center Ports	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
40GbE Server Connections (Kports)	17	37	98	208	660	1,399	7,407	18,303	26,859	27,857	24,843	
40GbE Switch Downlinks (Kports)	17	37	98	208	660	1,399	7,407	18,303	26,859	27,857	24,843	
40GbE TOTAL (Kports)	35	74	196	416	1,321	2,798	14,814	36,606	53,718	55,715	49,686	215,379
100GbE Switch Uplinks (Kports)	1	3	8	17	55	117	617	1,525	2,238	2,321	2,070	
100GbE Server Connections (Kports)	-	-	-	-	-	23	74	261	831	1,466	6,211	
100GbE Server Downlinks (Kports)	-	-	-	-	-	23	74	261	831	1,466	6,211	
100GbE TOTAL (Kports)	1	3	8	17	55	163	765	2,048	3,900	5,254	14,492	26,707

Source: Intel Corporation, July 2007

- **>200M port** opportunity for 40GbE in Servers & Switches
  - Assumptions: 2 ports/server, 1:1 server ports to switch downlinks
- 40G server connections drive demand for 100G switch uplinks
- Without 40GbE, higher-speed IB/FC/PCI-E may displace 10GbE before 100GbE is cost-effective, potentially reducing the long-term Ethernet TAM in the data center

from http://www.ieee802.org/3/hssg/public/apr07/hays_01_0407.pdf

Server Form Factors & Physical Networking



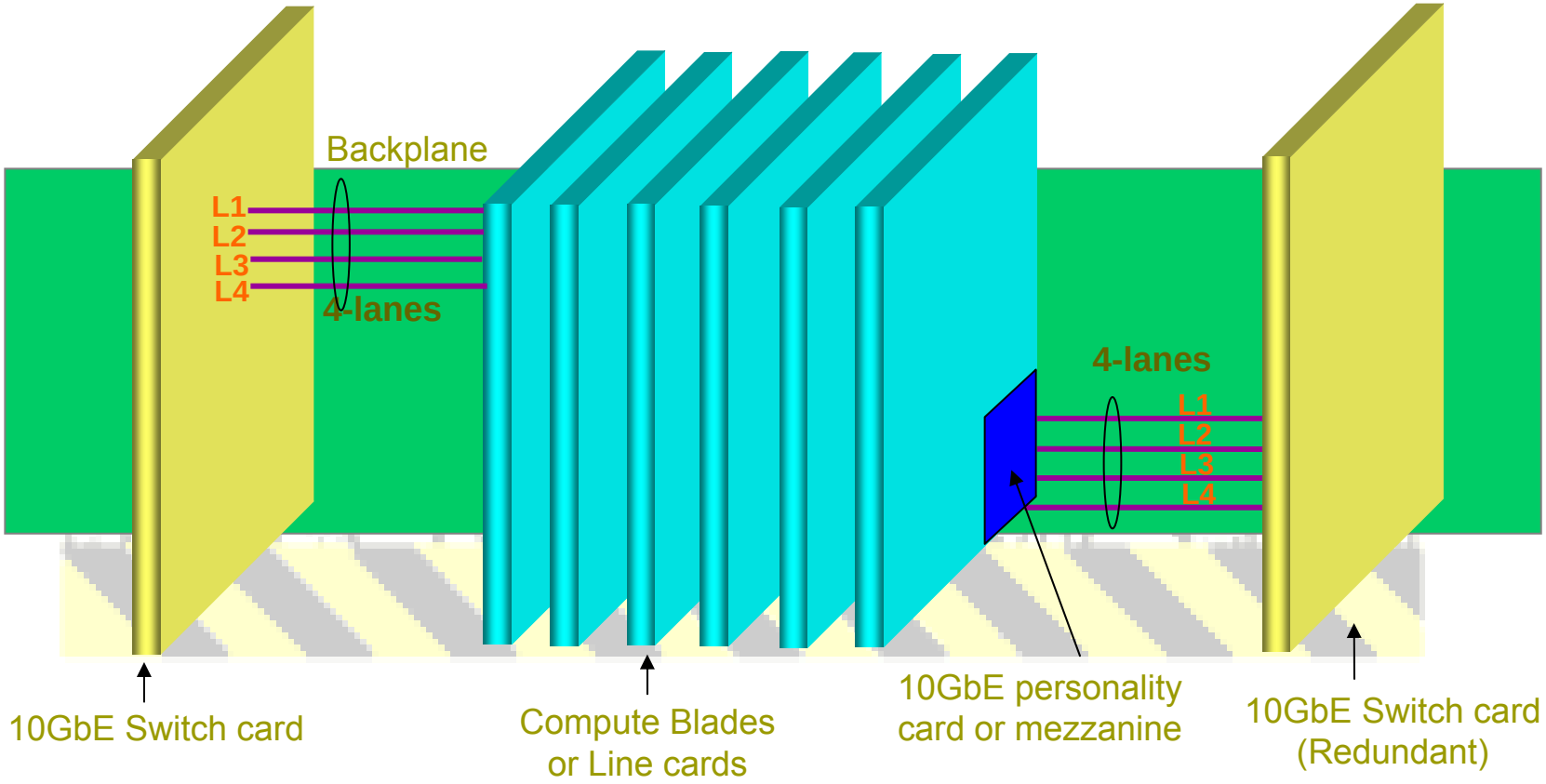
Pedestal Server:
- <100m to switch across data center
- Cu or Optical Cables
- BASE-T or -LR

Rack Server:
- <15m to switch at top-of-rack or in-row
- Copper Cables
- BASE-CX4 or -T

Blade Server:
- <1m to switch in back of chassis
- Copper PCB
- BASE-KX4 or -KR

from http://www.ieee802.org/3/hssg/public/apr07/hays_02_0407.pdf

Backplane System illustration

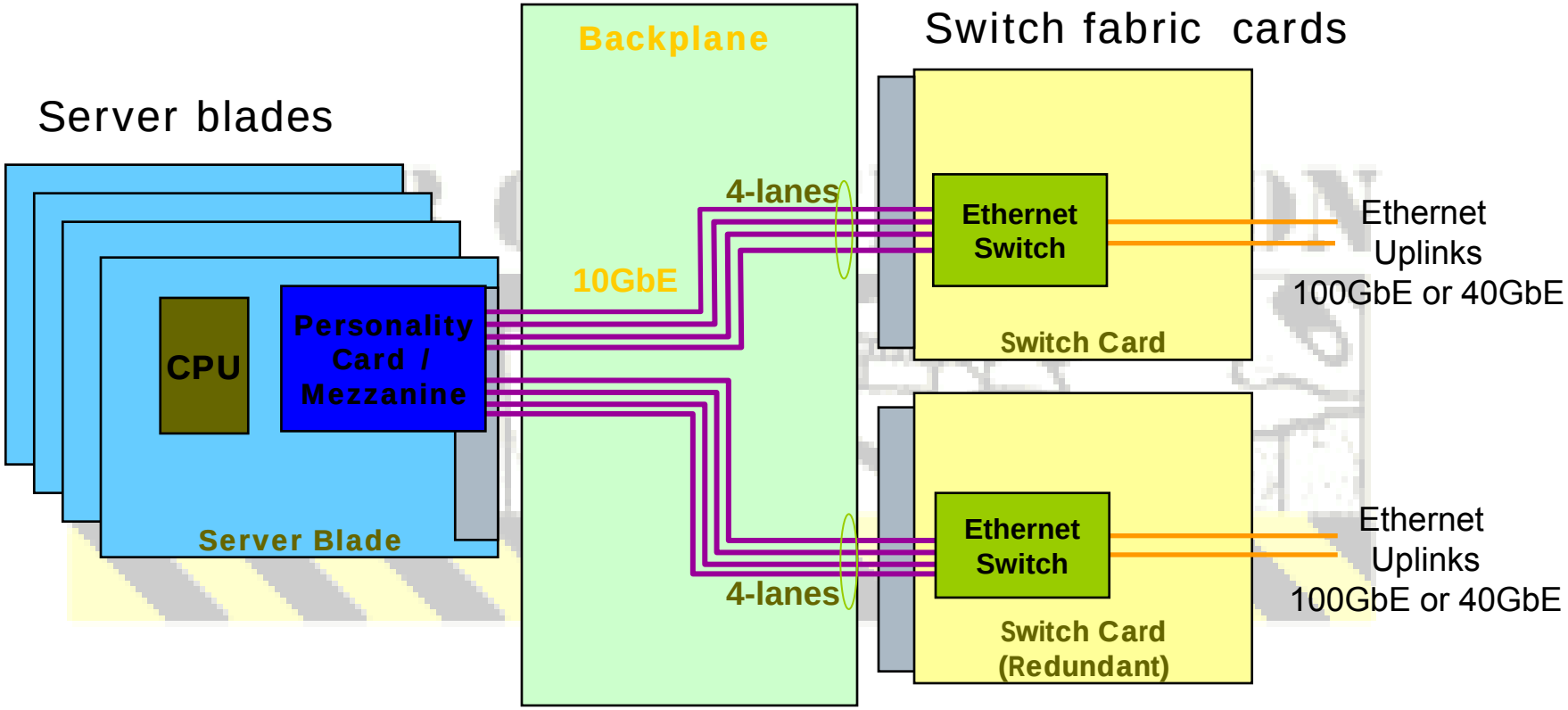


Note: The switch cards are shown at the chassis edge for simplicity.

In real systems there could be multiple fabrics located at the center, edge, or rear of the chassis

from http://www.ieee802.org/3/hssg/public/july07/ganga_01_0707.pdf

Backplane System schematic



- Backplane configuration
  - 4–lane Dual Star to support 10GBASE-KX4
  - Can support 10GBASE-KR per lane
  - Allows 40GbE MAC rate leveraging the KR PMD to support “40GBASE-KR4”

from http://www.ieee802.org/3/hssg/public/july07/ganga_01_0707.pdf

Server Aggregation

- Consider rack mount servers
  - Gigabit Ethernet I/O
    - Horizontal server architectures can justify 100G links for aggregation
    - 7 years after 1000 BASE-T
  - 10 Gigabit Ethernet I/O
    - Transition over the next 3 to 5 years
    - 10GBASE-T completion, implementation, deployments
    - Offload, bus, and virtualization technologies will mature to fully use bandwidth.
- Consider blade servers & backplanes
  - Per IDC ≈ 2.8M blade servers 2008 – 2009, estimate 500k chassis
  - Rule of Thumb
    - Create uplinks to support ½ the total aggregate bandwidth
  - 1000BASE-KX
    - ½ * 14 slots * GbE = 7G Uplink
    - ½ * 14 slots * 4 GbE = 28G Uplink
  - XAUI / 10GBASE-KX4 / 1 Port 10GBASE-KR-
    - ½ * 14 * 10GbE = 70G Uplink
    - **500K Chassis * 2 Ports = 1M 100G Ports Opportunity**
  - 4 Ports 10GBASE-KR
    - ½ * 14* 4 * 10 GbE = 280G Uplink
- **Existing network infrastructure is modeled on x10 aggregation scheme.**
- **Importance of cost for deployment**
  - **Today's cost**
  - **Long term ROI**

Based on conversations with and supported by Robert Winter, Dell.

Market Drivers for more bandwidth