Higher Speed Study Group
Call-For-Interest

IEEE 802.3 Working Group

San Diego, CA

July 18, 2006

Objective for this Meeting

- To ***measure the interest*** in starting a study group for ***Higher Speed Ethernet***

- We don't need to
  - Fully explore the problem
  - Debate strengths and weaknesses of solutions
  - Choose any one solution
  - Create PAR or five criteria
  - Create a standard or specification

- Anyone in the room may speak / vote
- RESPECT… give it, get it

Agenda

- **Presentations**
  - **“Higher Speed Ethernet - Overview,” Brad Booth.**
  - **“The Need for Higher Speed Ethernet,” Jeff Cain.**
  - **“The Technical Viability of Higher Speed Ethernet,” Petar Pepeljugoski.**
  - **“Higher Speed Ethernet - Why Now,” John D’Ambrosia.**
- **Discussion**
- **Call for Interest**
- **Future Work**
- **Cabo Wabo Time!**

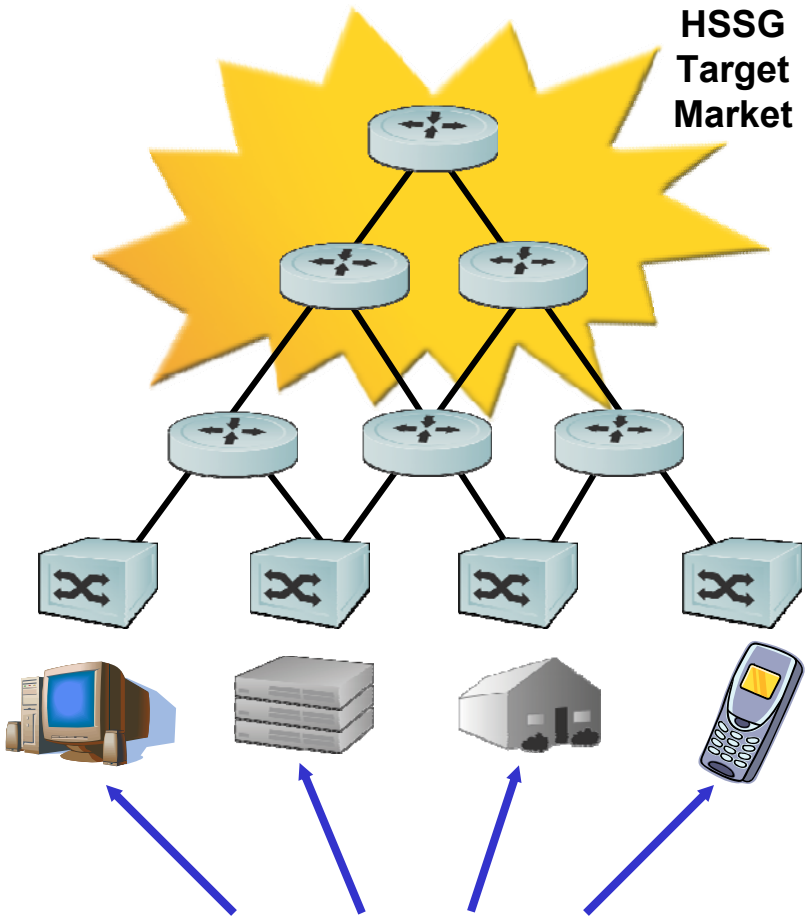
Higher Speed Ethernet - Overview

Presented by
Brad Booth, Quake Technologies

IEEE 802.3 Working Group
San Diego, CA
July 18, 2006

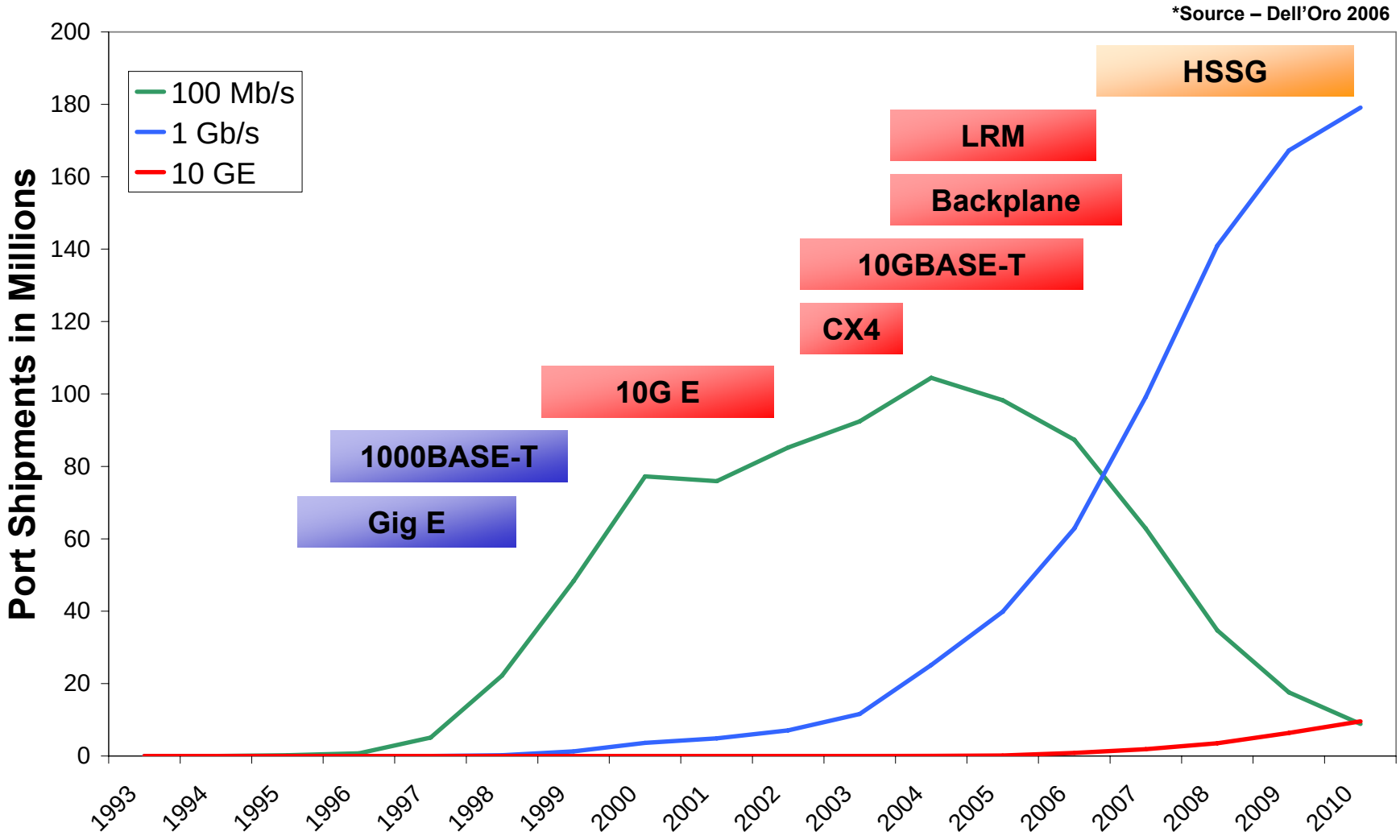
Value = Growth

- Metcalfe Law
  - Value of network proportional to square the number of users
- Permitted Ethernet to enter new markets
  - Metro, core, carrier
  - Access, consumer
  - Industrial, video, medical, aircraft
- Top of the network requirements
  - Bandwidth is a commodity
  - Resiliency
  - Reduced operational costs (OpEx)

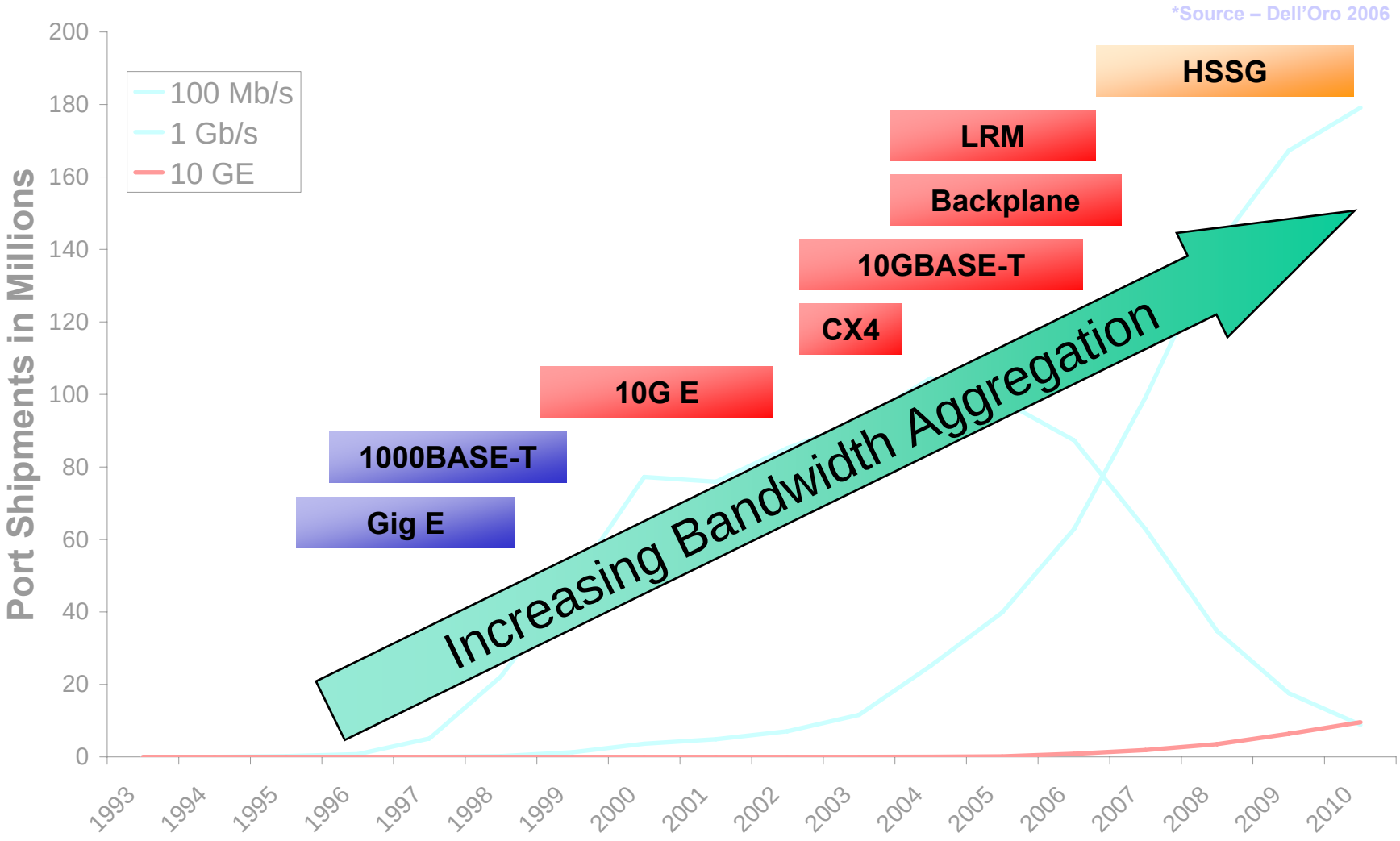


Increasing Performance & Bandwidth Demands

Managed Ethernet Switches



Managed Ethernet Switches

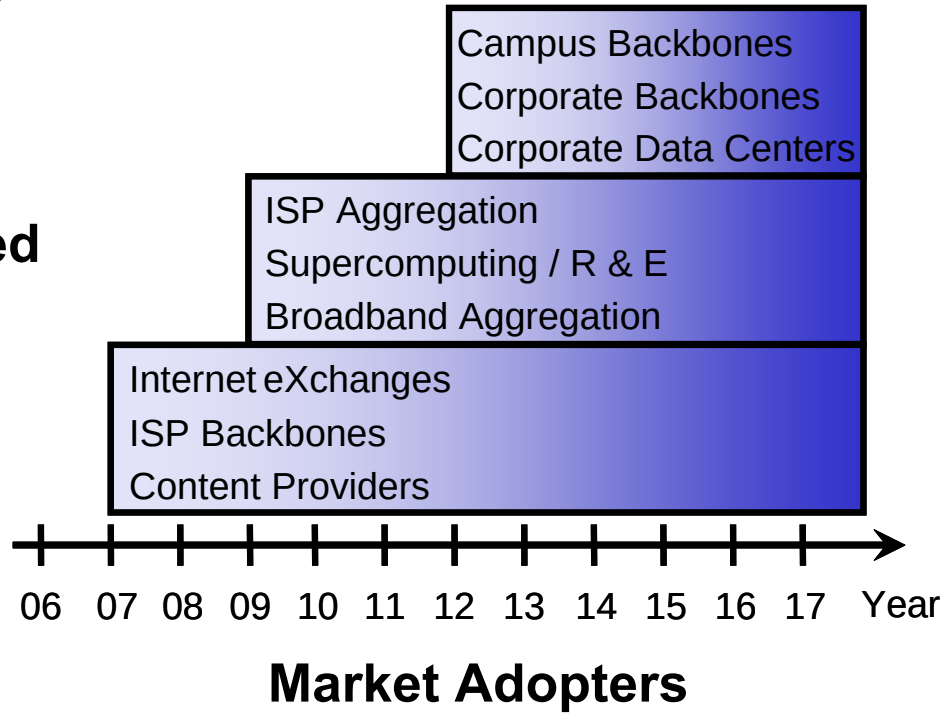


802.3ad Link Aggregation (LAG)

- Temporary fix for increased bandwidth demand
- Increased complexity
  - Difficult to plan for capacity and traffic engineering
  - Harder to manage & troubleshoot multiple physical links based on a single logical interface
  - Cable & link management
- Uneven distribution of traffic
  - Limitations in the standard
  - Inefficient distribution of large flows
  - Load balancing requires packet inspection or other knowledge

Market Drivers

- **Bird's eye view**
  - **Not targeted at the desktop**
  - **Multiple applications throughout the network**
  - **Cost model at aggregation points should be considered**
- **Possible reach targets**
  - **Backplane (1m)**
  - **Very short reach (< 25m)**
  - **Short reach (< 100m)**
  - **Long reach (< 40km)**
  - **Very long reach (> 40km)**
- **Implication…**
  - **Possibility of > 1 project**
  - **Study group may find multiple projects to initiate**



“The Need for Higher Speed Ethernet”

Presented by
Jeff Cain, Cisco Systems

IEEE 802.3 Working Group
San Diego, CA
July 18, 2006

Driving Bandwidth Requirements

- **Consumers**
  - **Increasing penetration**
  - **Increasing bandwidth**
- **Content**
  - **Increasing bandwidth requirements**
- **Networking – Pulling it all together**
  - **Service Providers**
  - **Internet eXchanges**
- **Other**
  - **High Performance Computing**
  - **Data Centers**
  - **Research & Development**

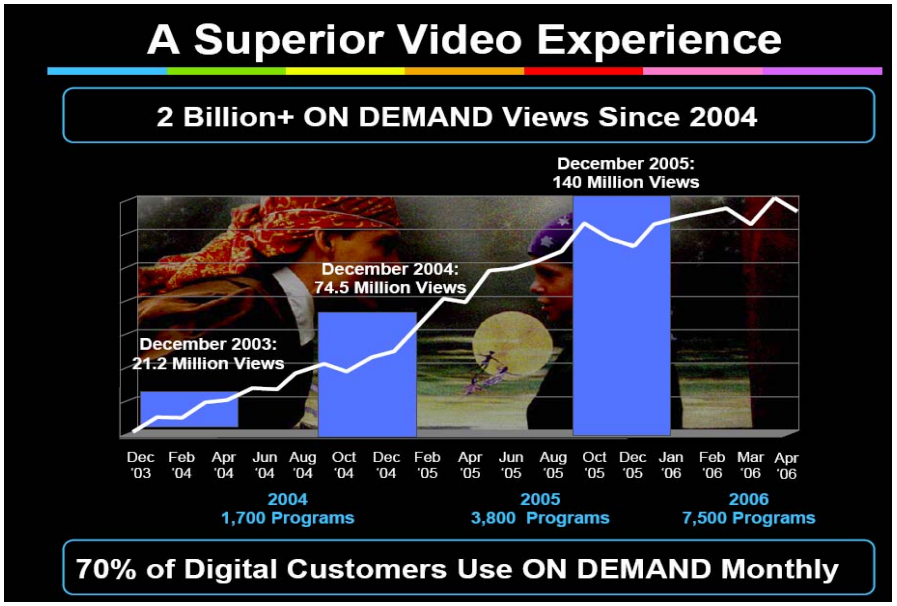
Consumers and Broadband Access

- **Applications need 50+ Mbps per home**
  - **File Sharing / Peer to Peer**
  - **IPTV / Video on Demand / High Definition TV**
  - **Gaming**
- **Broadband access**
  - **50 / 100 Mbps typical in Japan and Korea**
  - **US / Europe - 1 to 30 Mbps**
    - **Verizon fiber deployment for FIOS**
    - **Comcast to trial 100 Mbps service in 2006 (DOCSIS 3.0)**
    - **Centillium Communications announces VDSL2 OFFERING 100 Mbps symmetric data rates**
    - **Utah's Utopia in June's Spectrum**
      - **100M point-to-point network to a potential 2.5M customers**
      - **"easily" upgradeable to 1G**
    - **IEEE 10GEPON Study Group**

Content Providers

- **Personalized content is a killer application**
- **Google, Yahoo!, MSN and YouTube deliver high bandwidth personalized content**
  - **Yahoo!** (Source: Adam Bechtel, Yahoo!, "The Need for 100GE", Designcon 2006 Management Forum Panel)
    - **10 GbE used in data center, metro and WAN**
    - **6 x 10 GbE and 4 x 10 GbE LAG required Q106**
    - **Ethernet everywhere**
  - **Yahoo! Asia Pacific** (source: Chris Choi, Yahoo! Asia Pacific)
    - **40+ Gbps aggregate bandwidth need to support Major League Baseball streaming in Asia Pacific**
  - **YouTube** (Source: Colin Corbett, YouTube, "Peering of Video", NANOG 37)
    - **Personal video sharing site with over 20 Gbps traffic (50 million videos viewed per day)**
    - **~20% monthly traffic growth, all unicast traffic**
- **Online DVD rental initiatives by NetFlix and Amazon**

Video on Demand: Driving Bandwidth



Used with permission from Comcast

- Comcast – customer profile
  - 10M Digital subscribers - 47% of customer base
  - 70% Digital customers use Video on Demand
  - Only 28% are getting High Definition programming

Source: Ian Cox, Cisco Systems / Comcast

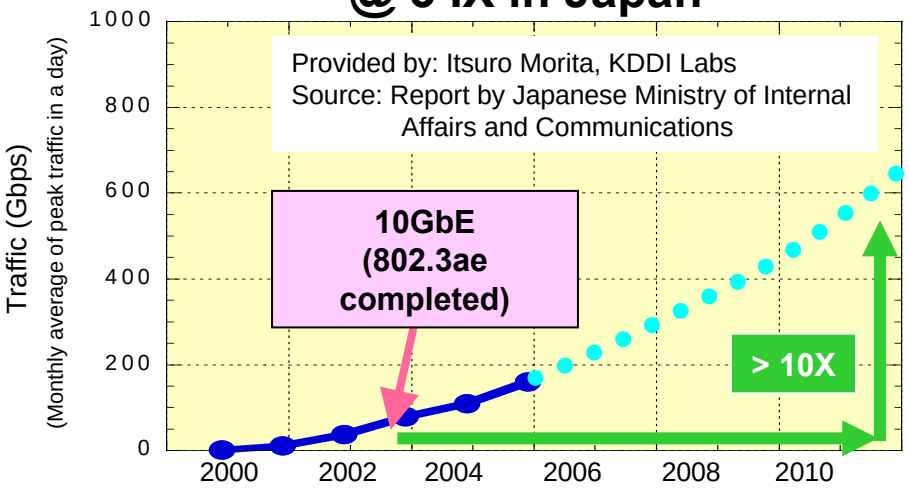
- Network profile
  - Driving bandwidth factor is Video on Demand unicast traffic
    - Standard Definition (SD): 3.5Mbps
    - High Definition (HD): 19 Mbps
  - Regional networks service a metro area
  - One 10 GbE link can only deliver 2500 SD streams
  - Sections of the regional network use up to 7 x 10 GbE links today
    - 1 to 2 bi-directional links for Voice, Broadcast TV, Internet
    - 2 to 5 uni-directional links Video on Demand
- Personalized Content – killer app!
  - Growth potential for digital subscribers
  - Growth potential for HD programming
  - Bandwidth requirements of HD

Service Providers

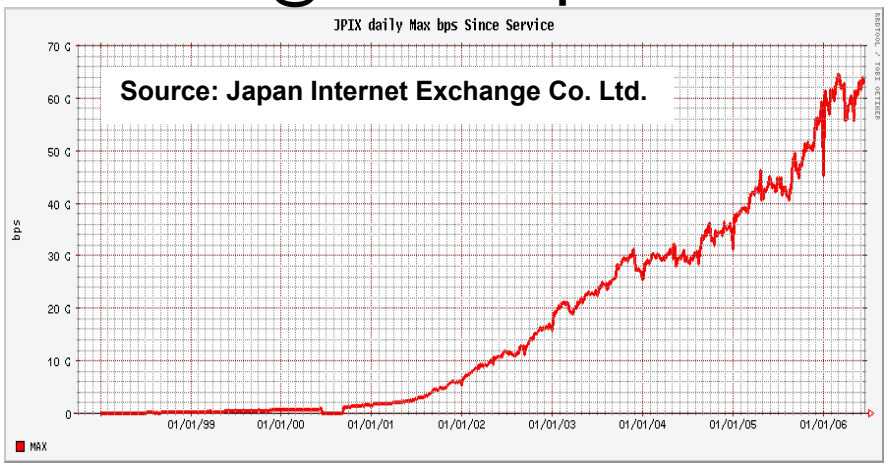
- **High growth projected as broadband penetration and speed increase**
  - **Core links and aggregation to core links: n x 10 GbE LAG**
  - **Customer links: 10 GbE and n x 10 GbE LAG**
- **Current interfaces and bundling technologies will be challenged to meet future demand in largest networks**
- **Level 3** (Source: Joe Lawrence, Level 3 Communications)
  - **Today (4Q05)**
    - **Edge router connectivity ranges from N x GbE to N x 10 GbE redundant**
    - **Backbone router connectivity can range up to 8 x 10 GbE**
    - **Most majors routes operate at 2-4 x 10 GbE**
    - **One major corridor operates at 8 x 10 GbE**
  - **Future (2010)**
    - **Backbone to LAN aggregation demands grow to >100 x 10 GbE**
    - **Most major routes projected to require 20-60 x 10 GbE**
    - **One major corridor projected to require over 100 x 10 GbE**

Traffic Patterns at Major IXs

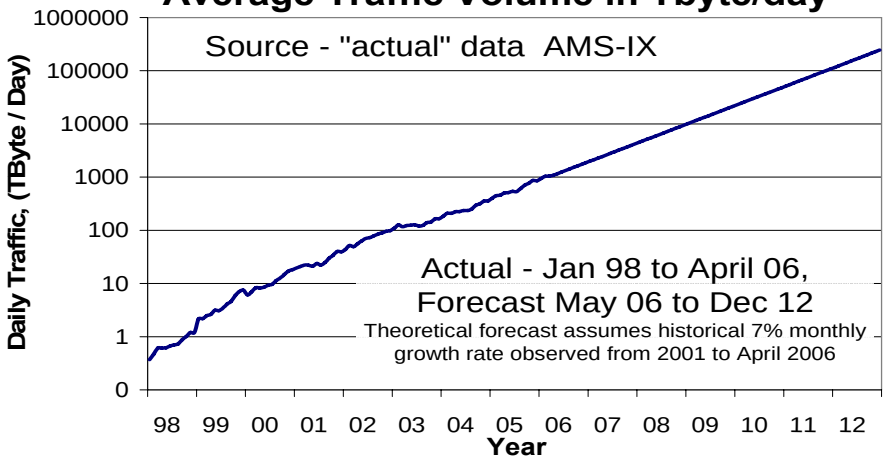
Monthly Average Peak Traffic @ 3 IX in Japan



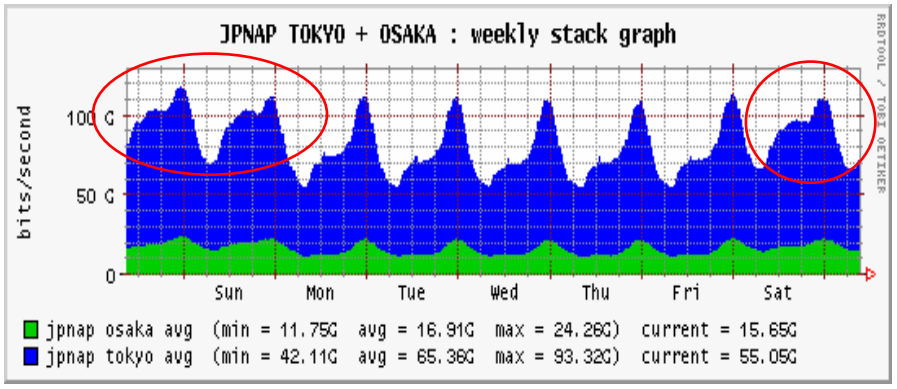
Daily Maximum Gbps @ JPIX in Japan



AMS-IX in Amsterdam Average Traffic Volume in Tbyte/day



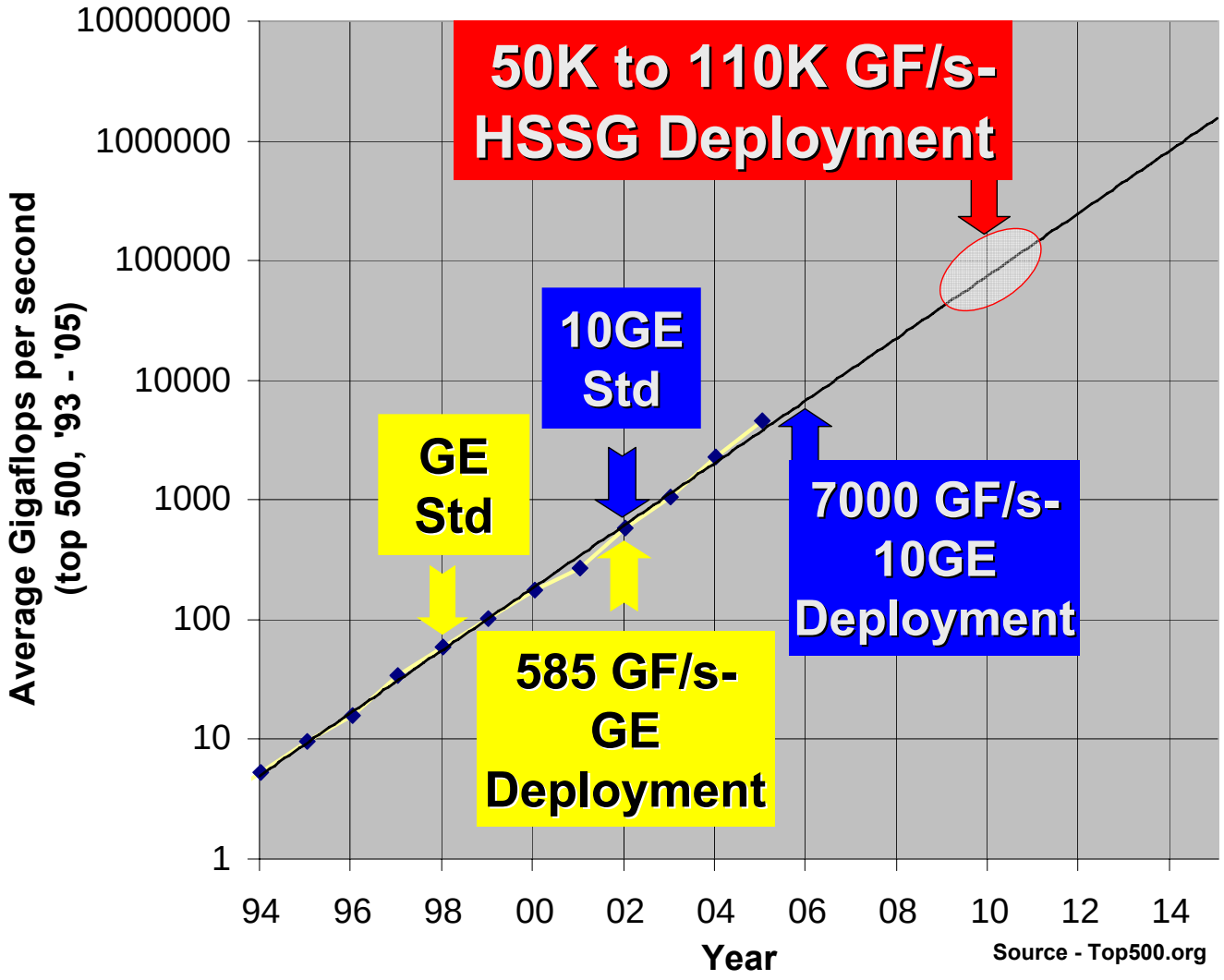
Daily Traffic @ JPNAP in Japan



Average traffic growth rate since 2003 6% per month

Used with permission from Internet Multifeed Co.

High Performance Computing



Historical Observations

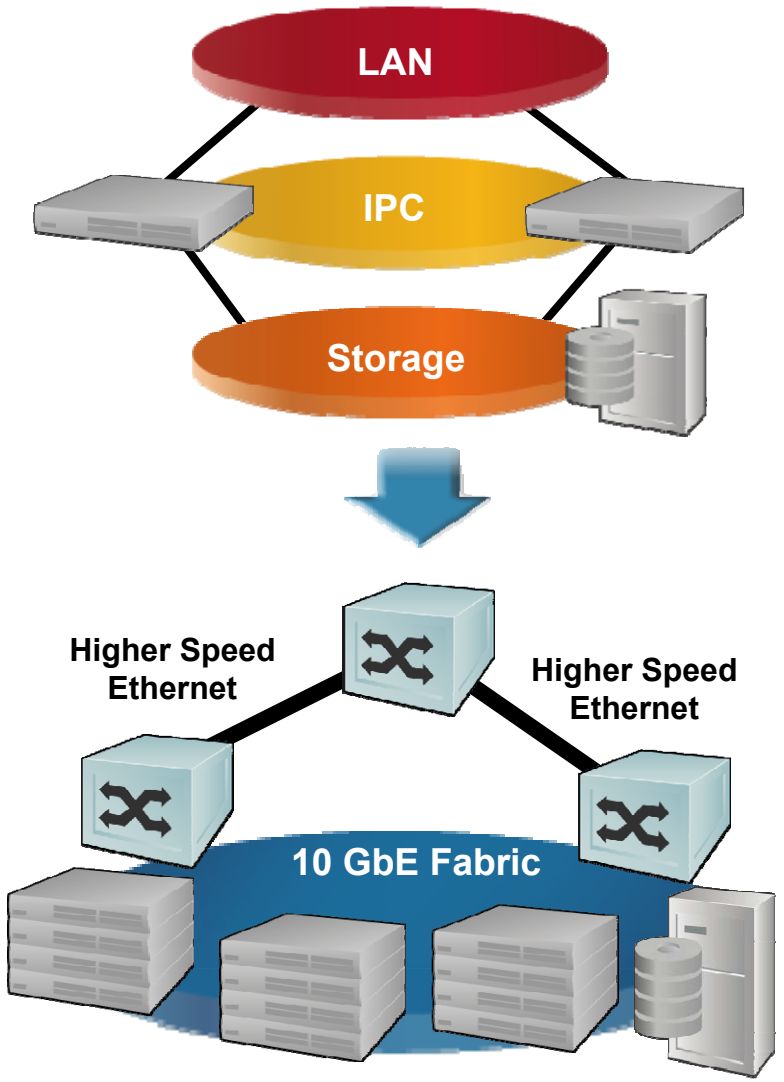
- 4 years between standard and deployment
- GE captures 50% of market in 3 years (2002 - 2005)
- 10GE enters top500 in 2006 (included in Gigabit Ethernet #'s)
- 12x increase in Avg Gigaflops: 10x increase in Ethernet Interconnect

Forecast

- 10x increase needed at 84,000 Gigaflops (2010)
- There will always be early adopters.

Data Center

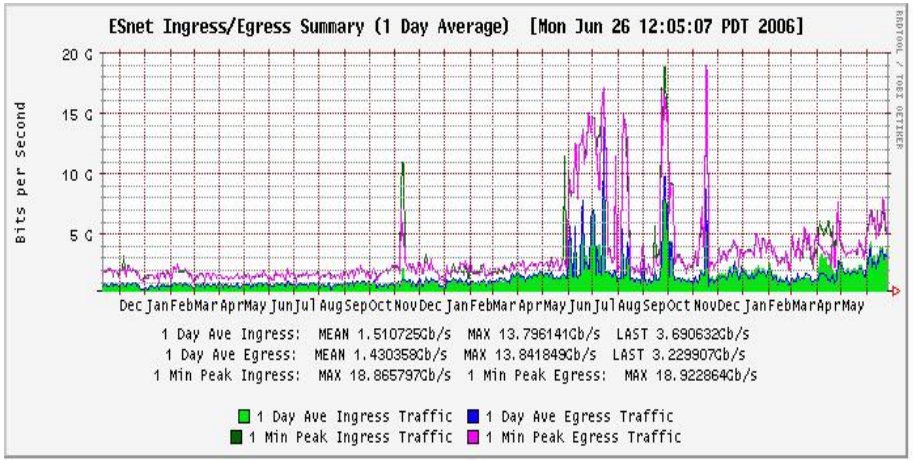
- Fabric consolidation
  - Converging LAN, storage and interprocessor communication (IPC)
  - TOE/TCP acceleration for LAN traffic
  - iSCSI for storage market
  - iWARP/RDMA for low latency IPC
- Modular platforms
  - Shift to blade server deployments
  - Bandwidth aggregation at switch blade
  - Multicore/multithreaded CPUs handle increase bandwidth availability
- Bandwidth hierarchy
  - Consolidation increases traffic bandwidth
  - Modular increases uplink bandwidth requirement



Consolidated Fabric

Research and Development

ESnet Daily Ingress / Egress Traffic



Source: ESnet / LBNL

- **Network Utilization 2006**
  - **From Dec 2003 to June 2006 – Mean 1.5 Gb/s**
  - **Min peak traffic has reached 18.9 Gb/s**
  - **Minimum of 5x of min peak to mean average needed**
- **Forecast 2010**
  - **Mean – 30 Gb/s**
  - **Min peak traffic forecast 150 Gb/s**

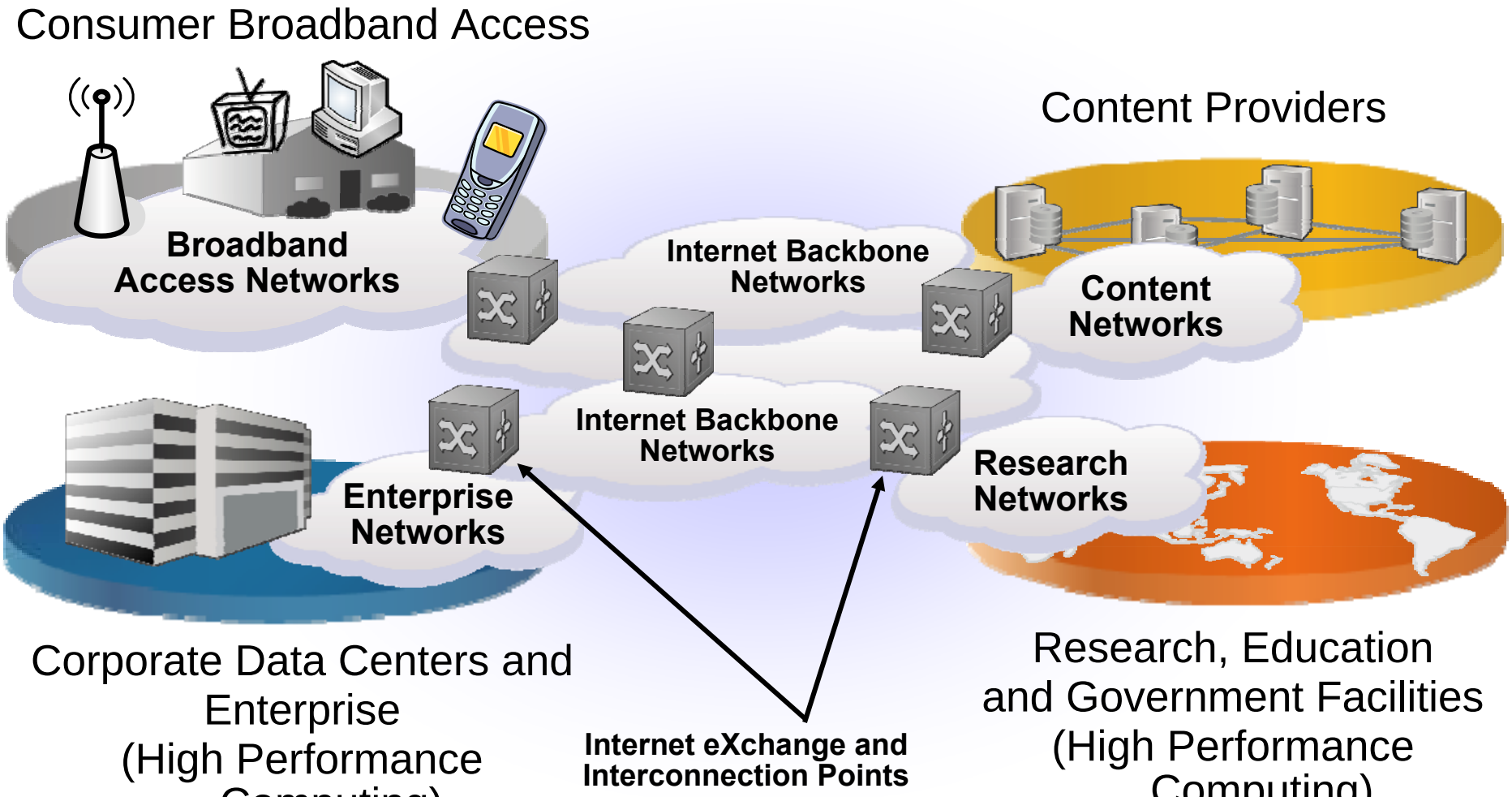
Evolving Quantitative Science Requirements for Networks

Science Areas	Today *End2End* Throughput	5 years End2End Throughput	5-10 Years End2End Throughput	Remarks
High Energy Physics	0.5 Gb/s	**100 Gb/s**	**1000 Gb/s**	high bulk throughput
Climate (Data & Computation)	0.5 Gb/s	**160-200 Gb/s**	**N x 1000 Gb/s**	high bulk throughput
SNS NanoScience	Not yet started	**1 Gb/s**	**1000 Gb/s + QoS for control channel**	remote control and time critical throughput
Fusion Energy	0.066 Gb/s (500 MB/s burst)	**0.198 Gb/s (500MB/ 20 sec. burst)**	**N x 1000 Gb/s**	time critical throughput
Astrophysics	0.013 Gb/s (1 TBy/week)	**N*N multicast**	**1000 Gb/s**	computational steering and collaborations
Genomics Data & Computation	0.091 Gb/s (1 TBy/day)	**100s of users**	**1000 Gb/s + QoS for control channel**	high throughput and steering

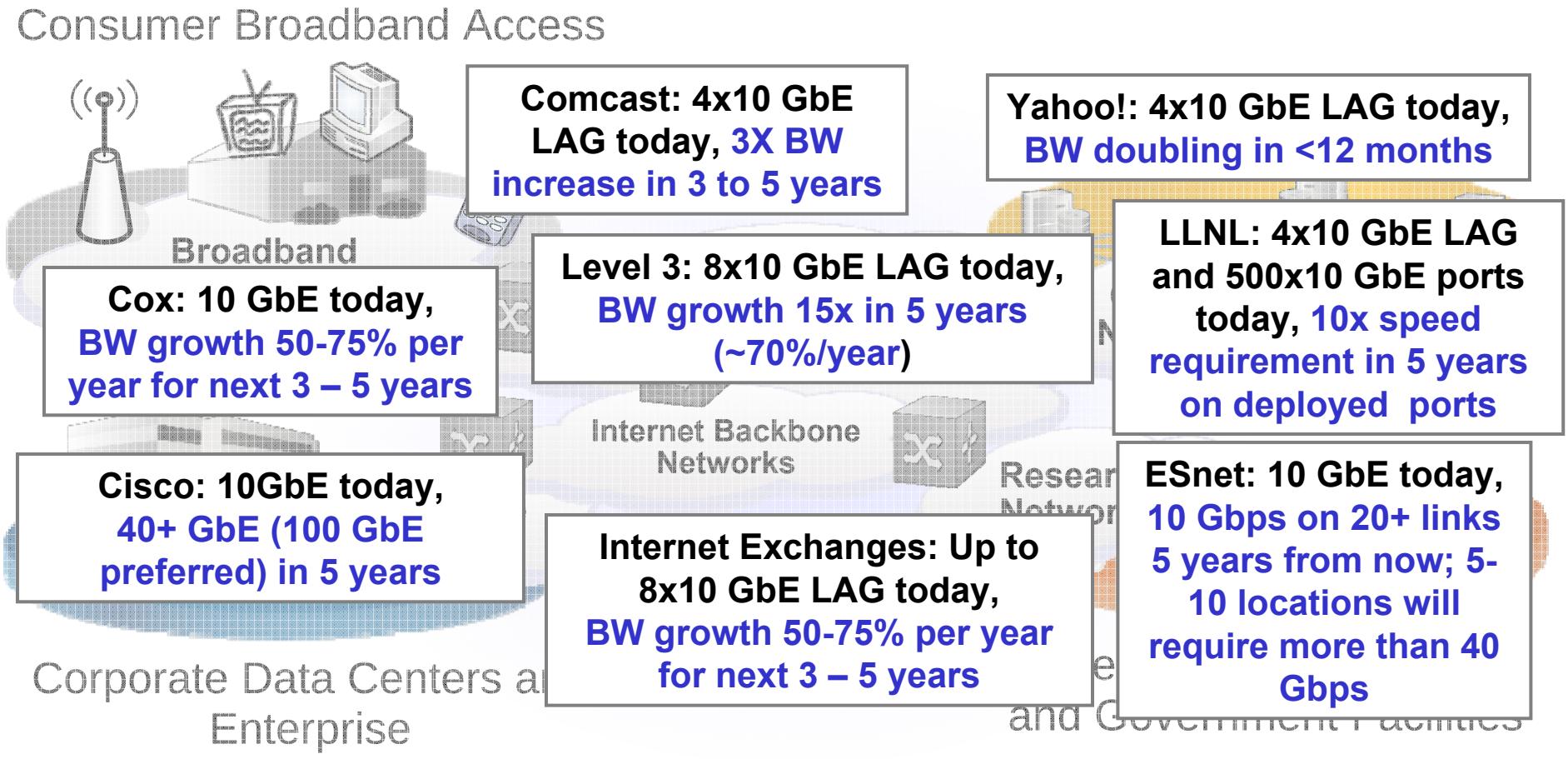
Used with permission from ESnet

Multiple projects that could have significant impact on Transport projection over the next 5 to 10 years.

The Ethernet Ecosystem



Bandwidth and Growth Projections



Comcast: 4x10 GbE LAG today, 3X BW increase in 3 to 5 years

Yahoo!: 4x10 GbE LAG today, BW doubling in <12 months

Cox: 10 GbE today, BW growth 50-75% per year for next 3 – 5 years

Level 3: 8x10 GbE LAG today, BW growth 15x in 5 years (~70%/year)

LLNL: 4x10 GbE LAG and 500x10 GbE ports today, 10x speed requirement in 5 years on deployed ports

Cisco: 10GbE today, 40+ GbE (100 GbE preferred) in 5 years

Internet Exchanges: Up to 8x10 GbE LAG today, BW growth 50-75% per year for next 3 – 5 years

ESnet: 10 GbE today, 10 Gbps on 20+ links 5 years from now; 5-10 locations will require more than 40 Gbps

Conclusions

- **Broadband penetration / access speeds / high bandwidth content driving core infrastructure bandwidth**
- **Multiple applications need higher speed**
  - **Portals / Content providers: bandwidth needs track with broadband growth**
  - **Video on Demand: Regional networks already in 10's of Gbps**
  - **Service Providers: 8 x10GbE links now, 10's x 10GbE in future**
  - **Internet eXchanges: aggregate point for Internet traffic**
  - **High Performance Computing: ≈12x improvement in processing drives 10x jump in Ethernet interconnect**
  - **Data Center: consolidation, server and storage consolidation, large clusters**
  - **Research & Development – 2010 Forecast- 30 Gbps mean traffic, 150 Gbps min peak traffic,**
- **10 GbE / w LAG is hitting scaling limits today**
- **Ethernet is the preferred technology in the end-to-end ecosystem**

The Technical Viability of Higher Speed Ethernet

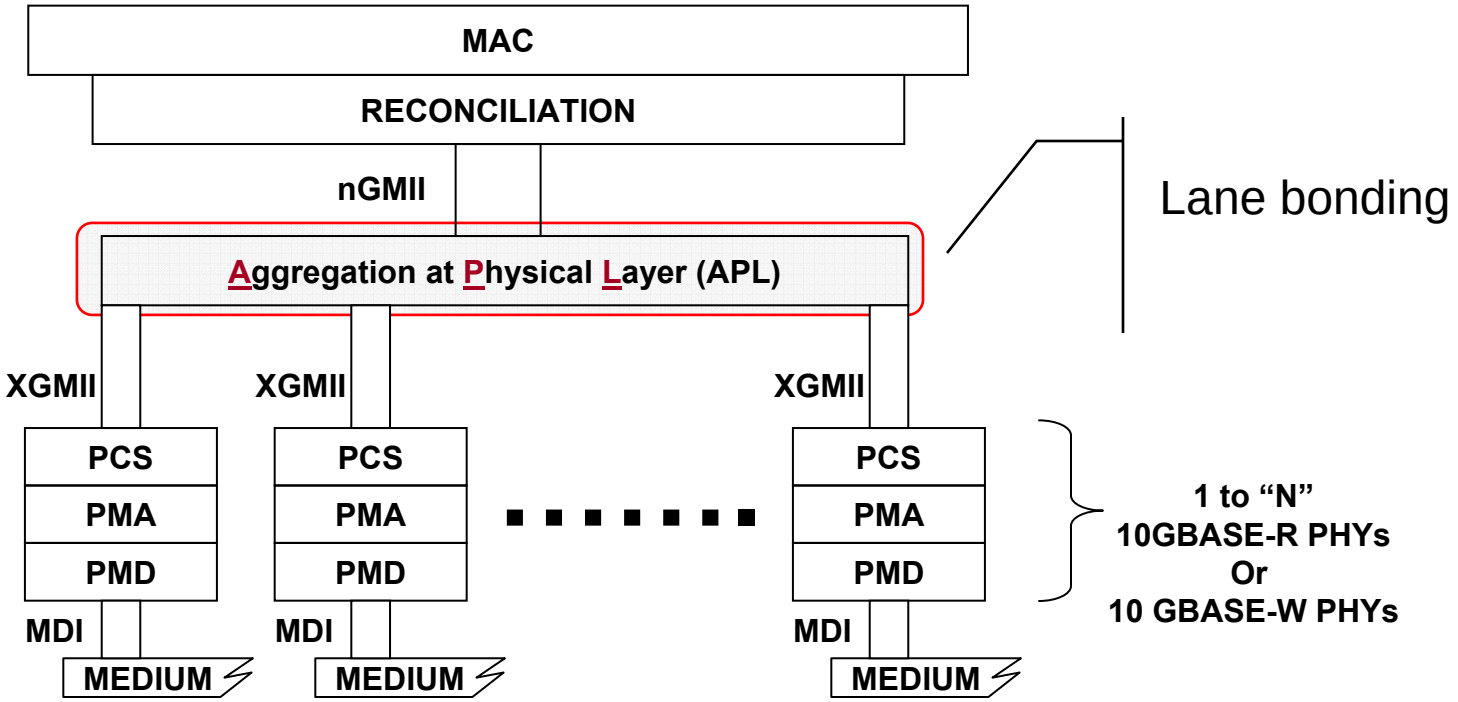
Presented by
Petar Pepeljugoski, IBM Research

IEEE 802.3 Working Group
San Diego, CA
July 18, 2006

Introduction

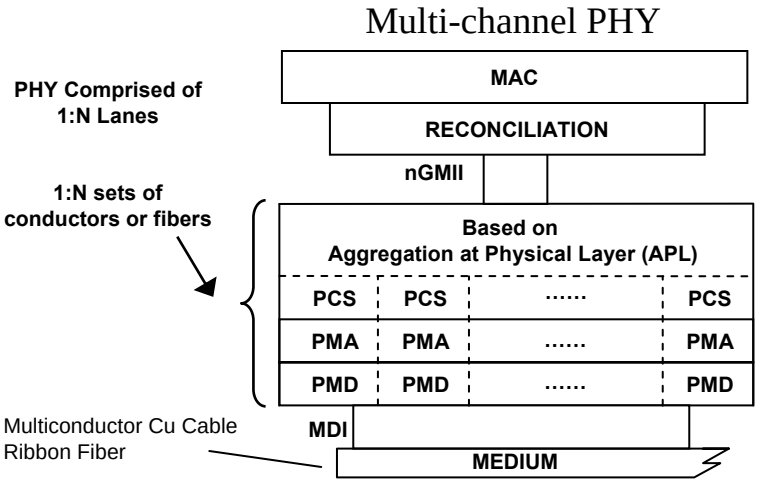
- **Marketing Input for application spaces**
  - **“LAN” reach (tens to hundreds of meters)**
  - **“WAN” reach (hundreds of meters to tens of kilometers)**
- **Solution space**
  - **Faster is an option, but not the only one**
  - **“Higher” speed means a “fatter” pipe**
    - **By speed**
    - **By number of wavelengths**
    - **By number of lanes**
- **Issues to consider**
  - **Architecture complexity**
  - **Component maturity**

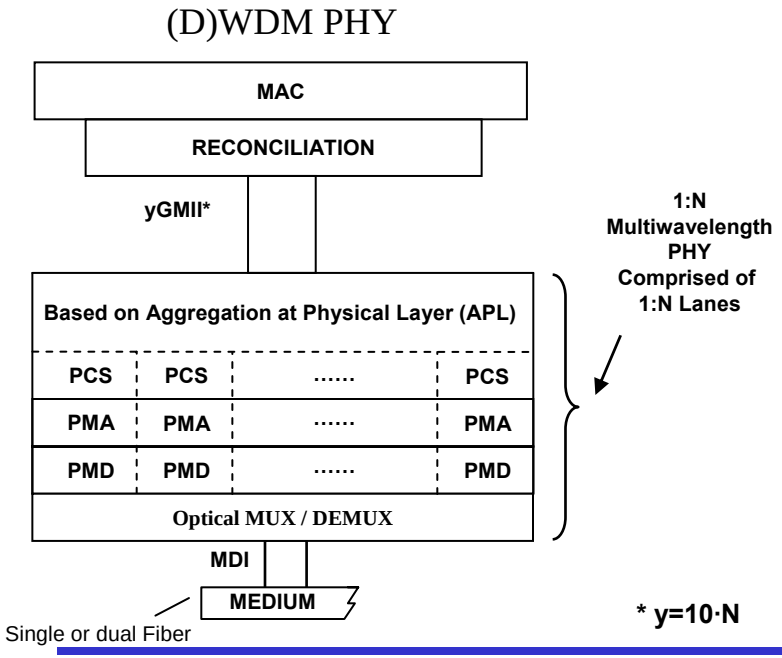
Example - Aggregation at the Physical Layer (APL)

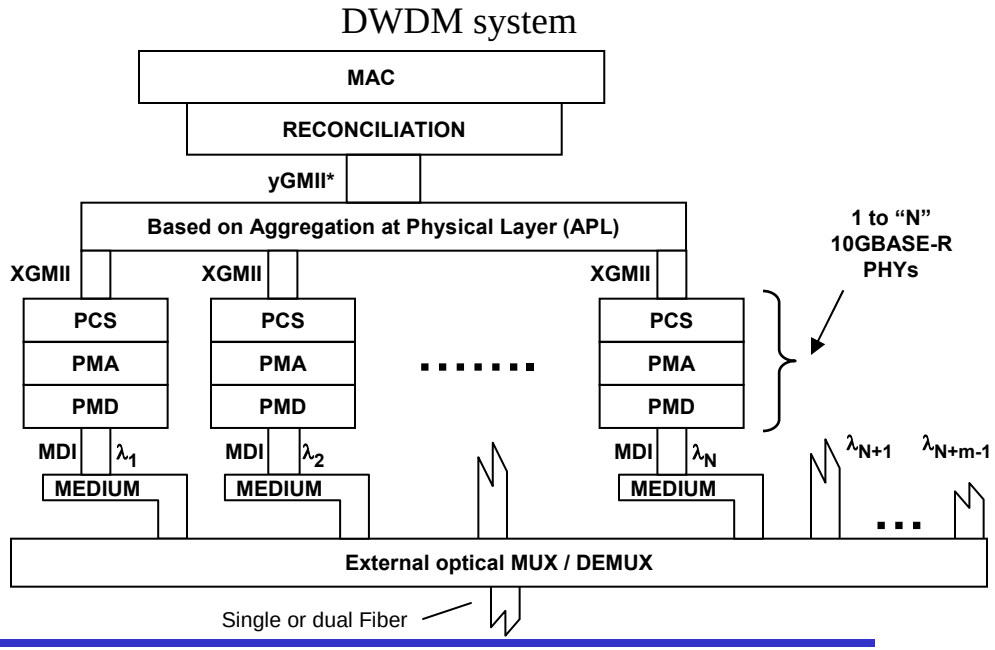


Using APL to Create a Fatter Pipe

- **Multi-channel PHY**
  - **Multicore cable, ribbon fiber**
  - **Parallel backplane channels**
- **Multi-wavelength (WDM) PHY**
  - **N wavelengths on single fiber pair**
- **Multi-wavelength (DWDM) system**
  - **Single wavelength per module**
  - **External optical MUX/DEMUX**

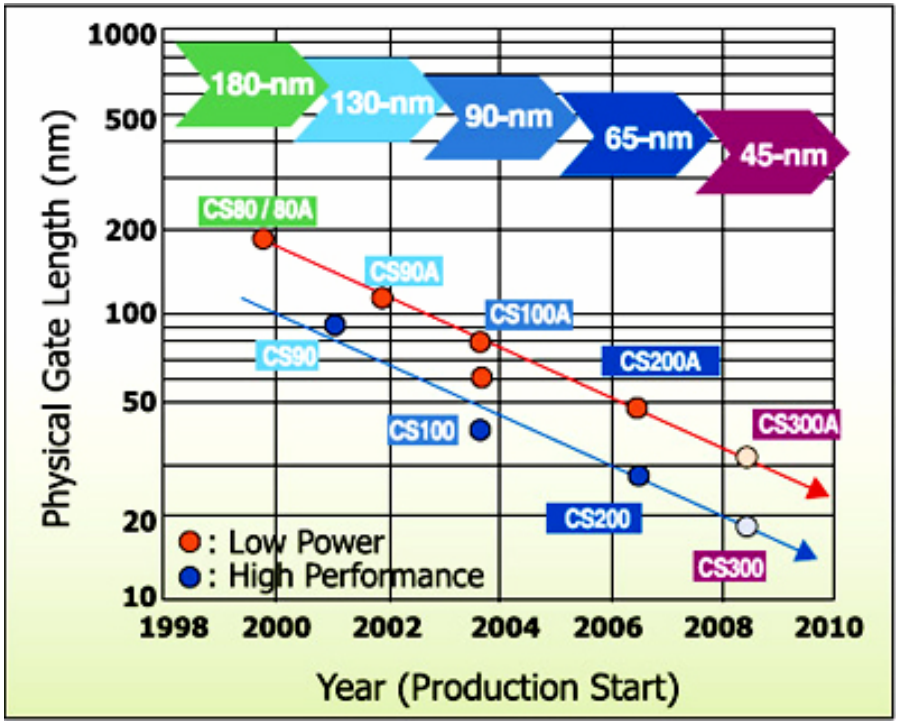






ASIC Technology Roadmap

- Typical 10 Gb/s MAC
  - 64-bit wide
  - 156.25 MHz
  - 90nm or 130 nm silicon
- MAC Integration with other circuitry
- Implementation discussions suggest
  - Speed improvement possible
    - by data path increase
    - by frequency increase
  - 45-90 nm ASIC processes



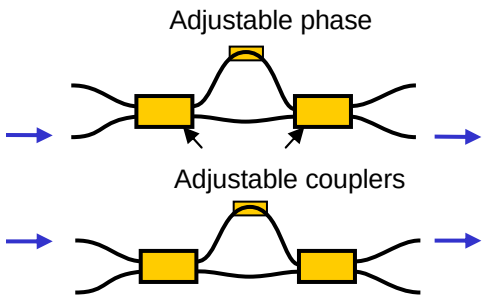
Used with permission from Fujitsu

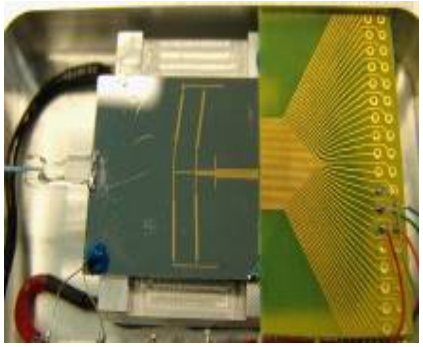
PMD Components

- PIN – bandwidths supporting speeds beyond 10 GBaud available
- Drivers and TIAs
  - Challenging for serial solution
- Optical connectors
  - LC, SC, ST, FC/PC and variations for serial/WDM, MTP/MPO for parallel links
  - Other connector types for reduced cost?
- WDM muxes and demuxes, optical equalizers etc.
  - Integrated photonics improves density, power consumption
- Sources:
  - Laser Diodes
  - 850nm 10 Gb/s VCSEL (singlets) widely available
    - Arrays demonstrated
  - 1300nm VCSELs up to 4 Gb/s,
    - progress expected towards 10 Gb/s
  - 1300nm and 1550nm edge emitting laser in wide use at 10 Gb/s
  - Directly modulated laser diodes – speeds beyond 10 Gb/s demonstrated
  - External modulators integrated in CMOS and InP
- Optical amplifiers
- Equalization
  - Links utilizing equalization to be standardized (802.3ap, 802.3aq)

Integrated Photonics

- **Optical equalizer implemented as silica planar lightwave circuit**



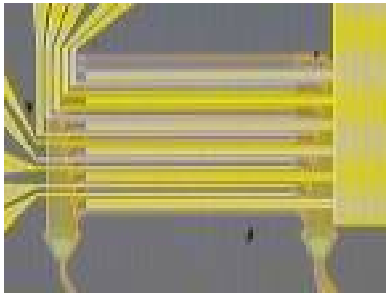


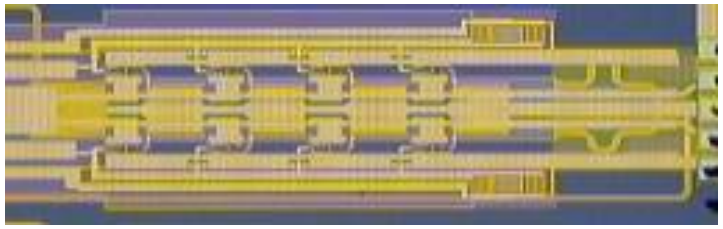
chip size: 35mm x 45mm

Doerr, used with permission from Lucent

- **DWDM muxes and demuxes (10-l) implemented as photonic integrated circuits**
- **Silicon optical filters (DWDM) multiplexing/demultiplexing**
  - **Electrically tunable integrated with control circuitry**

- **Silicon 10 Gb/s modulators**
  - **Low loss, low power consumption**
  - **Excellent signal integrity**



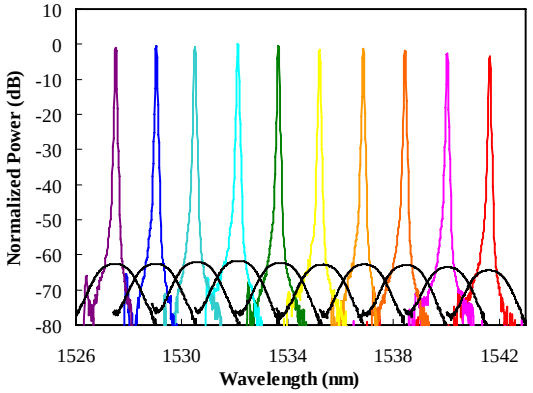


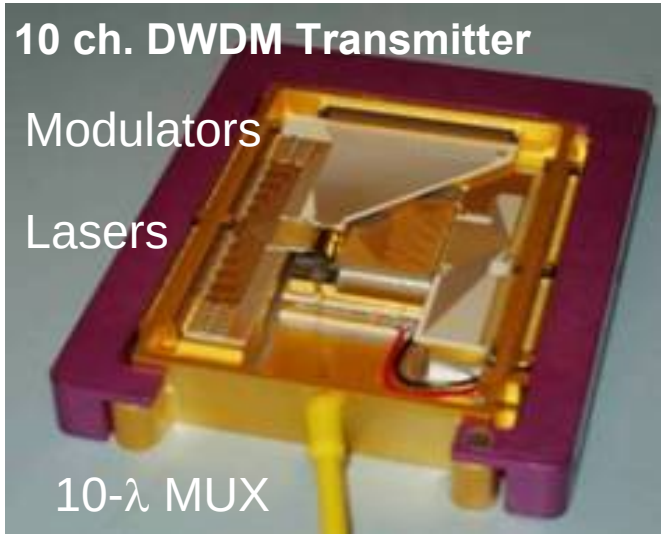
R. Merel, used with permission from Luxtera

DWDM Example: (N) Channels x (M) Gb/s (DWDM) Demo: 10x10Gb/s DWDM

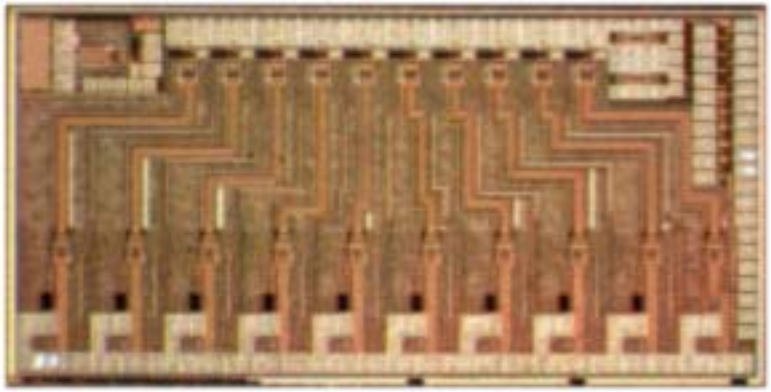
- **Transmitter and Receiver use Photonic Integrated Circuits**
  - Commercially deployed in carrier networks since 2004

10 ch. DWDM Spectrum





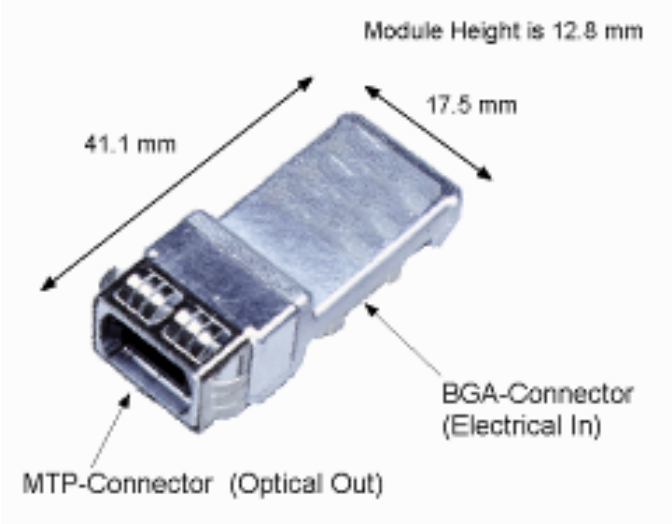
D. Perkins, used with permission from Infinera

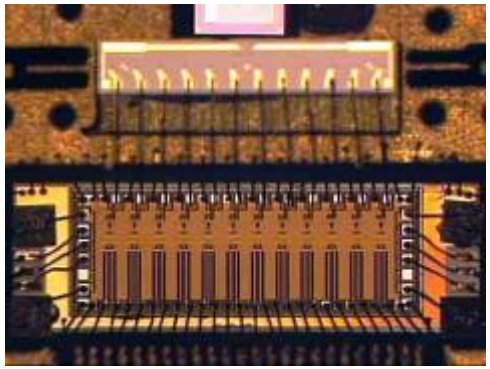


10 ch. TIA

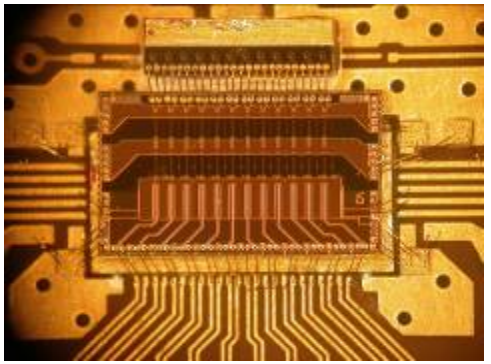
Multi-channel (Parallel) Example: (N) Ch. x (M) Gb/s Demo: 120 Gb/s, 12 x 10 Gb/s

- For 40 Gb/s, commercially available today
- SNAP12 (12x3.125Gb/s)
  - extension to 12x10Gb/s demonstrated in 2003:
  - SiGe LDD, SiGe Rx, 10Gb/s lasers
- 12 channels running at BER <10^{-12}, link length 316m using OM3 fiber
- 12 channels, 9.9-11.0 Gb/s/channel tester designed and built.





12 ch. TX OE chips



12 ch. RX OE chips

D. Kuchta, used with permission from IBM

Communications Industry Snapshot

- Established technology
  - Aggregation
    - VCAT and LCAS
    - 802.3ah type aggregation
    - XAUI in 802.3ae
    - Inverse Muxing over ATM
    - PPP Multilink
    - DOCSIS 3.0 channel bonding
  - ITU-T
    - DWDM wavelength plan
    - WDM component specs
    - "Black box" and "black link" DWDM specification methodologies
  - Ribbon fibers, connectors and lane definitions
- Industry Related Activities
  - SNAP12 and POP4 MSAs
  - Infiniband (IB) Standard for up to 96 Gb/s (120 GBaud)
    - Variable lane count (1, 4, 8, 12) at 2, 4 and 8 Gb/s lane speed
  - Optical Internetworking Forum
    - CEI-25 Project (20 – 25 Gb/s Electrical Signaling)
    - 100 – 160 Gb/s Interfaces
  - OC-768 Deployment

Conclusions

- **Aggregation at the Physical Layer (APL)**
  - **Lane Bonding in multiple 802.3 clauses**
  - **Scaleable 1 to "N" lane PHY's**
- **Building block technologies**
  - **Existing PMD technology**
  - **Multi-wavelength devices demonstrated**
  - **Multi-channel devices demonstrated**
  - **Repackaging of existing 10Gb/s optical technology**
- **Finer ASIC processes forecasted to be available if required for MAC and MAC clients**

- **Technology exists to enable Higher Speed Ethernet**

Higher Speed Ethernet - Why Now?

Presented by
John D'Ambrosia, Force10 Networks

IEEE 802.3 Working Group
San Diego, CA
July 18, 2006

The Need for Higher Speed

- **Work of the past 25 years is paying off!**
  - **(Traditional) LAN deployment**
  - **Non LAN deployment**
  - **Technology development**
- **Traffic is growing everywhere**
  - **The masses have more ways of faster access**
  - **Higher bandwidth content**
  - **New applications enabled**
- **Need for higher speed by multiple applications**

Summary

- Applications are challenging today's solutions
- "Higher" Speed needed throughout the entire Ecosystem
- Needed by 2010 for multiple applications

- Past efforts took 3 to 4 years
  - 10GE
  - EFM
  - 10GBASE-T

- We need to begin the process and study the problem

Supporters

David Law	3COM
Kam Patel	ADC
David Yanish	ADC
Adam Healey	Agere
Henk Steenman	AMS-IX
Howard Baumer	Broadcom
Kevin Brown	Broadcom
Wael Diab	Broadcom
Howard Frazier	Broadcom
Ali Ghiasi	Broadcom
Pat Thaler	Broadcom
Arthur Marris	Cadence
Alessandro Barbieri	Cisco
Hugh Barrass	Cisco
Mark Nowell	Cisco
Tom Lindsay	ClariPhy Communications
Menachem Abraham	Columbus Advisors
John Abbott	Corning
Steve Swanson	Corning
John Dallesasse	Emcore
Lane Patterson	Equinix
Arne Alping	Ericsson
Eddie Tsumura	Excelight Communications
Andy Moorwood	Extreme Networks
Jonathan King	Finisar
Joel Goergen	Force10 Networks
Pete Tomaszewski	Force10 Networks
Greg Chesson	Google
Petar Pepeljugoski	IBM Research
Larry Rubin	Independent
Drew Perkins	Infinera
Ilango Ganga	Intel
Gopal Hegde	Intel
Schelto van Doorn	Intel
Toshinori Ishii	Internet Multifeed Co.
Takejiro Takabayashi	Japan Internet Exchange Co. Ltd.
Wenbin Jiang	JDSU
Mike McConnell	KeyEye Communications
Itsuro Morita	KDDI Labs
Alan Flatman	LAN Technologies
Mike Bennett	LBNL
Joe Lawrence	Level3
Roger Merel	Luxtera
Chris Di Minico	MC Communications
Rolf McClellan	McClellan Consulting
Gourgen Oganessyan	Molex
Jimmy Sheffield	Newport Enterprises
Peter Schoenmaker	NTT America
David Martin	Nortel
Kirk Bovill	OCP
Robert Lingle	OFS
George Oulundsen	OFS
Rich Taborek	Onde Technology
Ed Cornejo	Opnext
Matt Traverso	Opnext
Mike Dudek	Picolight
Jack Jewell	Picolight
Brian Holden	PMC-Sierra
Dan Dove	ProCurve by HP
Brad Booth	Quake Technologies
Rami Kanama	Redfem Integrated Optics
John Monson	Scintera Networks
George Zimmerman	Solarflare
Rick Rabinovich	Spirent Communications
Shimon Muller	Sun
Stephan Korsback	Switchcore
Glen Kramer	Teknovus
Andre Szczepanek	Texas Instruments
Ralf-Peter Braun	T-Systems
Suveer Dhamejani	Tyco Electronics
Eric Lynskey	UNH-IOL
Bob Noseworthy	UNH-IOL
Martin Carroll	Verizon
Frank Chang	Vitesse
Kevin Daines	World Wide Packets
Tom Palkert	Xilinx
Adam Bechtel	Yahoo!

Straw Polls

Call-For-Interest

- **Should a Study Group be formed for "Higher Speed Ethernet"?**

  Y: 147 N: 9 A: 31

Participation

- **I would participate in the "Higher Speed" Study Group in IEEE 802.3.**

  Tally: 108

- **My company would support participation in the "Higher Speed" Study Group in IEEE 802.3**

  Tally: 76

Future Work

- **Ask 802.3 to form Higher Speed SG on Thursday**
- **If approved**
  - **802 EC informed of Higher Speed SG on Friday**
  - **First HSSG meeting, week of September 2006 IEEE 802.3 Interim.**

Thank You!

Contributors

- David Law, 3COM
- Henk Steenman, AMS-IX
- Howard Frazier, Broadcom
- Pat Thaler, Broadcom
- Hugh Barrass, Cisco
- Ian Cox, Cisco
- Andy Moorwood, Extreme Networks
- Dave Hawley, Extreme Networks
- Steve Garrison, Force10 Networks
- Joel Goergen, Force10 Networks
- Greg Hankins, Force10 Networks
- Petar Pepeljugoski, IBM Research
- Drew Perkins, Infinera
- Itsuro Morita, KDDI Labs
- Mike Bennett, LBNL
- Joe Lawrence, Level(3) Communications
- Marcus Duelk, Lucent
- Roger Merel, Luxtera
- Rolf McClellan, McClellan Consulting
- Brad Booth, Quake Technologies