400 Gigabit Ethernet Call-For-Interest Consensus

IEEE 802.3 Ethernet Working Group
IEEE 802 March 2013 Plenary, Orlando, FL

Objective for this Meeting

- To ***measure the interest*** in starting a study group to address ***400 Gb/s Ethernet*** *interconnect*
  - *Defining Carrier Transport beyond the scope of this effort*
- 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

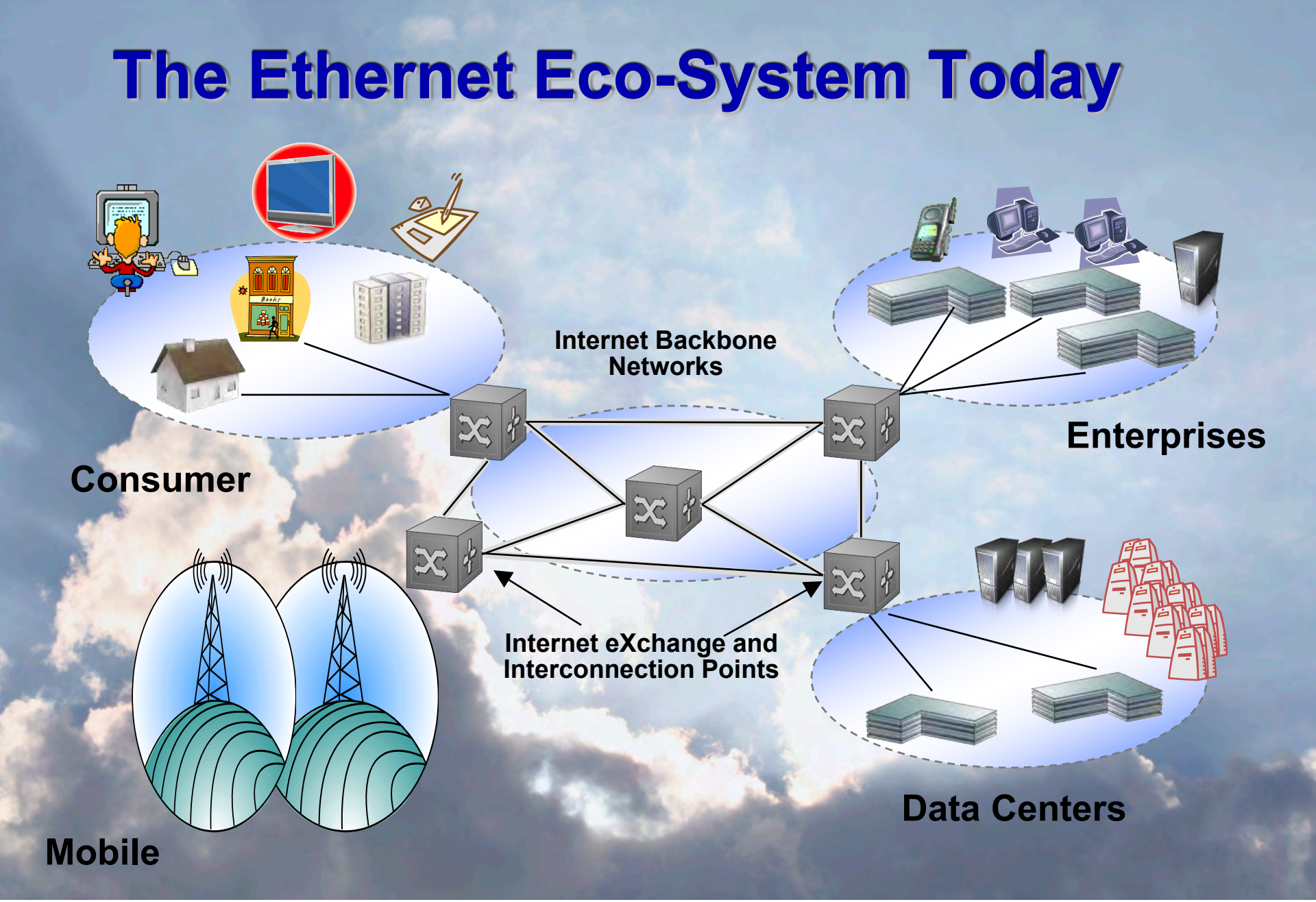
- **Presentations**
  - **“The Bandwidth Explosion,” David Ofelt, Juniper**
  - **“Beyond 100 Gigabit Ethernet, Technical Challenges,” Mark Nowell, Cisco.**
  - **“400 Gigabit Ethernet- Why Now,” John D’Ambrosia, Dell.**
- **Straw Polls**

Presented by

IEEE 802.3 Working Group
Orlando, FL, USA
March 19, 2013

THE BANDWIDTH EXPLOSION

The Ethernet Eco-System Today



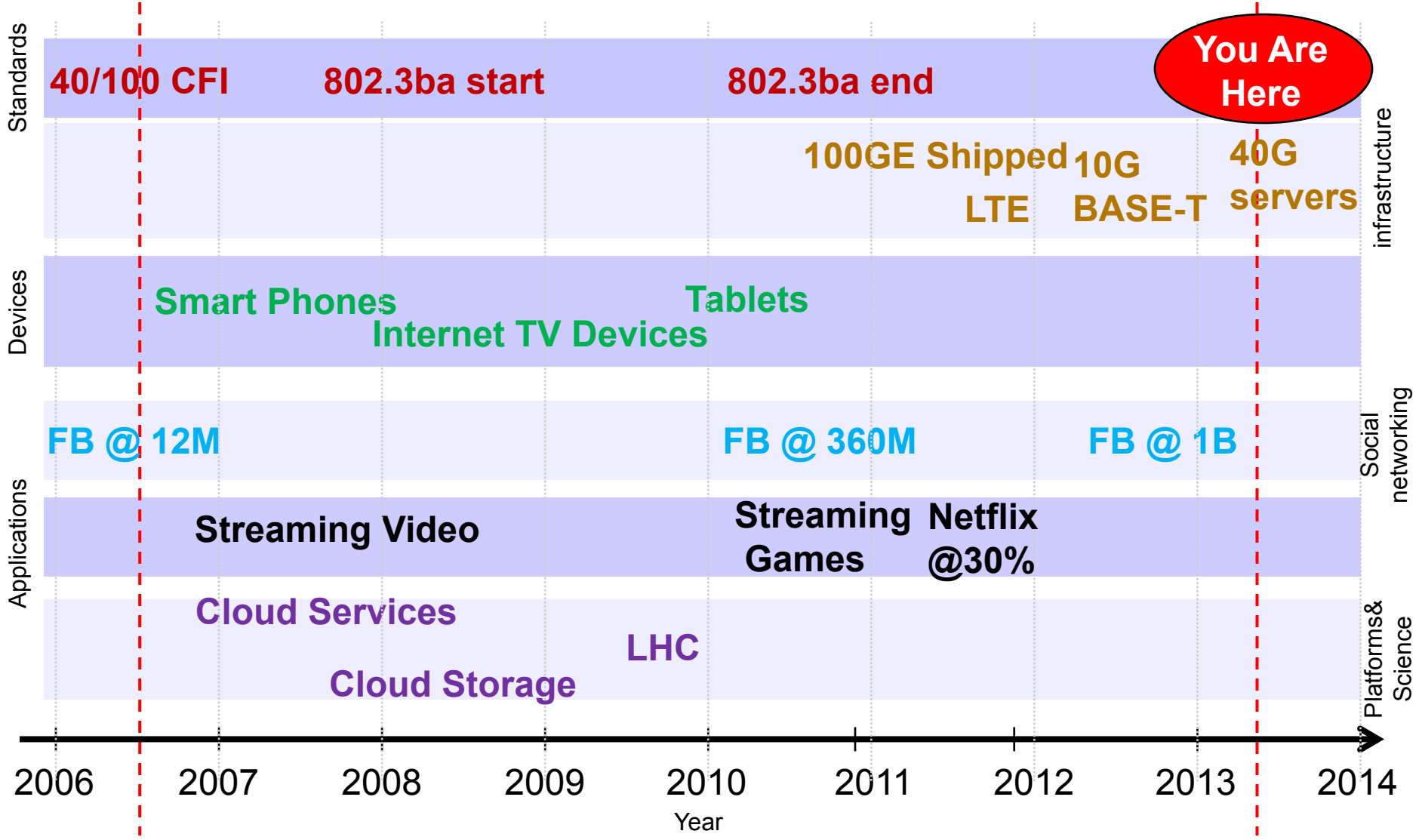
Changes since 2007 HSSG

- **Infrastructure / Devices**
  - Smart Phones
  - Tablets
  - Wi-Fi Deployments
  - 3G / 4G / LTE
  - 10G Server Deployment
  - Internet Enabled TV
  - The “Cloud”
- **Applications**
  - Cloud-based Businesses
  - Practical Cloud Storage
  - Ubiquitous Video Streaming
  - Social Media Explosion
  - Video Calling Commonplace
  - New Database Technology
  - Online Gaming

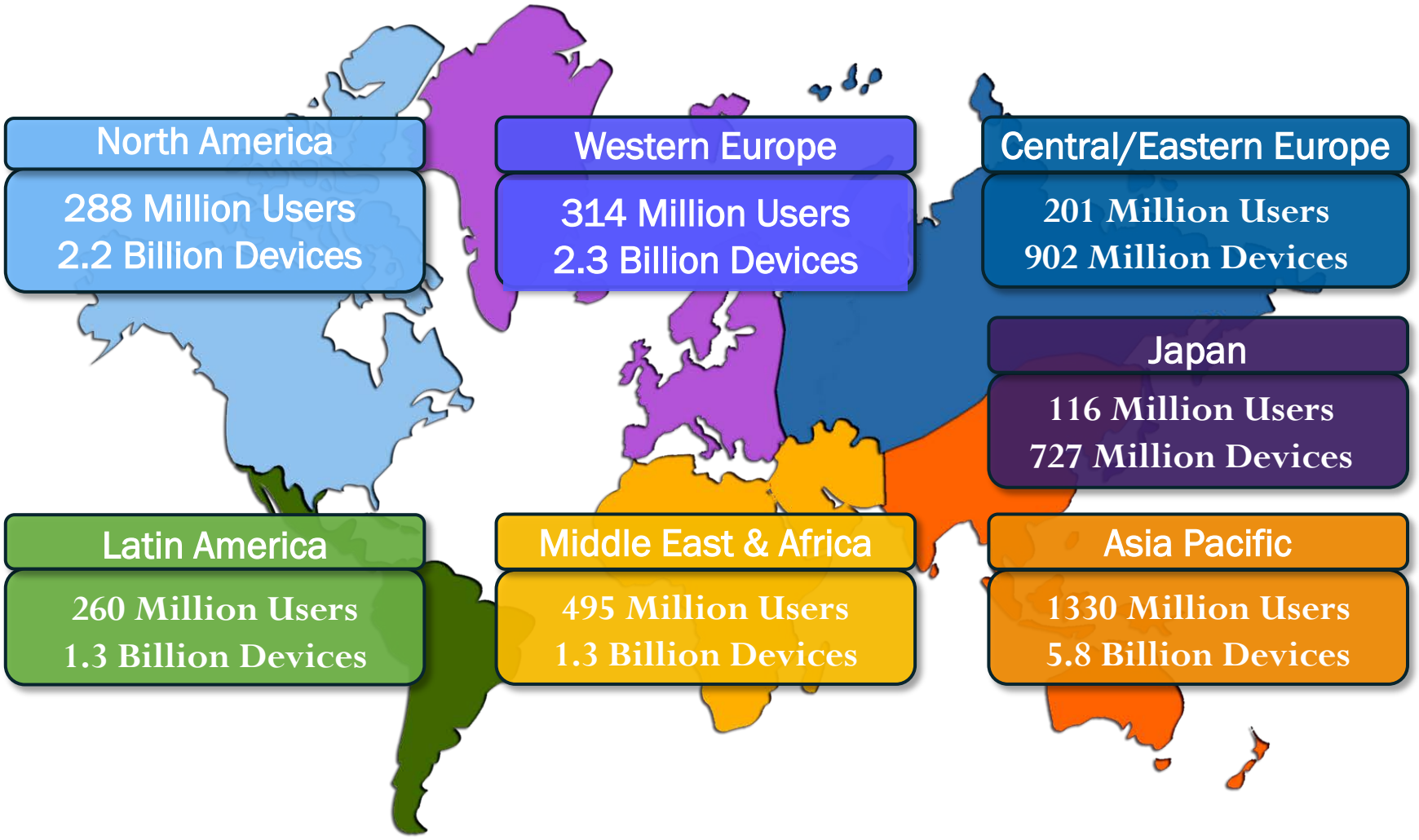
Device	Traffic Multiplier
Tablet	1.1
64-bit laptop	1.9
Internet Enabled TV	2.9
Gaming Console	3.0
Internet 3D TV	3.2

Compared against a 32 bit laptop

Highlights since IEEE P802.3ba



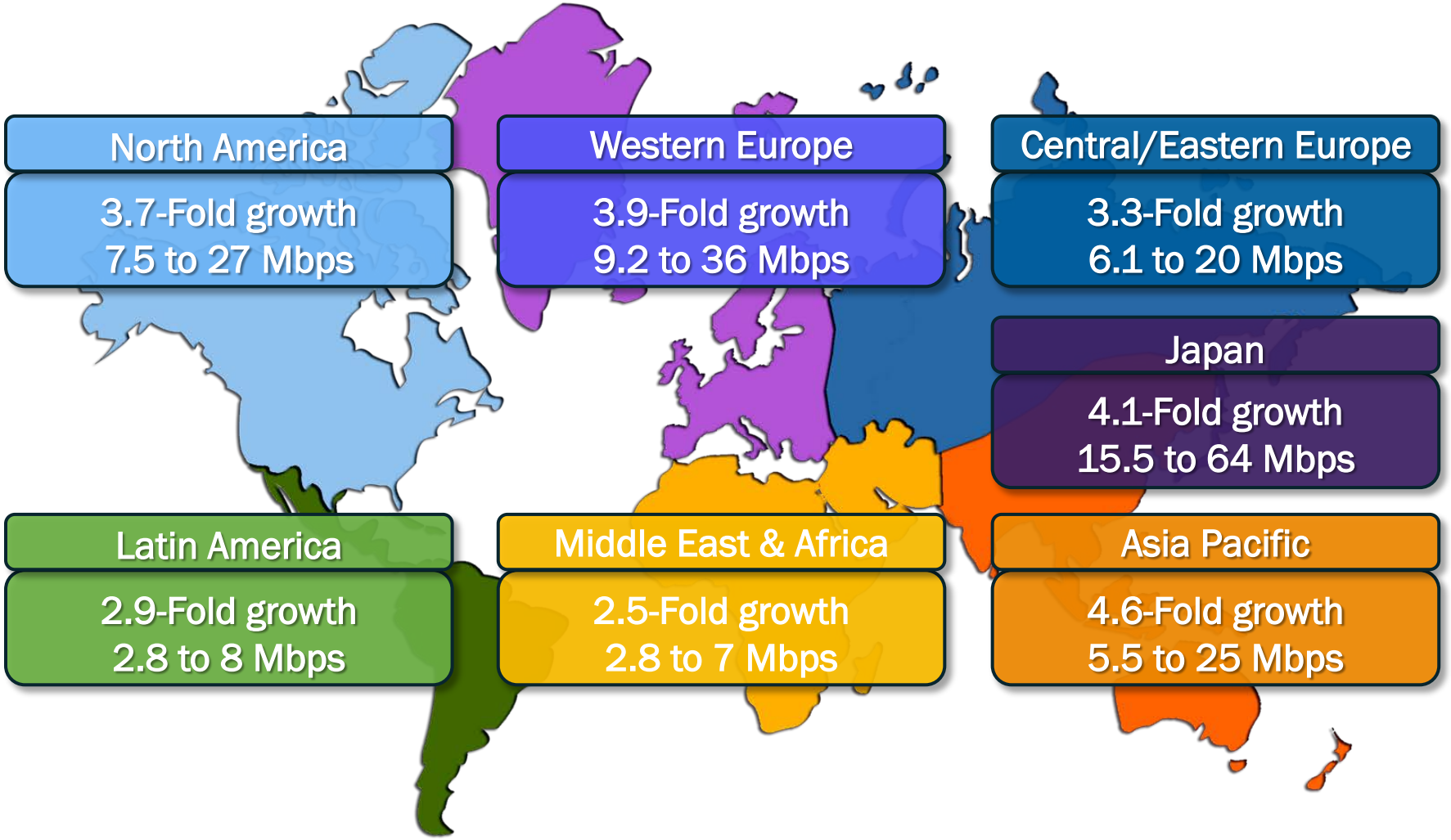
2015 Global Users and Network Connections



Source: nowell_01_0911.pdf citing Cisco Visual Networking Index (VNI) Global IP Traffic Forecast, 2010–2015, http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/nowell_01_0911.pdf

Global Broadband Speed 2010-2015

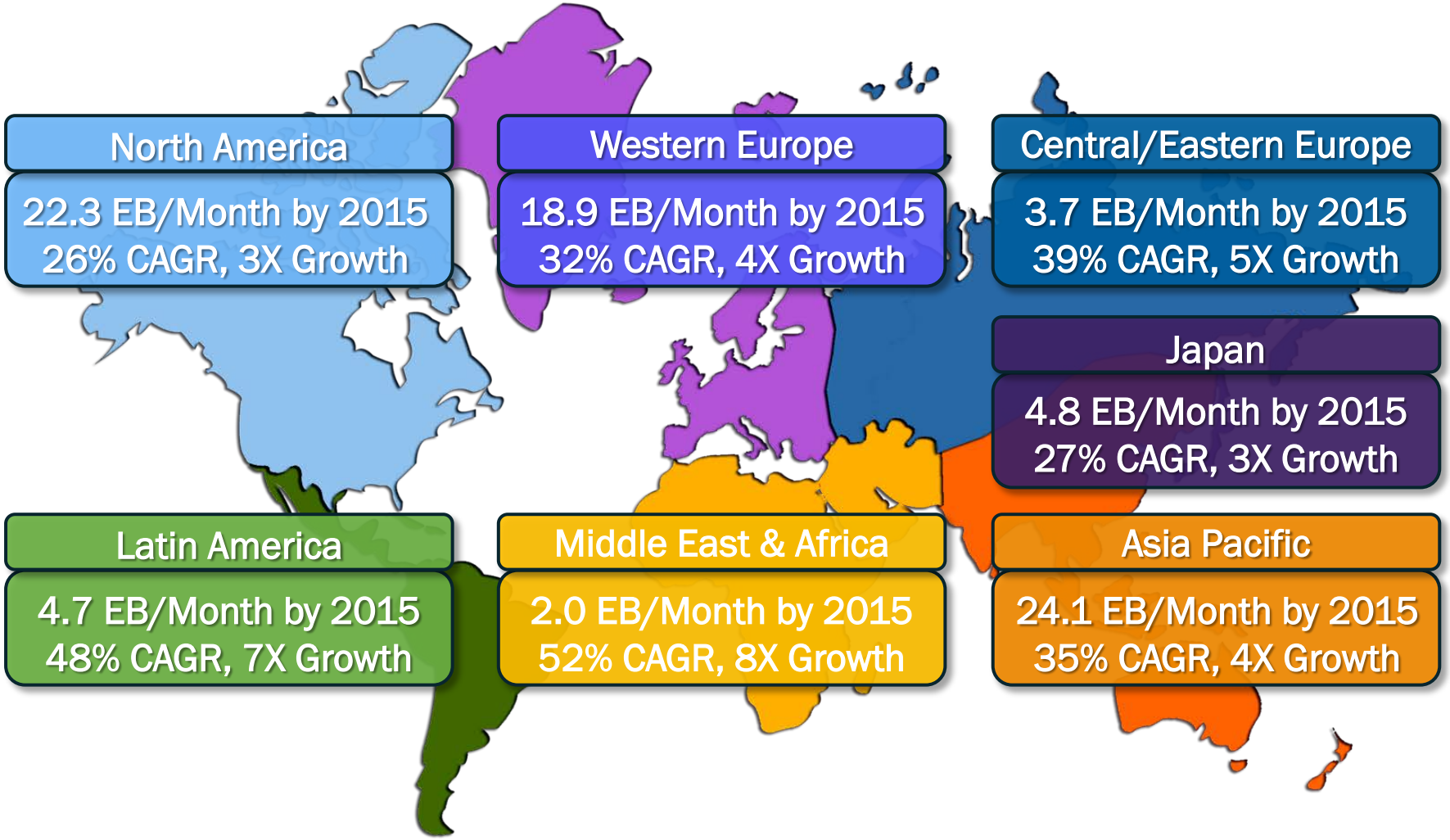
Average broadband speed will grow 4X; from 7 to 28 Mbps



Source: nowell_01_0911.pdf citing Cisco Visual Networking Index (VNI) Global IP Traffic Forecast, 2010–2015, http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/nowell_01_0911.pdf

Global IP Traffic Growth, 2010–2015

Regional contributions to the Zettabyte journey



Source: nowell_01_0911.pdf citing Cisco Visual Networking Index (VNI) Global IP Traffic Forecast, 2010–2015, http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/nowell_01_0911.pdf

Video (Some Interesting Facts & Forecasts)

- **Forecast May 30, 2012***
  - **By 2016 – 1.2 million video minutes traveling Internet every second**
  - **# of Internet video users: 0.792B (2011) to 1.5B (2016)**
  - **Mobile video – fastest growing mobile service, 0.271B (2011) to 1.6B (2016)**
  - **Desktop videoconferencing – 26.4M (2011) to 218.9M (2016)**
- **Forecast Feb 2013 ****
  - **Mobile video to represent 66% of all mobile data traffic by 2017**
- **YouTube Statistics *****
  - **Every minute - 72 hours of video are uploaded**
  - **Each month 4 billion hours of video are watched**
  - **25% of global views come from mobile devices**
  - **Traffic from mobile devices tripled in 2011**

- *** May 30, Press Release, "2012, Cisco VNI Forecast," http://newsroom.cisco.com/press-release-content?articleId=888280.**
- **** Visual Networking Index , Cisco, http://www.cisco.com/en/US/netsol/ns827/networking_solutions_sub_solution.html#~forecast, Feb, 2013.**
- *****Youtube Statistics, http://www.youtube.com/t/press_statistics, data obtained Feb 15, 2013.**

Science: Big Data Sources

CERN is "the tip of the iceberg"

Today

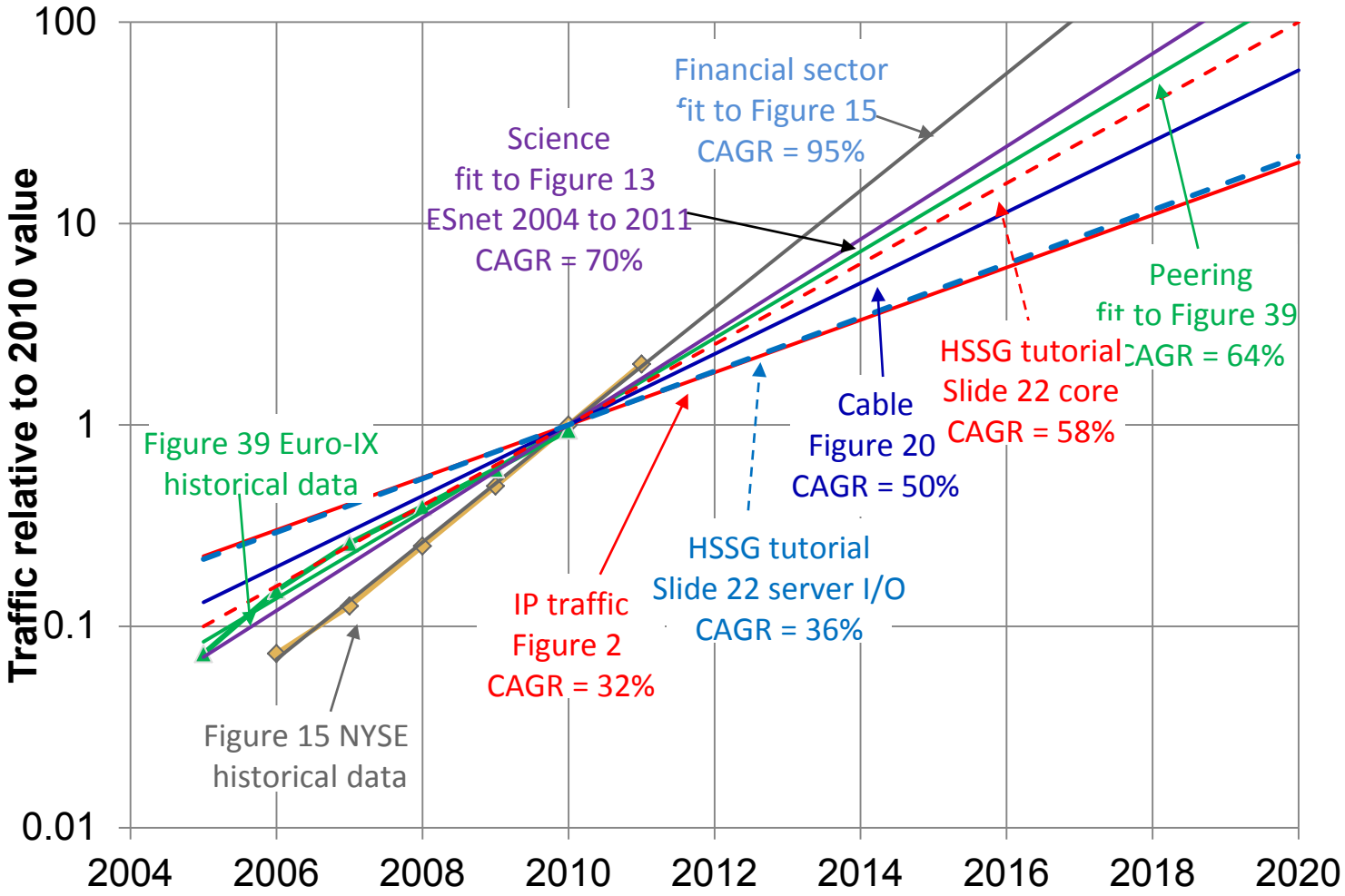
- CERN
  - Atlas detector in LHC (Large Hadron Collider) generates ~1 petabyte/sec
  - Trigger farm reduces to 450MB/sec
    - Tens of Gb/s of outbound traffic to analysis centers
- Genome sequencing
  - Per-instrument data rate strongly ↗ (~10x over 5 years)
  - Data costs plummeting → vastly increased data volume
  - http://www.genome.gov/sequencingcosts/

Future

- Belle-II
  - 250PB of experimental data in first 5 years of operation
- Square Kilometer Array (SKA)
  - ~2800 receivers in telescope array
  - 2 petabytes/sec to central correlator
    - sending @ ~100 Gb/s to analysis centers

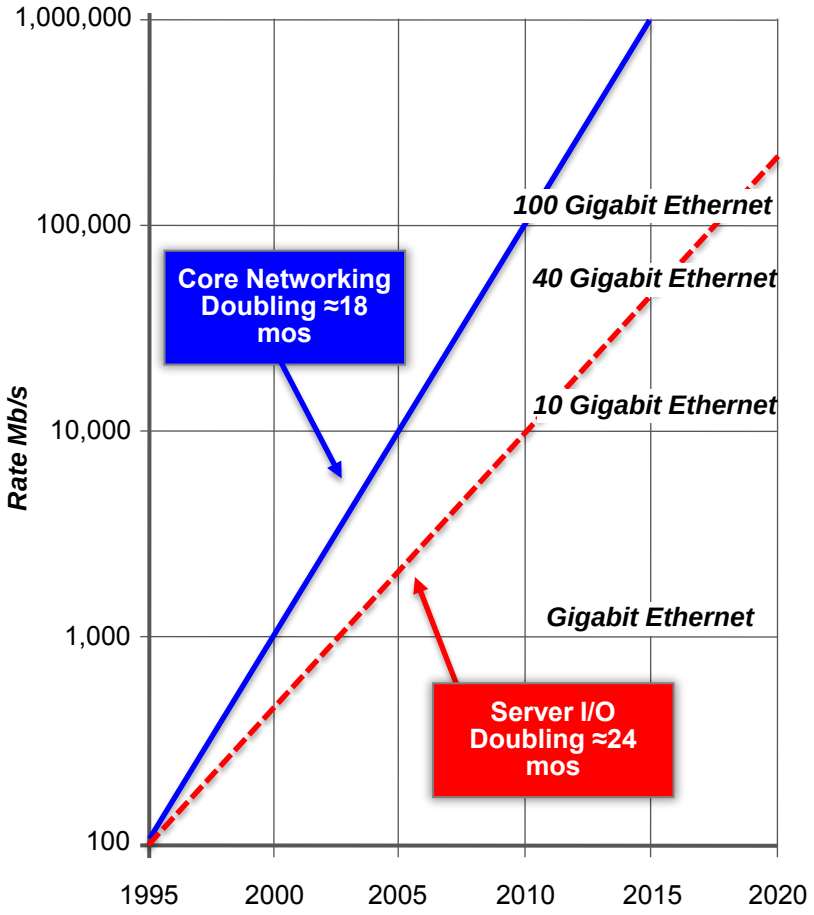
Source: http://www.ieee802.org/3/ad_hoc/bwa/public/dec11/dart_01_1211.pdf (updated: interview Eli Dart, August 29, 2012)

Findings of IEEE 802.3 BWA Ad Hoc



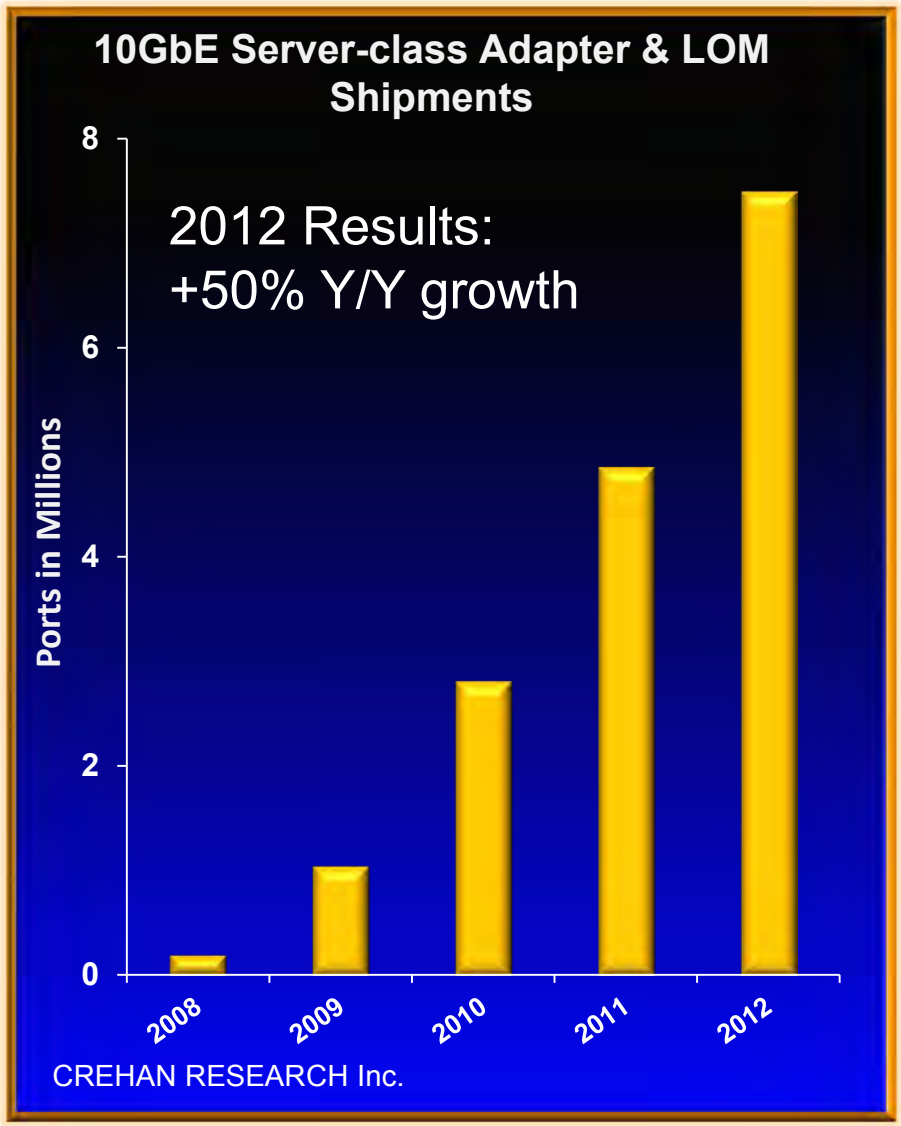
Source: http://www.ieee802.org/3/ad_hoc/bwa/BWA_Report.pdf

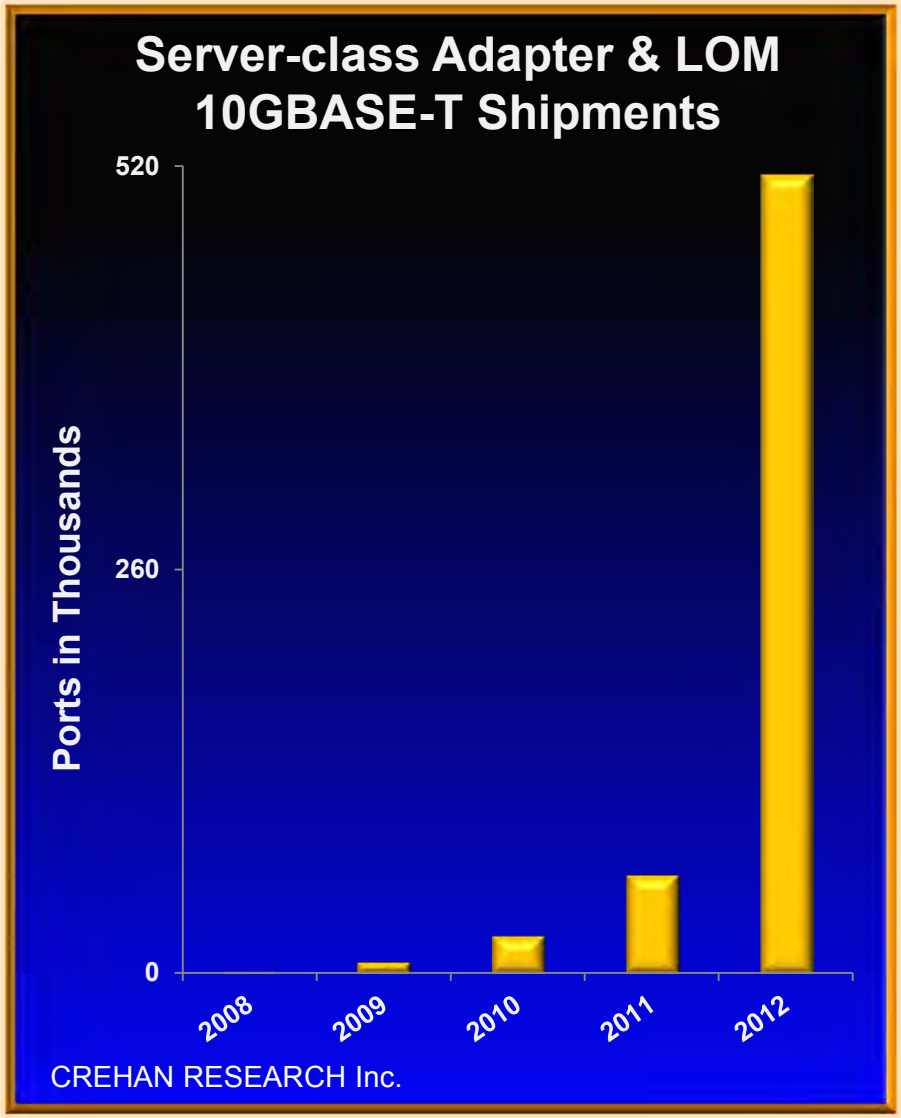
The Server Roadmap



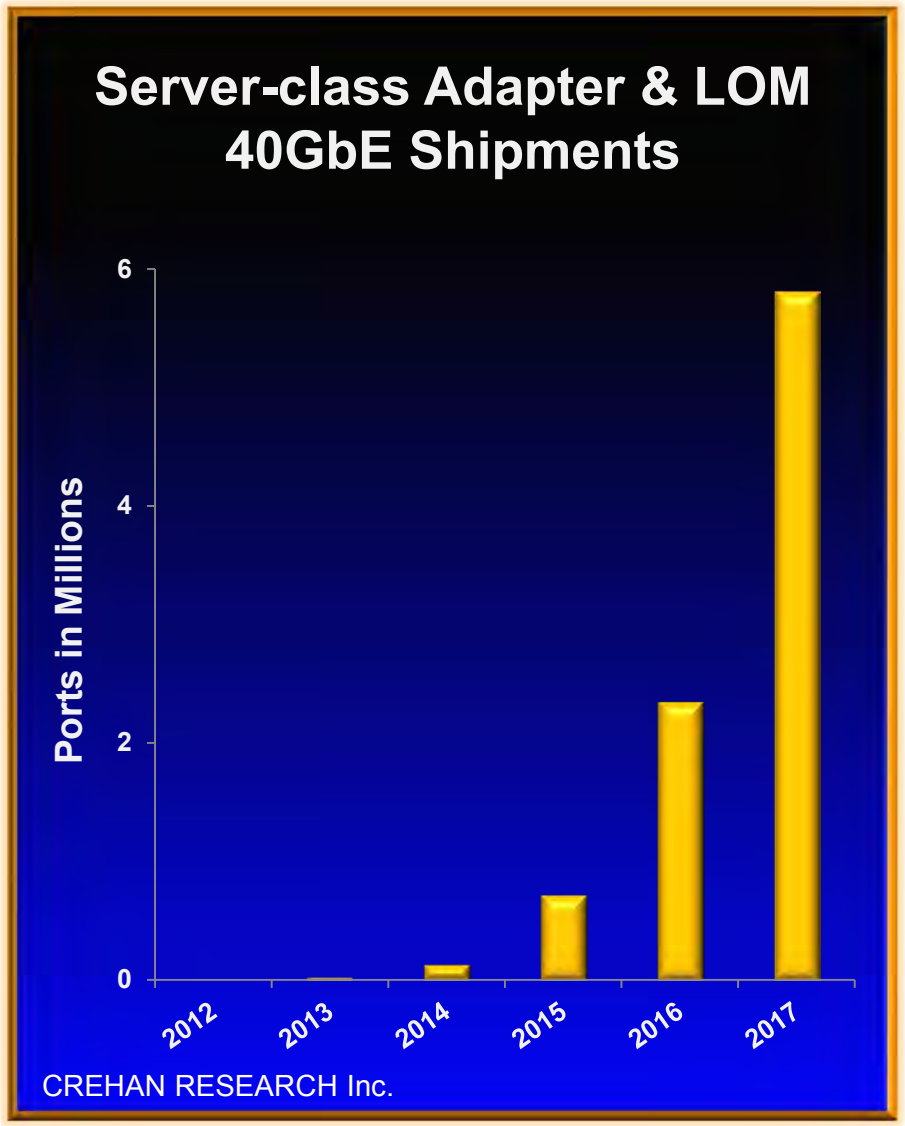
- Server Upgrade Path
  - 2014: 40 GbE
  - 2017: 100 GbE
- Blade Servers
  - 802.3ba: 10 GbE to 40 GbE
  - 802.3bj: 40 GbE to 100 GbE
- Other Future Server I/Os
  - 40GBASE-T
  - 100GbE over 20m MMF

10GbE Server Deployments

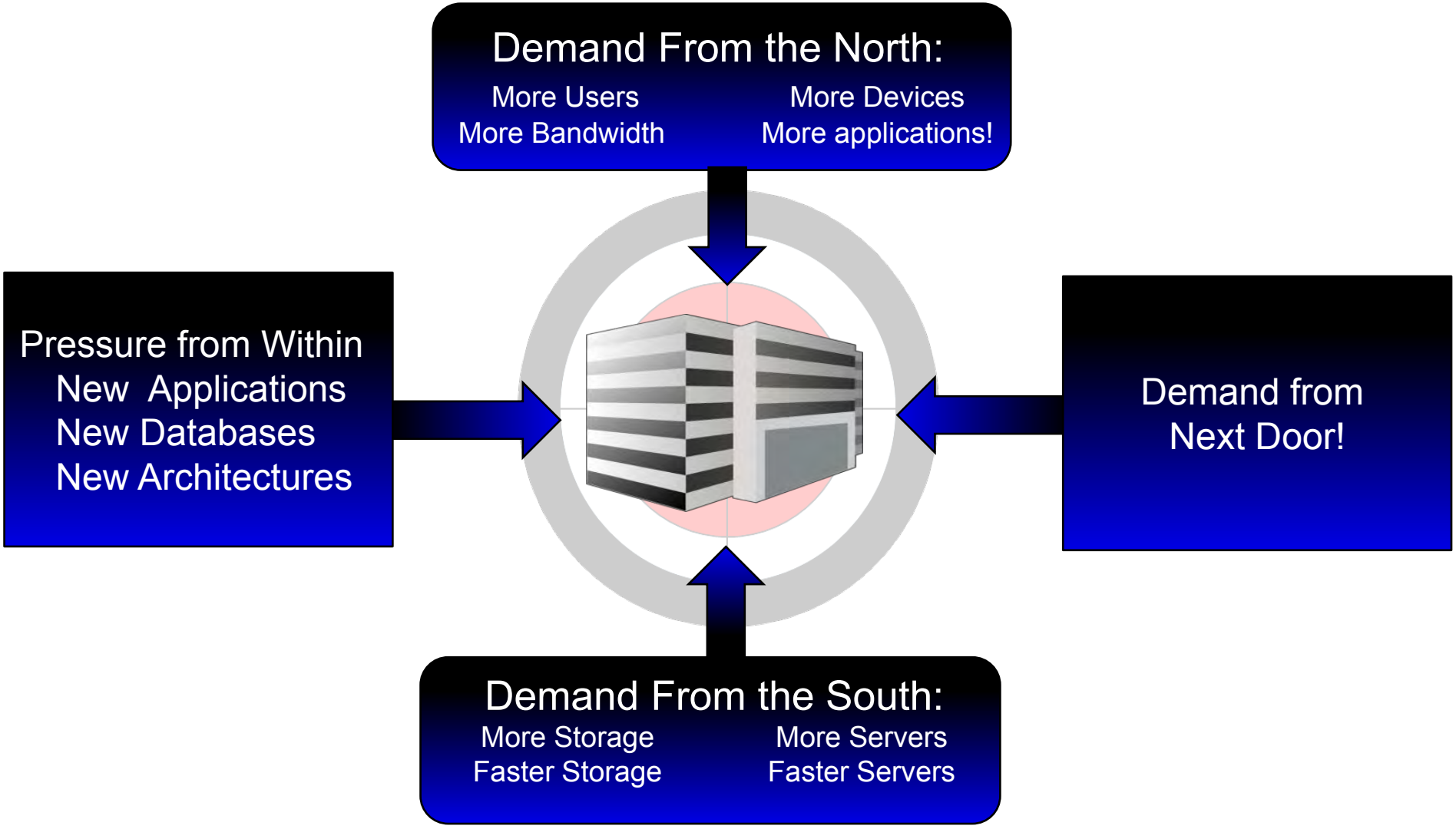




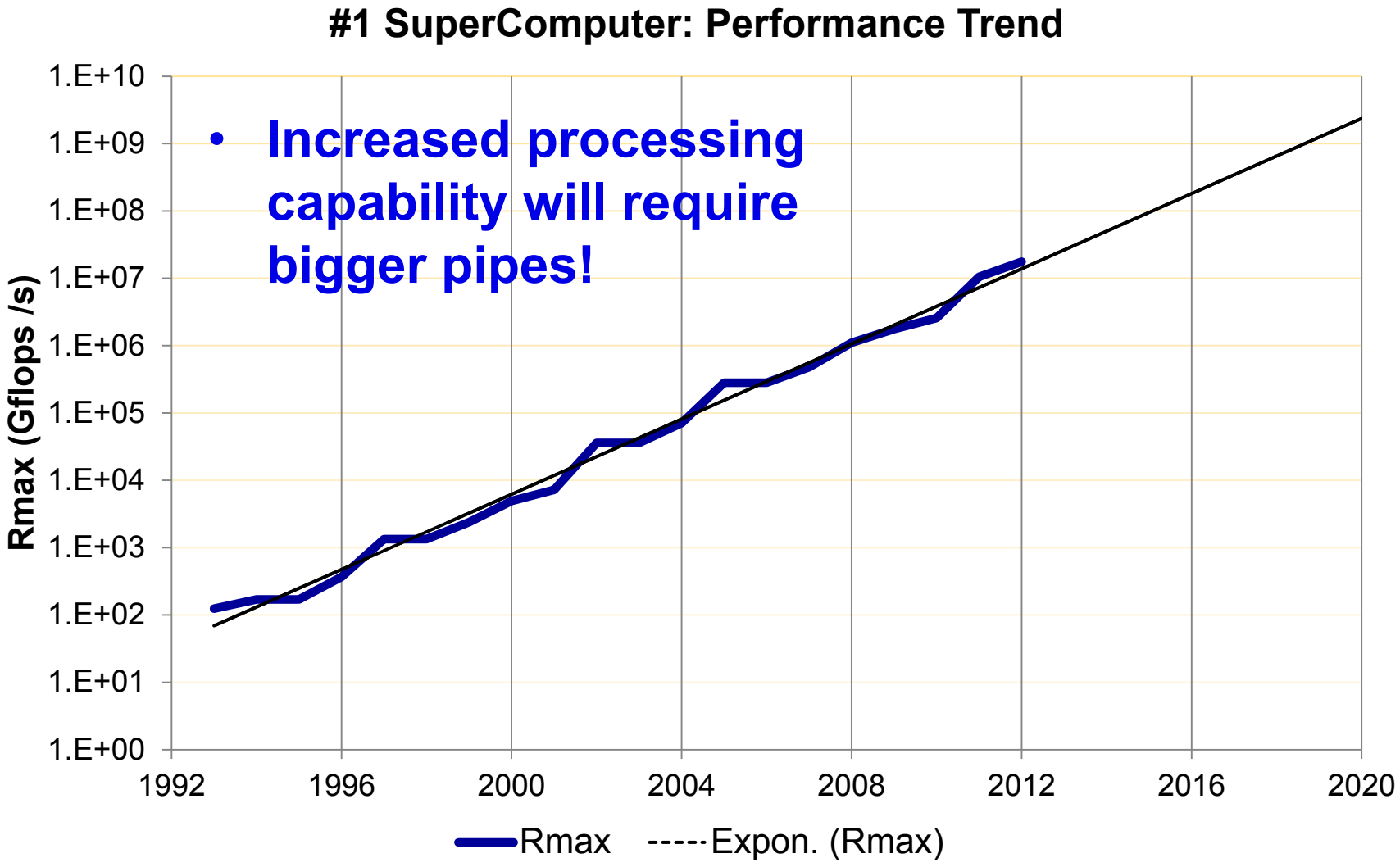
40GbE Server Deployment Forecast



Center of the Storm – Data Centers

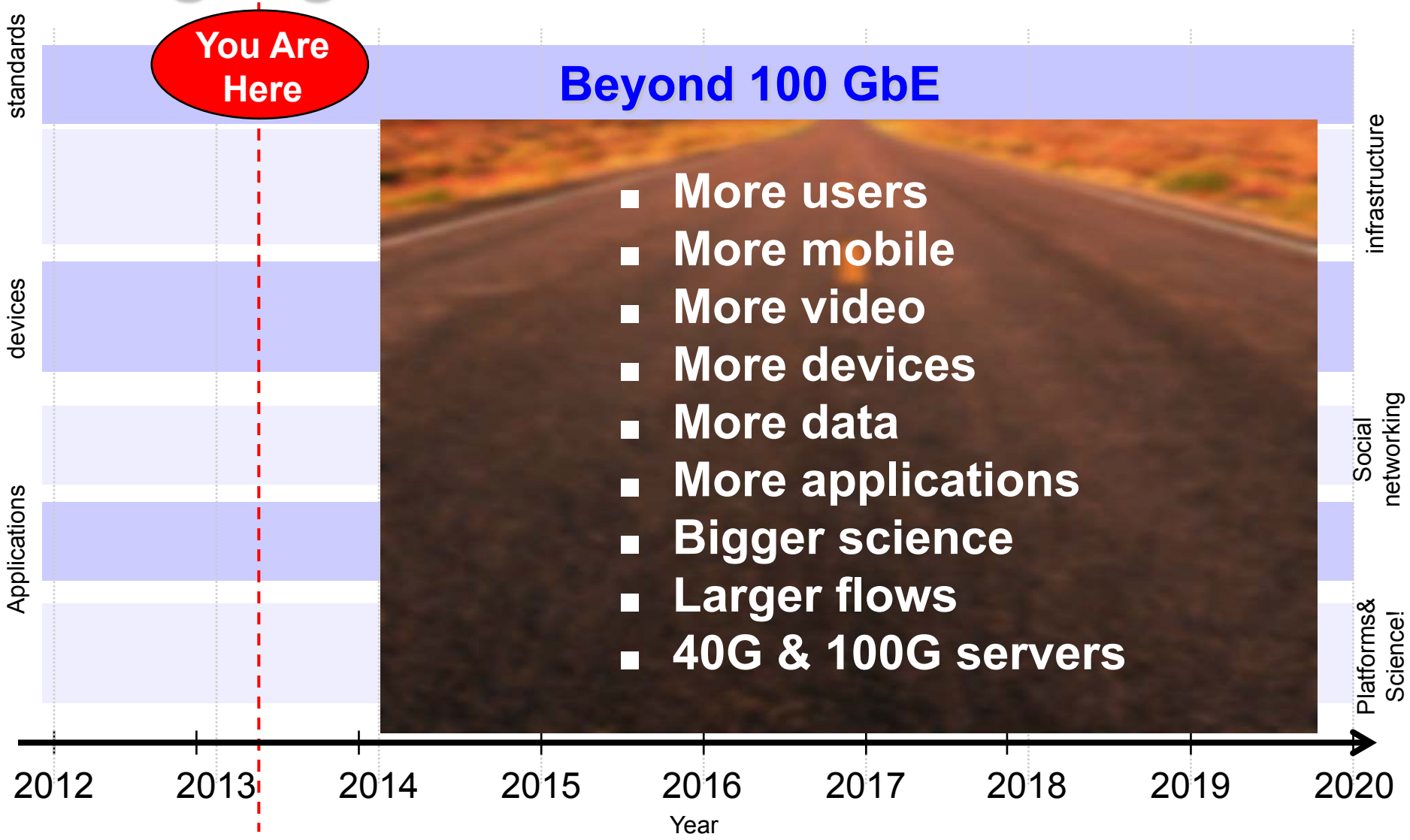


High Performance Computing

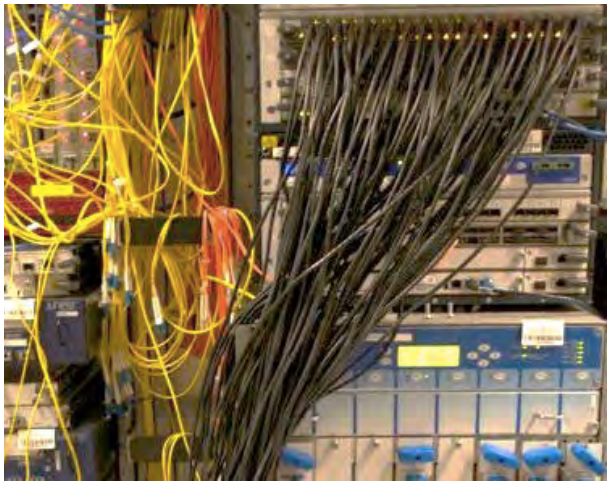


- **Increased processing capability will require bigger pipes!**

Highlights since IEEE P802.3ba



Link Aggregation



Courtesy, David Ofelt, Juniper.

For Illustrative Purposes
- **Example of an "80xN" LAG using 80 links**
- **Uses 25% of front panel for single "LAG"**
- **OpEx Challenges**

- Problem: Need to scale the Network (density & cost)
- Temporary Solution: Link Aggregation
- Pros: Addresses bandwidth requirements between releases of faster links
- Cons:
  - Non-deterministic performance
  - Fastest flow limited to individual link speed
  - Exponential bandwidth growth implies:
    - Exponential growth in number of links
    - Growth in operational & management issues
- Faster links address these issues ***and they will be LAGGed!***

	BW	100GbE	400GbE
2015	Terabit	10 x 100	3 x 400
2020	10 Terabit	100 x100	25 x 400

Section Summary

- Unrelenting bandwidth growth everywhere
  - End users / devices!
  - Applications!
  - Video!
- Carrier Transport – Beyond scope of this effort
- Ethernet Interconnect
  - Must scale to support bandwidth growth
  - Example – Data centers
    - Above
    - Below – server bandwidth growth (10GbE> 40GbE>100GbE)
    - East / West – Internal / External

Enabling bandwidth
Enables applications
Drives bandwidth

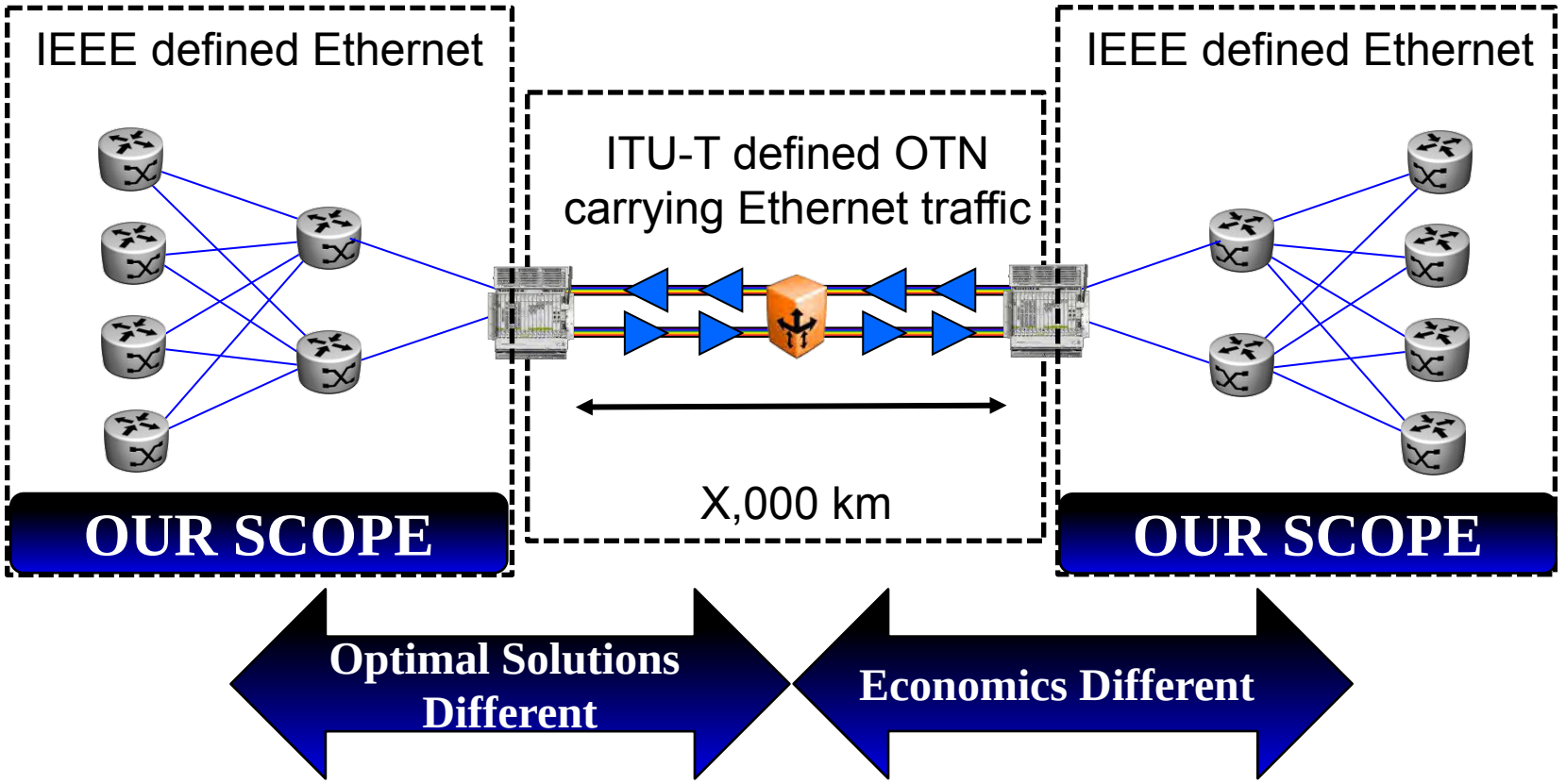
Presented by

IEEE 802.3 Working Group
Orlando, FL, USA
March 19, 2013

BEYOND 100 GIGABIT ETHERNET TECHNICAL CHALLENGES

What Are We Talking About?

- **At highest rates Ethernet is becoming dominant traffic for both client-side and line-side (OTN)**
- **Interdependent problems, but not interchangeable solutions**



CMOS Roadmap

- CMOS IC features have shrunk by ~2x since 100Gb/s MAC/PCS was defined in 802.3ba
- CMOS International Technology Roadmap for Semiconductors, 2011 Revision Overview:

- ITRS Sponsoring Industry Associations (IAs): European Semiconductor IA, Japan Electronics and Information Technology Association, Korea Semiconductor IA, Taiwan Semiconductor IA, (US) Semiconductor IA

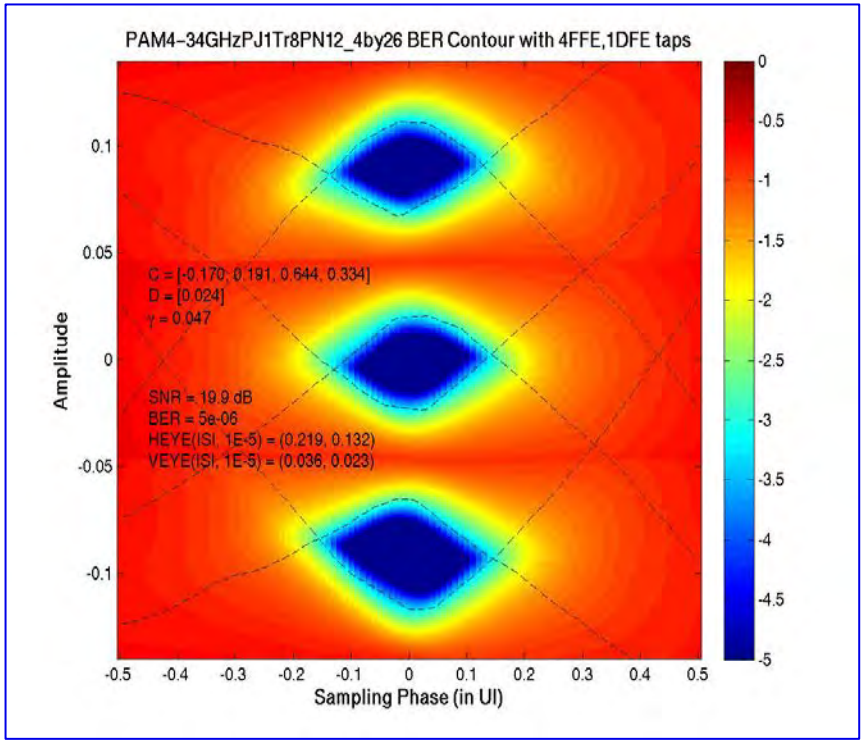
Technology Building Blocks (1 of 2)

Electrical Signaling Development

- 25 Gb/s
  - IEEE P802.3bj (NRZ / PAM-4)
  - IEEE P802.3bm
  - OIF CEI-28G
  - 32G Fibre Channel
- OIF CEI-56G

x16 Connector

Courtesy, Tom Palkert, Molex



Courtesy, Ali Ghiasi, Broadcom

Modulation

Technology Building Blocks (2 of 2)

Courtesy, Atsushi Takai, Oclaro

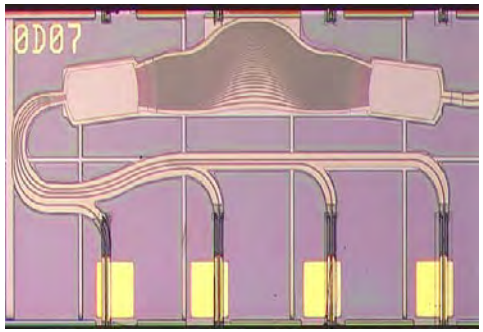
x16 Optical Connector & Cable

Courtesy, US Conec

GaAs Photonics

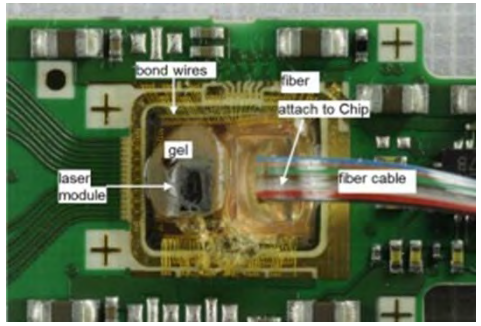
Courtesy, Finisar Corp.

InP Photonics



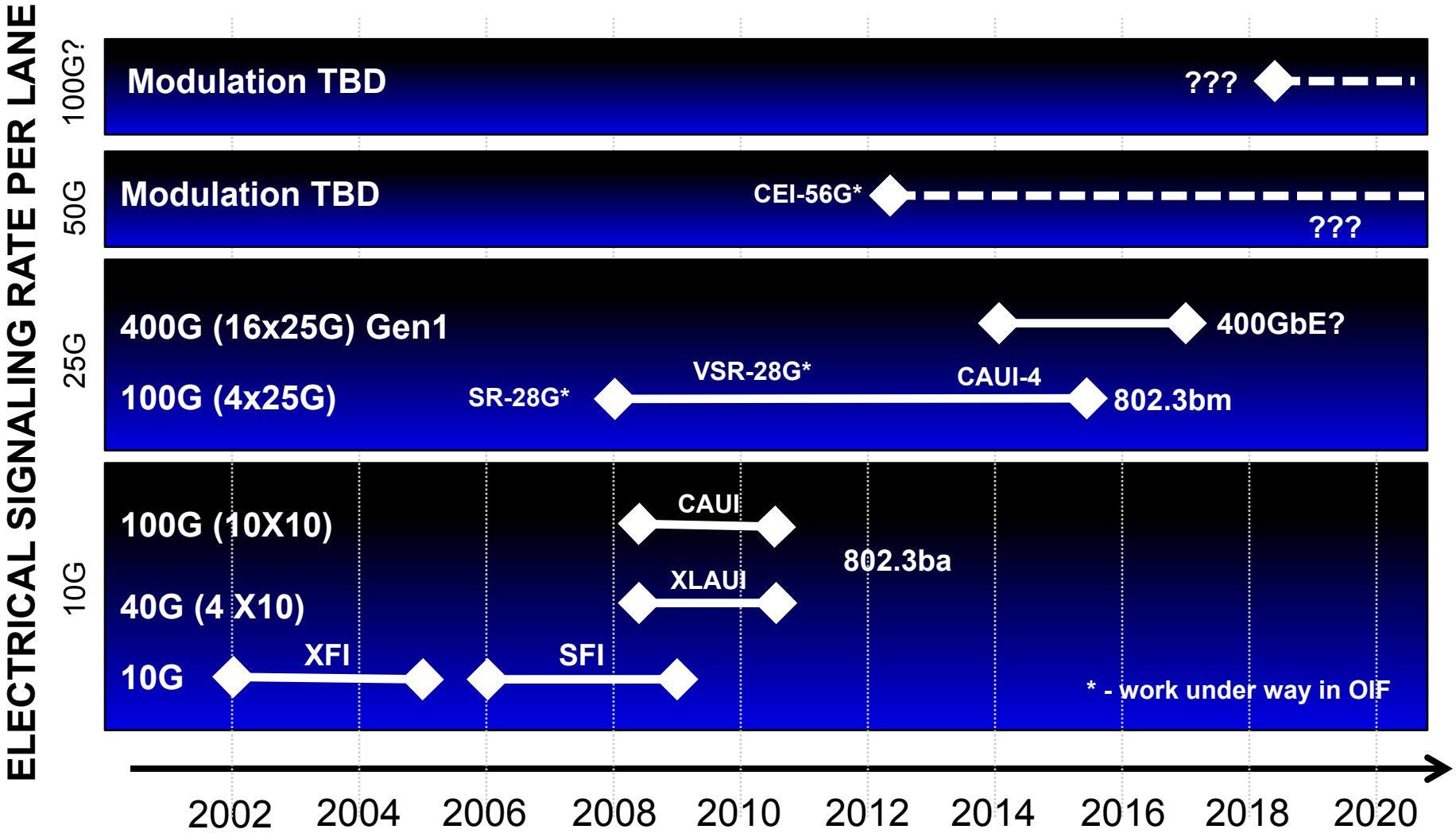
Courtesy, CyOptics Inc.

Silicon Photonics

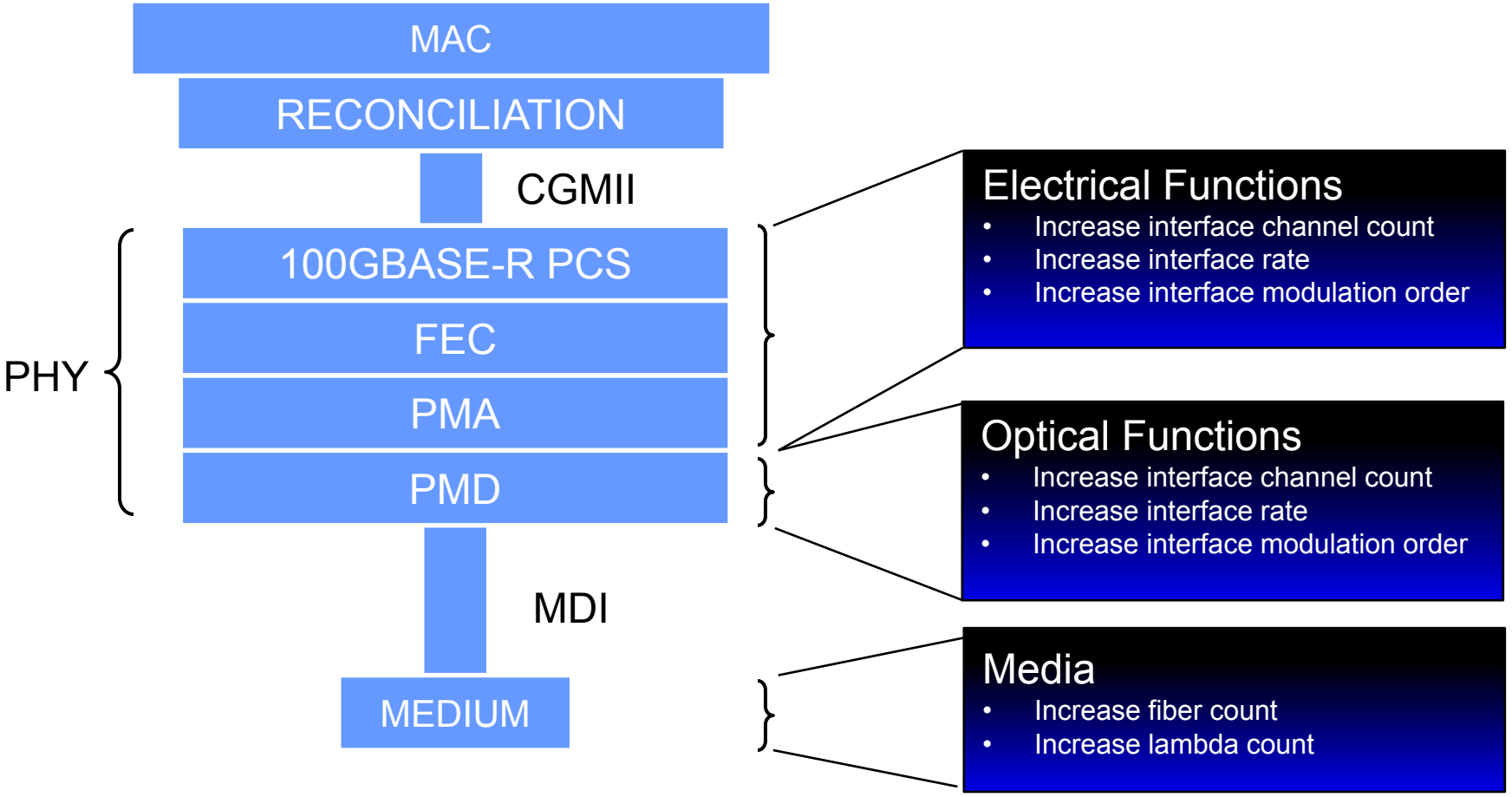


Courtesy, Tom Palkert, Luxtera

Module Electrical Interface Industry Development Efforts



An Ethernet Overview of the Problem

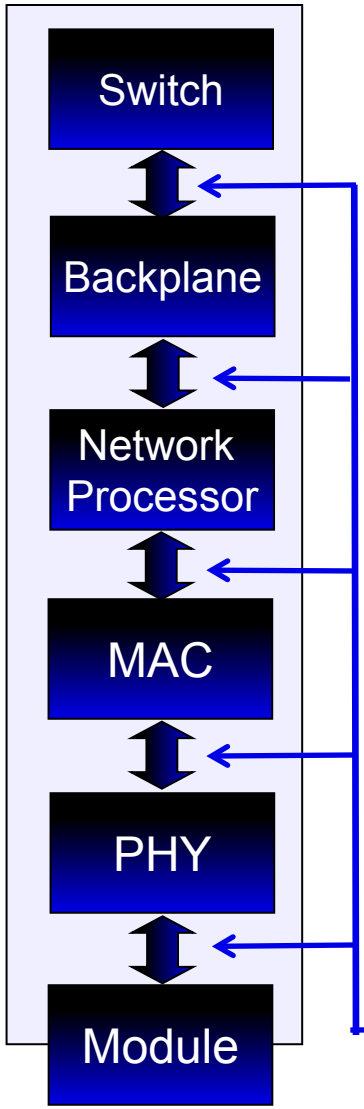


MAC Technical Feasibility

400 Gb/s and 1 Tb/s first generation implementation, with a separate MAC/PCS device (1 Tb/s interface widths are very challenging)

400 Gb/s and 1 Tb/s MAC options					
MAC Rate	Node	Technology	Bus Width	Clock Rate	Notes
100 Gb/s	45, 40nm	ASIC	160 bits	644 MHz	
100 Gb/s	45, 40nm	FPGA	512	195 MHz	
400 Gb/s	28, 20nm	ASIC	400	1 GHz	
400 Gb/s	28, 20nm	FPGA	1024	400 MHz	
1 Tb/s	28, 20nm	ASIC	1024	1 GHz	
1 Tb/s	28, 20nm	FPGA	2560	400 MHz	Challenging

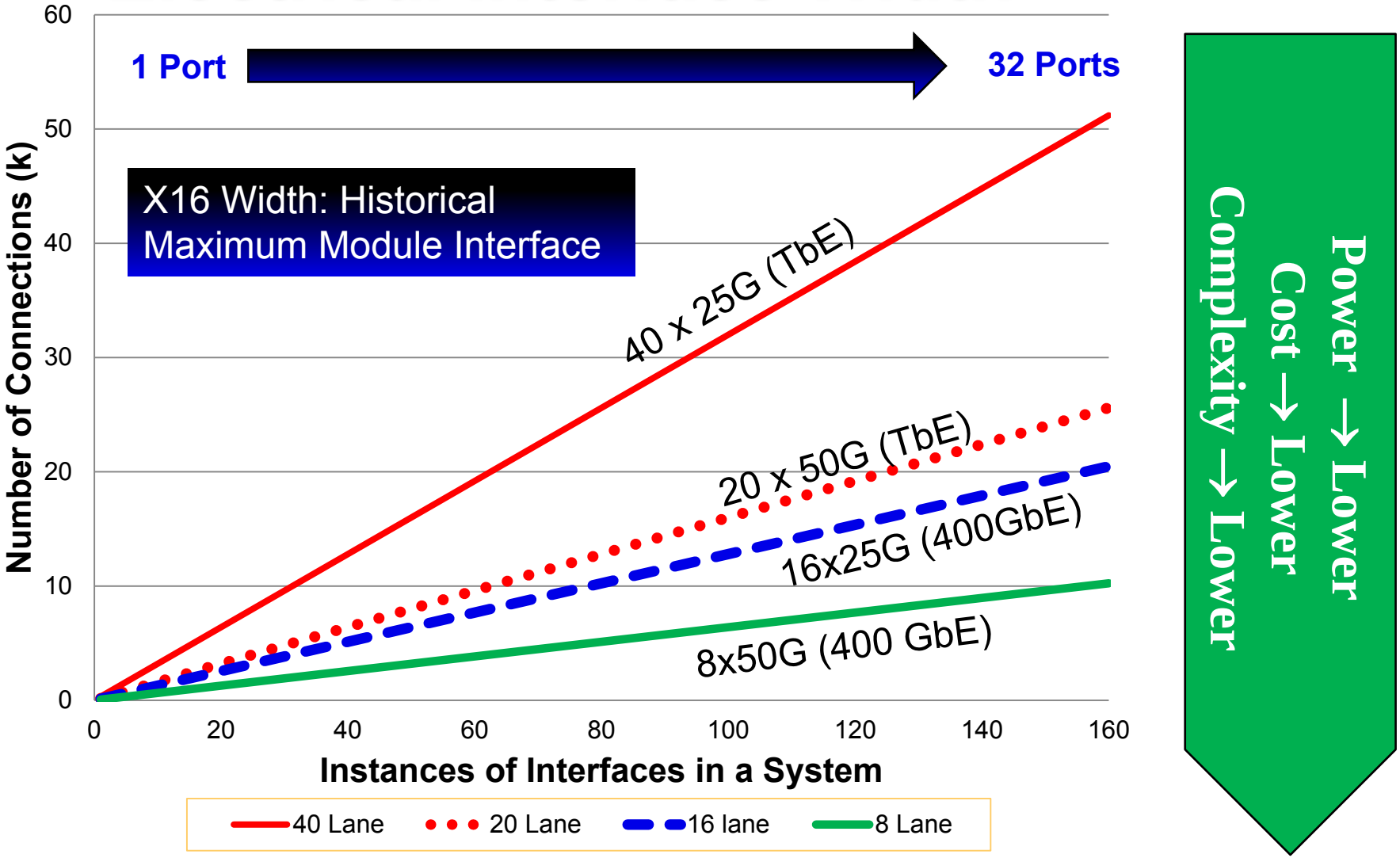
A System Perspective



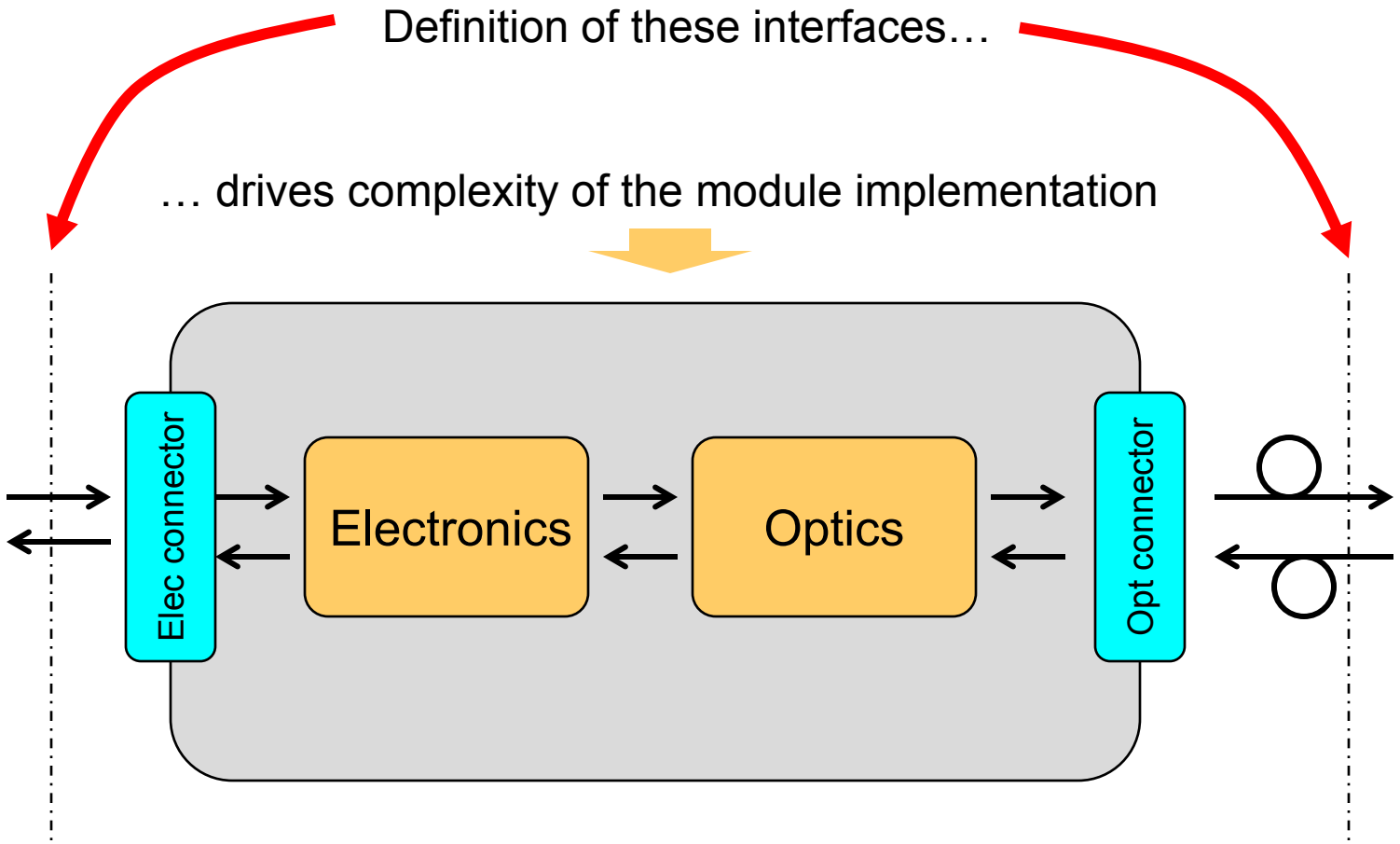
- Interface width is not just a module consideration, but is an overall is a system issue
- Wider interfaces > more pins
- More pins > More traces to route
  - More power
  - More cost
  - More complexity

Sample architecture highlights all places that might be needed to support interface

The Impact of the Electrical Interface Width

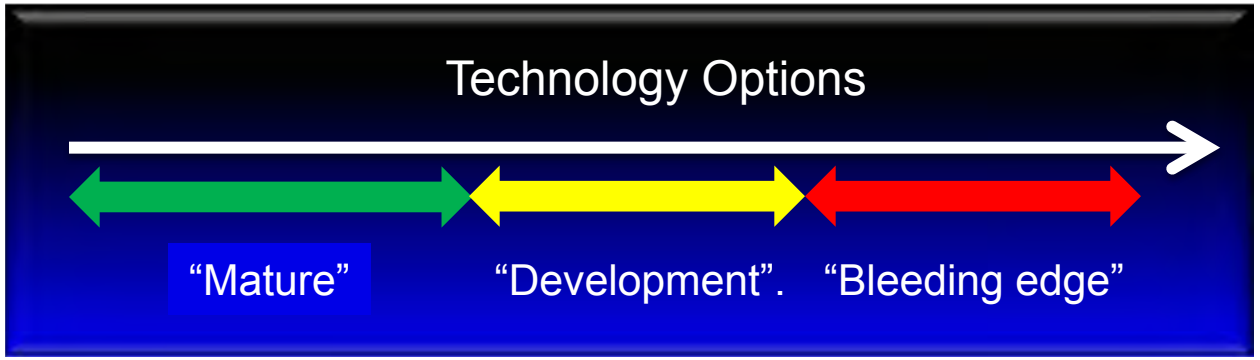


Example: Anatomy of an Optical Module Implementation

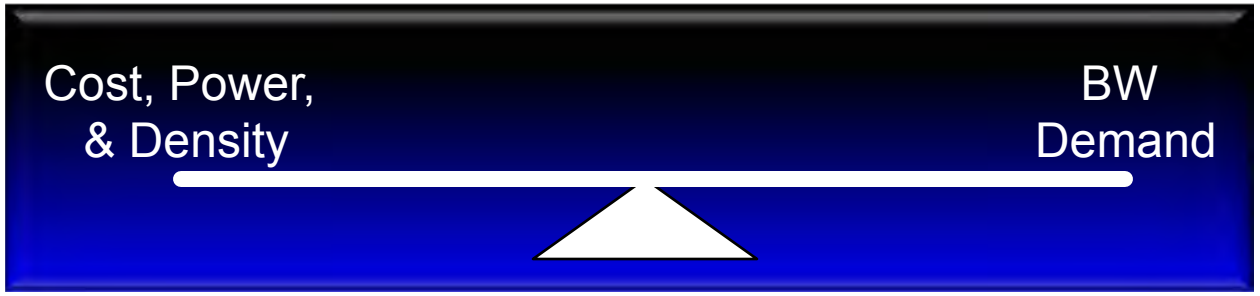


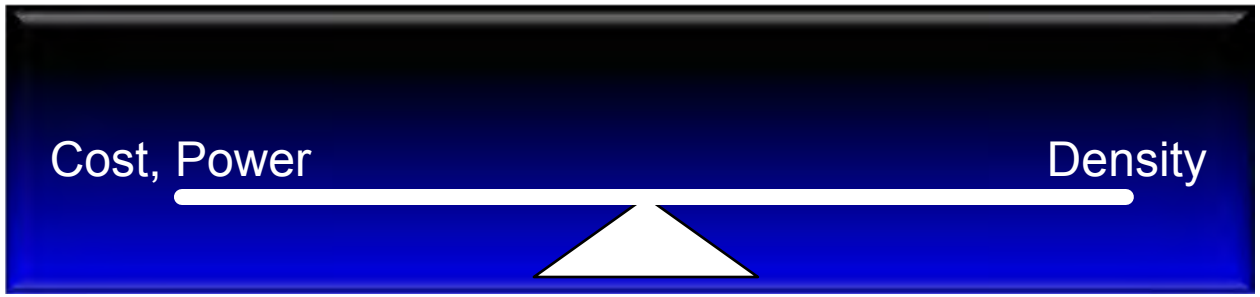
Matching Needs with Capabilities

It's all going to change with time

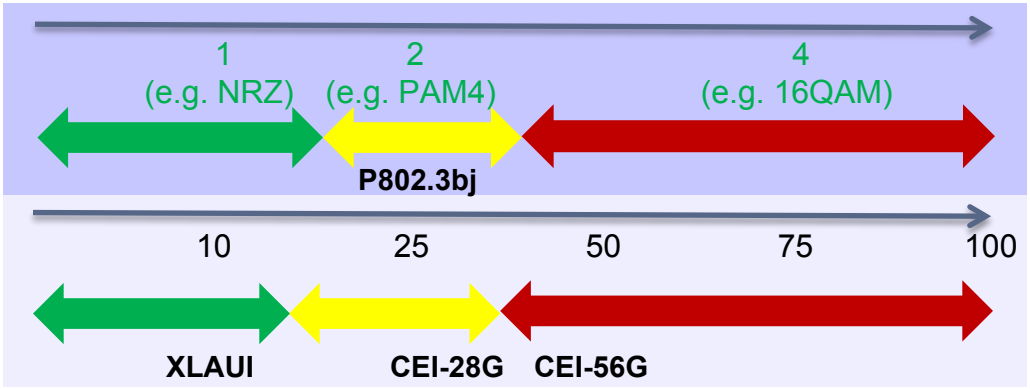


The never ending balancing acts!



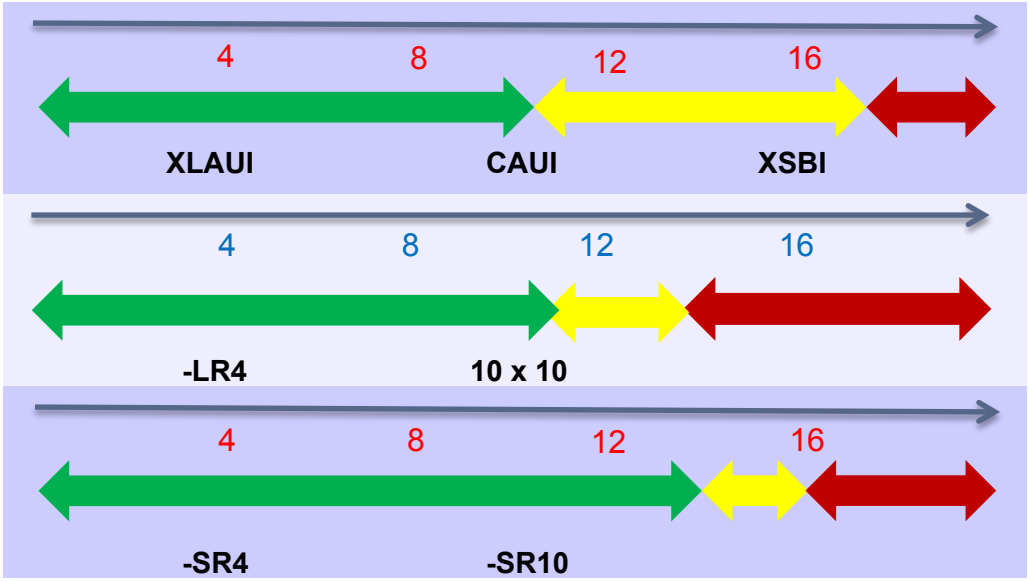


Potential Technology Axes for increasing Gbit/s in the Optical and Electrical Domains



Modulation
(i.e. Bits per Symbol)

Signaling Rate
(i.e. Gbaud)

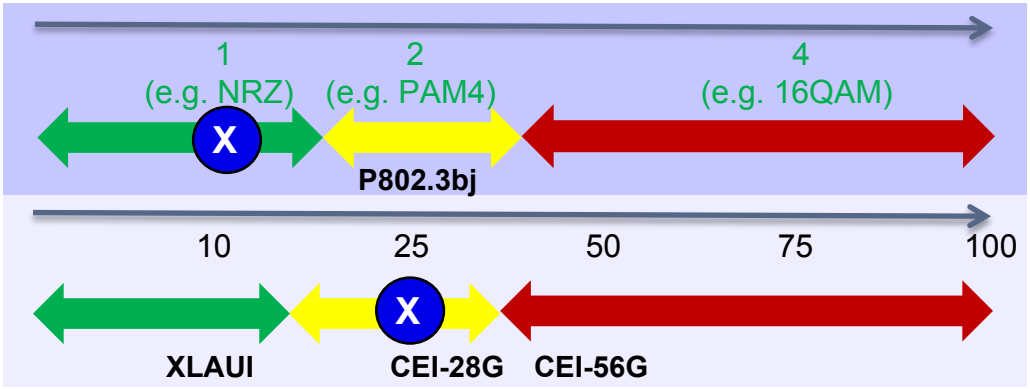


Space Division Multiplexing
(i.e. Multiple Electrical Channels)

Wavelength Division
Multiplexing (i.e. λs)

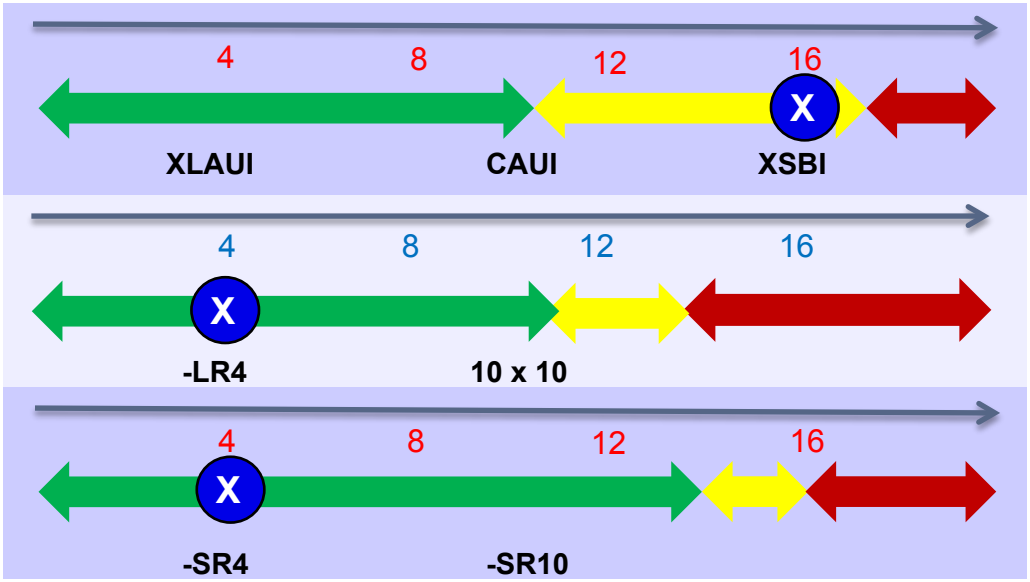
Space Division Multiplexing
(i.e. Multiple Optical Fibers)

Example: Finding a Path to 400Gbit "Gen 1"



Modulation
(i.e. Bits per Symbol)

Signaling Rate
(i.e. Gbaud)

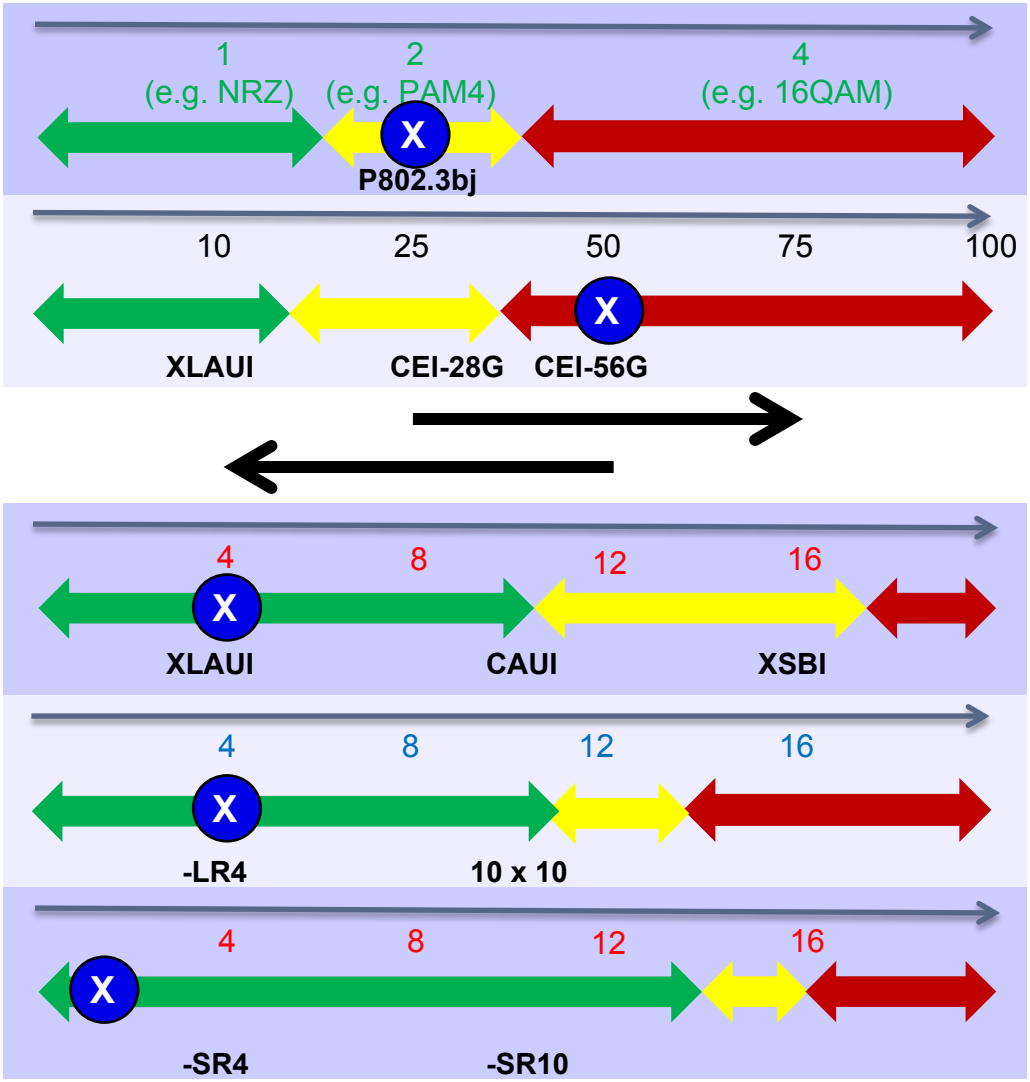


Space Division Multiplexing
(i.e. Multiple Electrical Channels)

Wavelength Division
Multiplexing (i.e. λs)

Space Division Multiplexing
(i.e. Multiple Optical Fibers)

Example: Finding a Path to 400Gbit "Future"



Modulation
(i.e. Bits per Symbol)

Signaling Rate
(i.e. Gbaud)

Space Division Multiplexing
(i.e. Multiple Electrical Channels)

Wavelength Division
Multiplexing (i.e. λs)

Space Division Multiplexing
(i.e. Multiple Optical Fibers)

Summary

- **Time is not on our side:**
  - **2015 Capacity Requirements: 10x 2010: 1 Terabit**
  - **2020 Capacity Requirements: 100 x 2010: 10 Terabit**
- **Not a matter of "can it be done" but "can it be done at right cost!"**
  - **Power, cost, density**
- **Technology for 400 Gigabit Ethernet**
  - **Leverage 100GbE building blocks**
  - **Plausible implementations for today and next generation**
  - **Fits with dense 100GbE system roadmap**

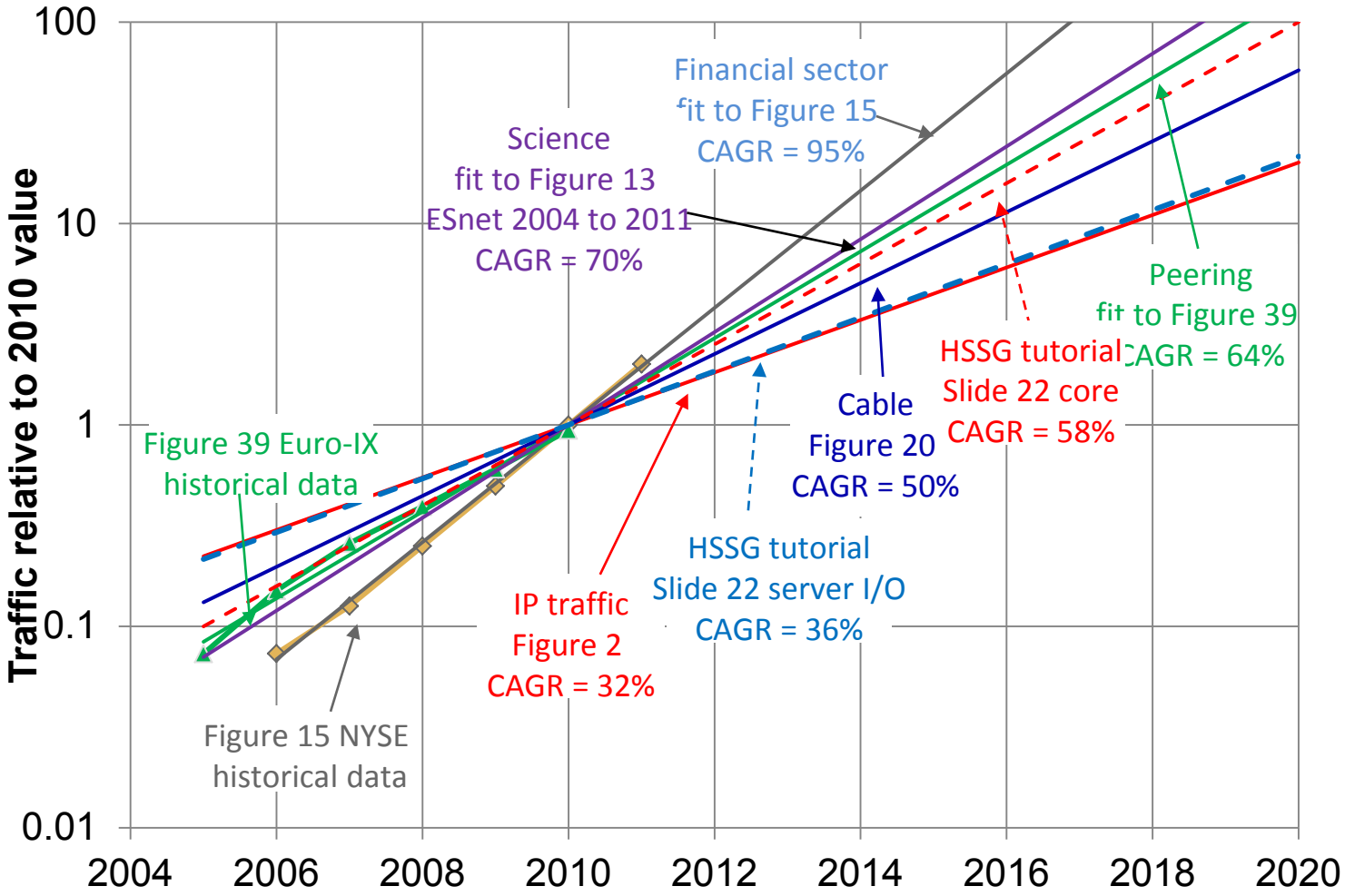
- **400Gb/s Ethernet: A cost effective solution!**

Presented by
John D'Ambrosia, Dell

IEEE 802.3 Working Group
Orlando, FL, USA
March 19, 2013

400 GIGABIT ETHERNET - WHY NOW?

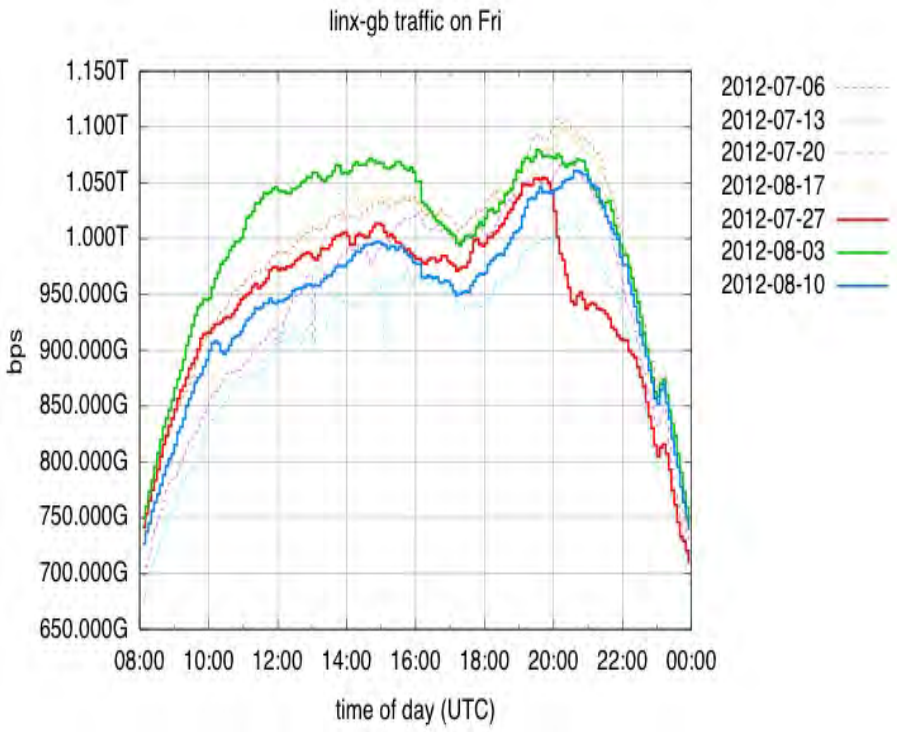
Findings of IEEE 802.3 BWA Ad Hoc



Source: http://www.ieee802.org/3/ad_hoc/bwa/BWA_Report.pdf

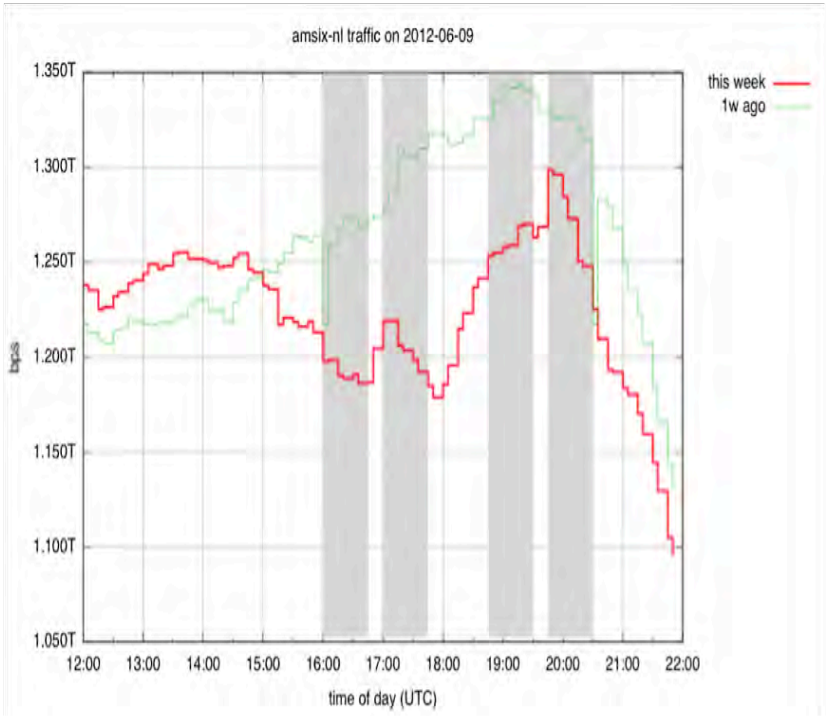
The Future is Here…

2012 Summer Olympics



Source: https://labs.ripe.net/Members/fergalc/internet-traffic-during-olympics-2012

After First Round of Euro 2012 Matches



Source: https://labs.ripe.net/Members/fergalc/internet-traffic-after-first-round-of-euro-2012-matches/AMSIXNL.png

Thanks to Bijal Sanghani, Euro-IX.

The Need for 400 Gb/s Ethernet

- **Traffic is growing everywhere**
  - **More Internet users**
  - **More ways to access the internet faster**
  - **Higher bandwidth content**
  - **New applications enabled**
  - **And it goes on**
- **IEEE 802.3 BWA Forecast**
  - **2015 Capacity, 10x requirements of 2010 - Terabit**
  - **2020 Capacity 100x requirements of 2010 – 10 Terabit**
- **Time is not on our side…**

Summary

- Bandwidth – exponential growth continues!
- New bandwidth generating applications constantly being introduced
- Higher Speed @ lower cost per bit needed by Ethernet Interconnect
- Past efforts took 3 to 4 years
  - 10 Gigabit Ethernet
  - Ethernet First Mile
  - 40 Gigabit and 100 Gigabit Ethernet
- We need to begin the process to study the problem

- "The Next Speed is not the Last Speed"

Contributors

- Pete Anslow, Ciena
- Chris Cole, Finisar
- Kai Cui, Huawei
- John D'Ambrosia, Dell
- Mark Gustlin, Xilinx
- Mike Li, Altera
- Jeff Maki, Juniper
- Andy Moorwood, Infinera
- Gary Nicholl, Cisco
- Mark Nowell, Cisco
- David Ofelt, Juniper
- Brian Teipen, ADVA
- Jim Theodoras, ADVA
- Steve Trowbridge, Alcatel-Lucent
- IEEE 802.3 Higher Speed Ethernet Consensus Ad Hoc

Supporters

- Ghani Abbas, Ericsson
- Pete Anslow, Ciena
- Thananya Baldwin, Ixia
- Stephen Bates, PMC-Sierra
- Mike Bennett, LBNL
- Ralf-Peter Braun, Deutsche Telecom
- Derek Cassidy, BT
- Chris Cole, Finisar
- Kai Cui, Huawei
- John D'Ambrosia, Dell
- Arash Farhood, Cortina
- Harry Forbes, Nexans
- Mark Gustlin, Xilinx
- Hiroshi Hamano, Fujitsu Labs
- Brian Holden, Kandou Bus, S.A.
- Xi Huang, Huawei
- Hideki Isono, Fujitsu Optical Components
- Masashi Kono, Hitachi
- Jeff Maki, Juniper
- Tom McDermott, Fujitsu Network Communications
- Andy Moorwood, Infinera
- Gary Nicholl, Cisco
- Mark Nowell, Cisco
- David Ofelt, Juniper
- Jerry Pepper, Ixia
- Rick Rabinovich, Alcatel-Lucent
- Sam Sambasivan, AT&T
- K. Seto, Hitachi Cable
- Song Shang, Semtech
- Kapil Shrikhande, Dell
- Scott Sommers, Molex
- Xiaolu Song, Huawei
- Ted Sprague, Infinera
- Peter Stassar, Huawei
- William Szeto, Xtera Communications
- Tomoo Takahara, Fujitsu Labs
- Toshiki Tanaka, Fujitsu Labs
- Katsuhisa, Tawa, Sumitomo Electric
- Brian Teipen, ADVA
- Jim Theodoras, ADVA
- Steve Trowbridge, Alcatel-Lucent
- Chengbin, Wu, ZTE
- Helen Xuyu, Huawei
- Wenyu Zhao, CATR

STRAW POLLS

Call-For-Interest

- **Should a Study Group be formed for “ 400 Gb/s Ethernet”?**

  Y: **N:** **A:**

Participation

- **I would participate in the "400 Gb/s Ethernet" Study Group in IEEE 802.3.**

  Tally: xx

- **My company would support participation in the "400 Gb/s Ethernet" Study Group in IEEE 802.3**

  Tally: xx

Future Work

- **Ask 802.3 on Thursday**
  - **Form 400 Gb/s Ethernet SG**
  - **Pending approval of 400 Gb/s Ethernet SG, Request that 802.3 WG request 802 EC HSE Consensus Ad Hoc be disbanded.**
- **If approved, on Friday**
  - **Request 802 EC informed of 400 Gb/s Ethernet SG**
  - **Pending approval of 400 Gb/s Ethernet SG, Request that 802 EC HSE Consensus Ad Hoc be disbanded.**
  - **First 400 Gb/s Ethernet SG meeting, week of May 2013 IEEE 802.3 Interim.**

THANK YOU!