The Findings of the IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc

IEEE 802 Plenary

San Diego, CA, USA

July 16, 2012

Presenters

- John D'Ambrosia, Dell, IEEE 802.3 BWA Chair
- Peter Anslow, Ciena, IEEE 802.3 BWA Editor
- Mark Nowell, Cisco
- Scott Kipp, Brocade
- Peter Stassar, Huawei

Agenda

- Introduction (John D'Ambrosia)
- Findings
  - Overview (Mark Nowell)
  - The Data Center (Scott Kipp)
  - Transport Networks (Peter Stassar)
- Summary (Peter Anslow)

Disclaimers

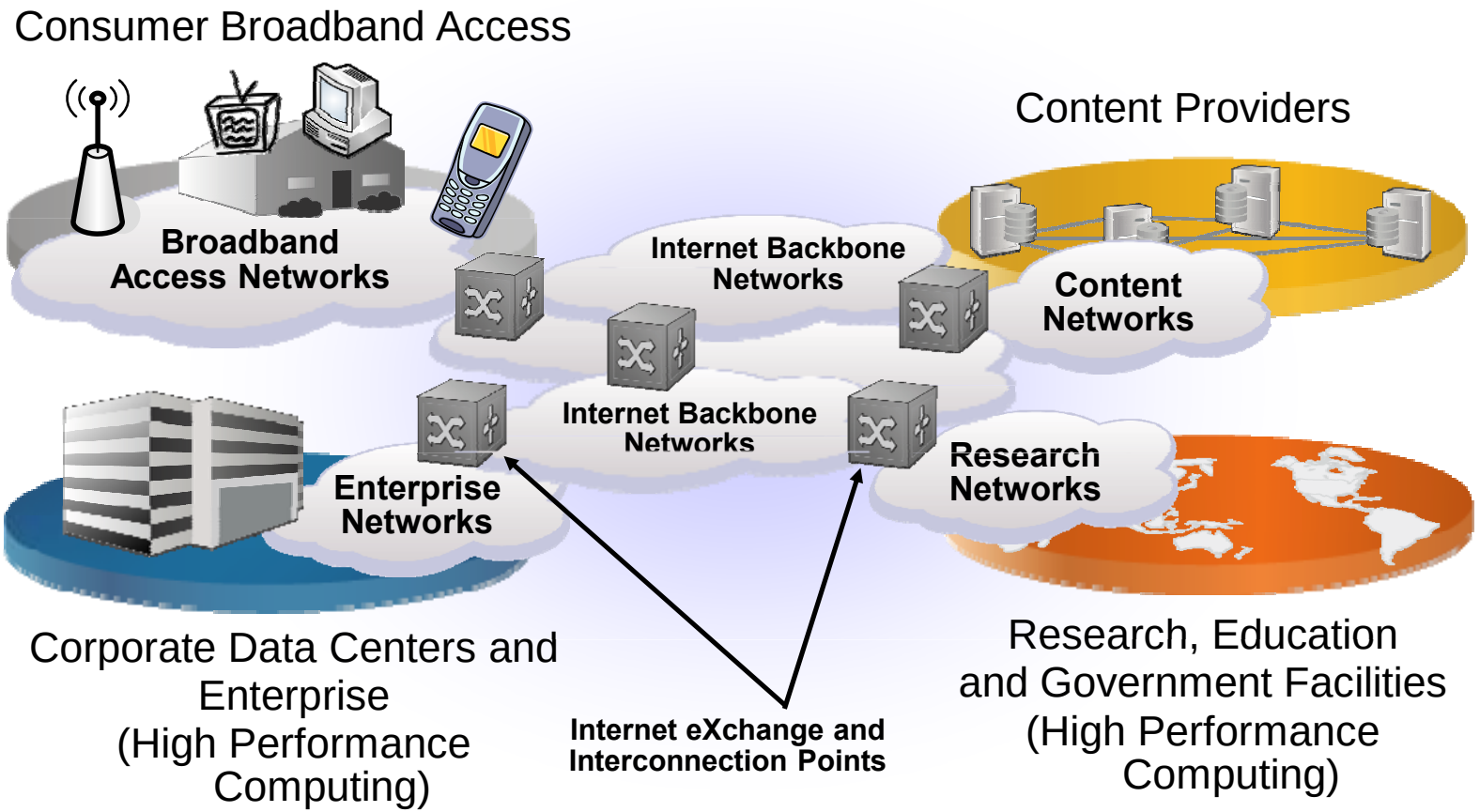
- This presentation is a supplement to the IEEE Industry Connections Ethernet Bandwidth Assessment D1.2, which is pending final approval (this week) by the IEEE 802.3 Working Group
- All contributed information is solely the perspective of the respective contributors.
- The views expressed in the Assessment solely represent the views of the IEEE 802.3 Working Group, and do not necessarily represent a position of the IEEE, the IEEE Standards Association, or IEEE 802.

INTRODUCTION

JOHN D'AMBROSIA, DELL

The 2006 HSSG Call-For-Interest

The Ethernet Ecosystem



Why Higher Speed Ethernet?

Fundamental bottlenecks are happening everywhere

Increased # of users + **Increased access rates and methods** + **Increased services** = **Bandwidth explosion everywhere**

As demonstrated by the number of ISPs: Comcast, AOL, YahooBB, NTT, Cox, EasyNet, Rogers, BT, ...

EFM, xDSL, WiMax, xPON, Cable, WiFi, 3G/4G...

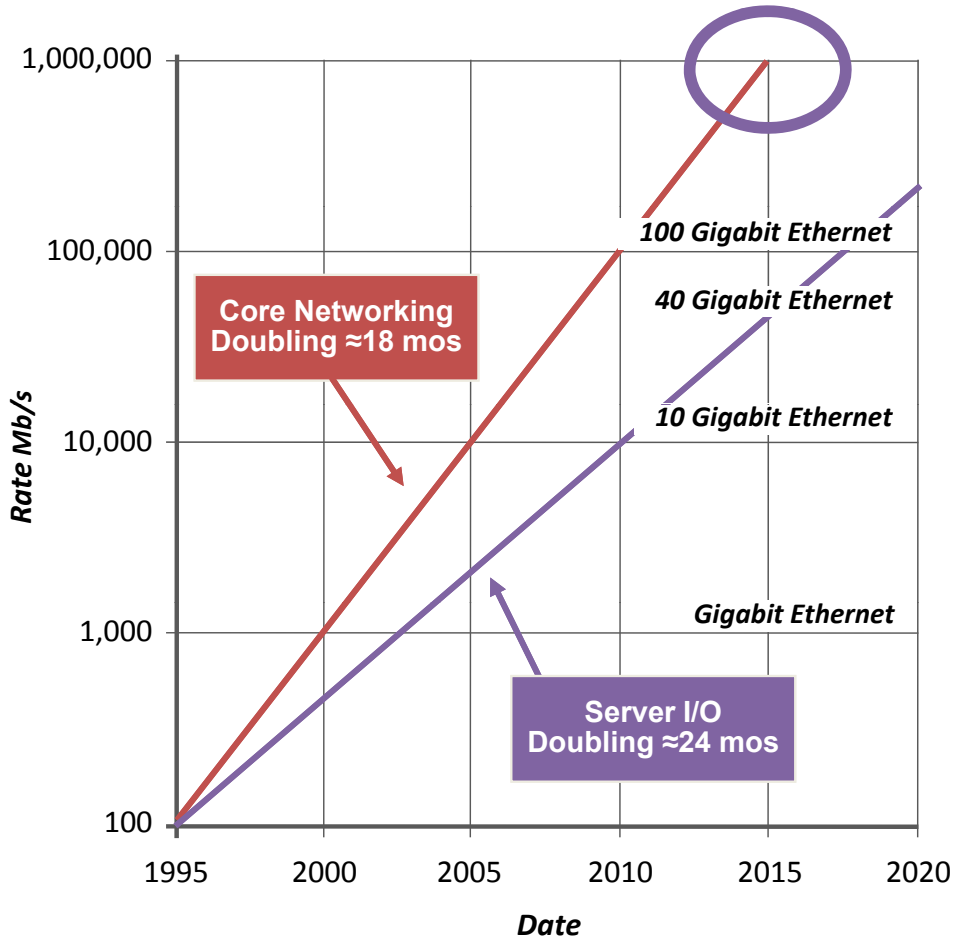
YouTube, BitTorrent, VOD, Facebook, Kazaa, Netflix, iTunes, 2nd life, Gaming...

Life after IEEE P802.3ba

- End-users through the prior HSSG: <u>The next speed of Ethernet must begin when 100GbE done!</u>
- HSSG Bandwidth Forecast for "Core Networking"
  - 2013: 400 Gb/s
  - 2015: 1 Tb/s
- Other bandwidth trends?

- 2011 Formation of:

  IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc



IEEE 802.3 BWA
Web & Reflector Information

- Charter and Scope
  - Evaluate Ethernet wireline bandwidth needs of the industry
  - Reference material for a future activity
  - The role of this ad hoc is to gather information, not make recommendations or create a CFI

- Webpage - http://www.ieee802.org/3/ad_hoc/bwa/index.html
- Reflector - http://www.ieee802.org/3/ad_hoc/bwa/reflector.html
- Public request for data - http://www.ieee802.org/3/ad_hoc/bwa/public/anslow_01a_0411.pdf

Summary of Data Submissions

- Scott Kipp, Brocade, “Data Center Bandwidth Scenarios”
  - http://www.ieee802.org/3/ad_hoc/bwa/public/may11/kipp_01_0511.pdf
- Andy Bach, NYSE Euronext, “Bandwidth Demand in the Financial Industry - The Growth Continues”
  - http://www.ieee802.org/3/ad_hoc/bwa/public/jun11/bach_01a_0611.pdf
- Kimball Brown, LightCounting, “Server Bandwidth Scenarios - Signposts for 40G/100G Server Connections”
  - http://www.ieee802.org/3/ad_hoc/bwa/public/jul11/brown_01a_0711.pdf
- Tom Cloonan, Arris, “Bandwidth Trends on the Internet... A Cable Data Vendor's Perspective”
  - http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/cloonan_01a_0911.pdf
- Scott Kipp, Brocade, “Storage Growth and Ethernet”
  - http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/kipp_01a_0911.pdf
- Mark Nowell, Cisco, “Cisco Visual Networking Index (VNI) Global IP Traffic Forecast Update; 2010 - 2015”
  - http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/nowell_01_0911.pdf
- Petar Pepeljugoski and Paul Coteus, IBM, “Bandwidth needs in HPC taking into account link redundancy”
  - http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/pepeljugoski_01_1111.pdf
- Huang Xi, Huawei, “Bandwidth Needs in Core and Aggregation nodes in the Optical Transport Network”
  - http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/huang_01_1111.pdf
- Henk Steenman, AMS-IX / Euro-IX, “The European Peering Scene”
  - http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/steenman_01_1111.pdf
- Lone Hansen, BSRIA, “Global Data Centres Presentation IEEE”
  - http://www.ieee802.org/3/ad_hoc/bwa/public/dec11/hansen_01_1211.pdf
- Eli Dart, ESnet, “Data Intensive Science Impact on Networks”
  - http://www.ieee802.org/3/ad_hoc/bwa/public/dec11/dart_01_1211.pdf

Assessment Limitations

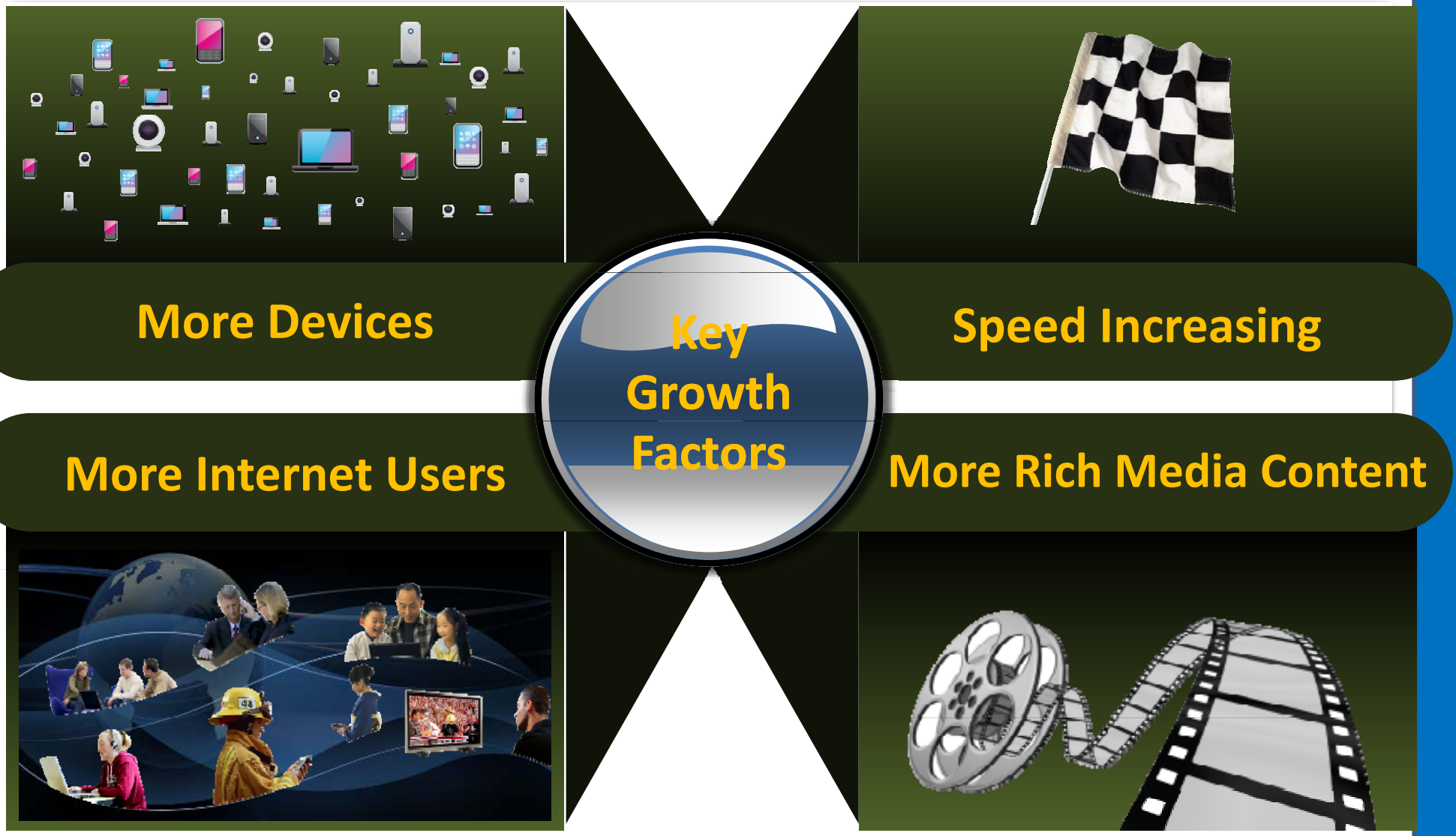
- Assessment Duration: 18 months maximum
  - Limited study time
  - Prevent data from becoming dated
  - Information provided snapshot at time of submission
- Past trends may not be an accurate predictor of the future
  - Emerging applications
  - Technology
  - Standardization Efforts
  - Will Ethernet cost per gigabit continue to decrease?
- Underlying assumptions
  - Market adoption
  - Continuation of applications that require increasing bandwidth

FINDINGS

OVERVIEW

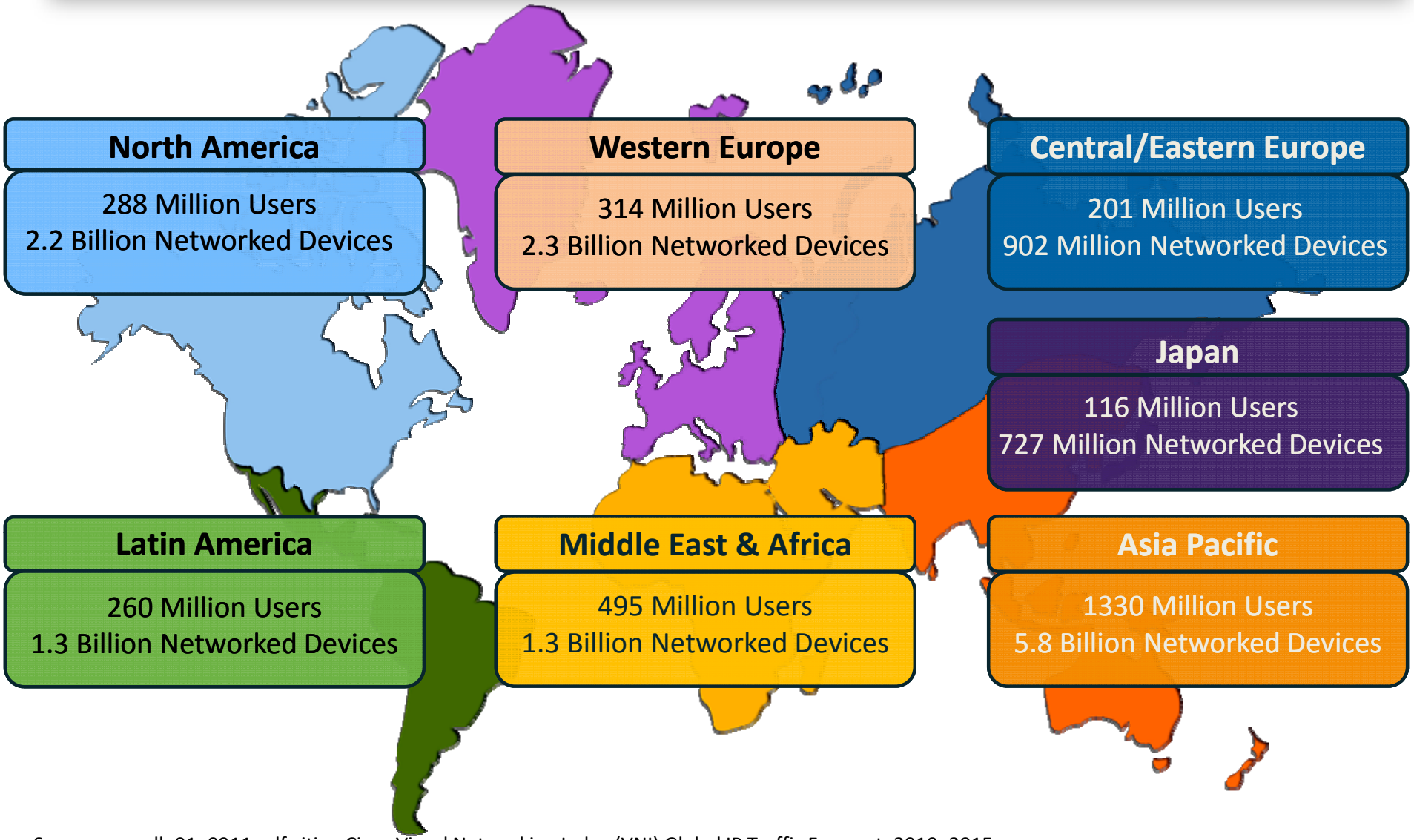
MARK NOWELL, CISCO

The Equation Remains the Same



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

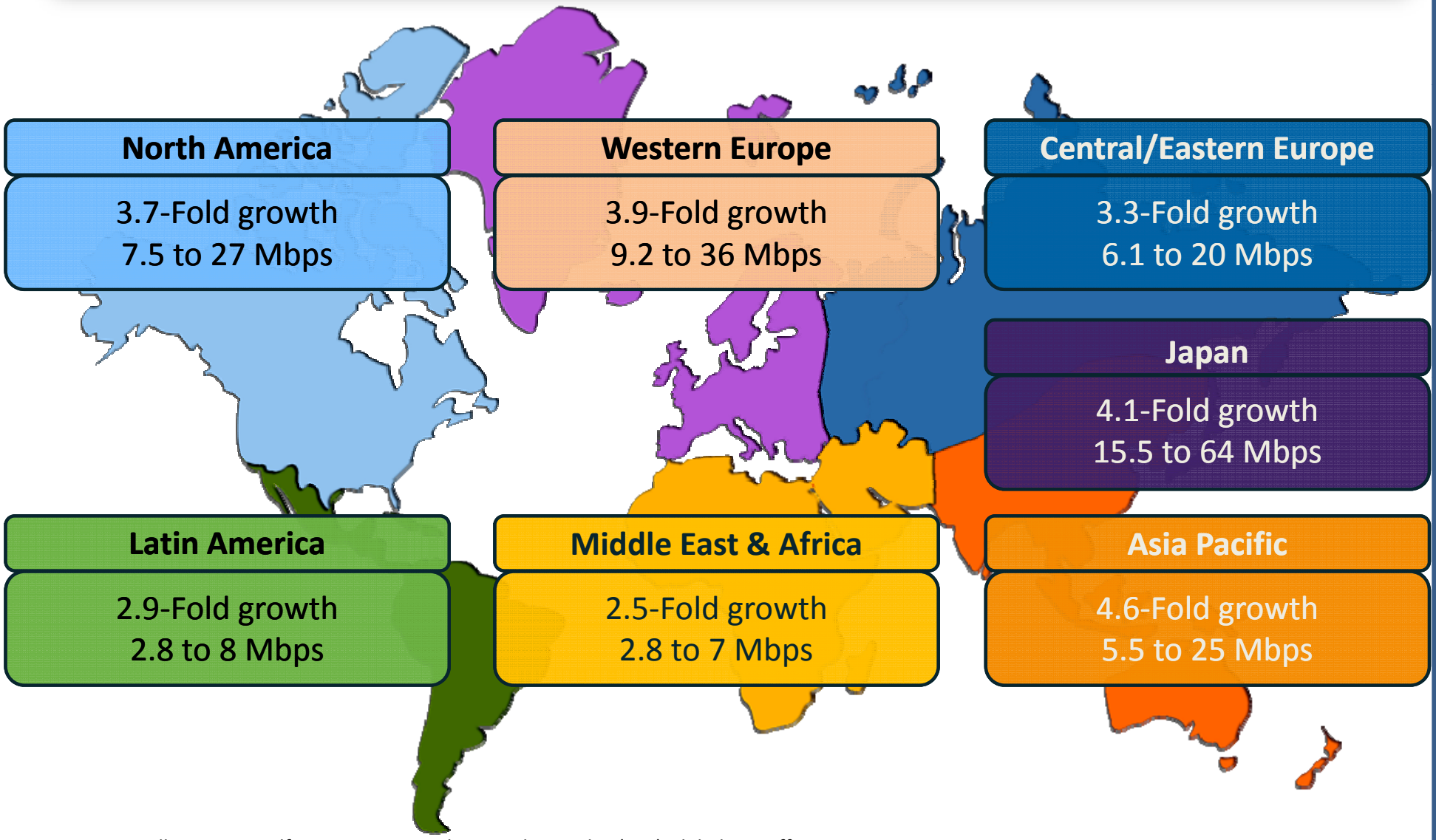
2015 Global Users and Network Connections



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

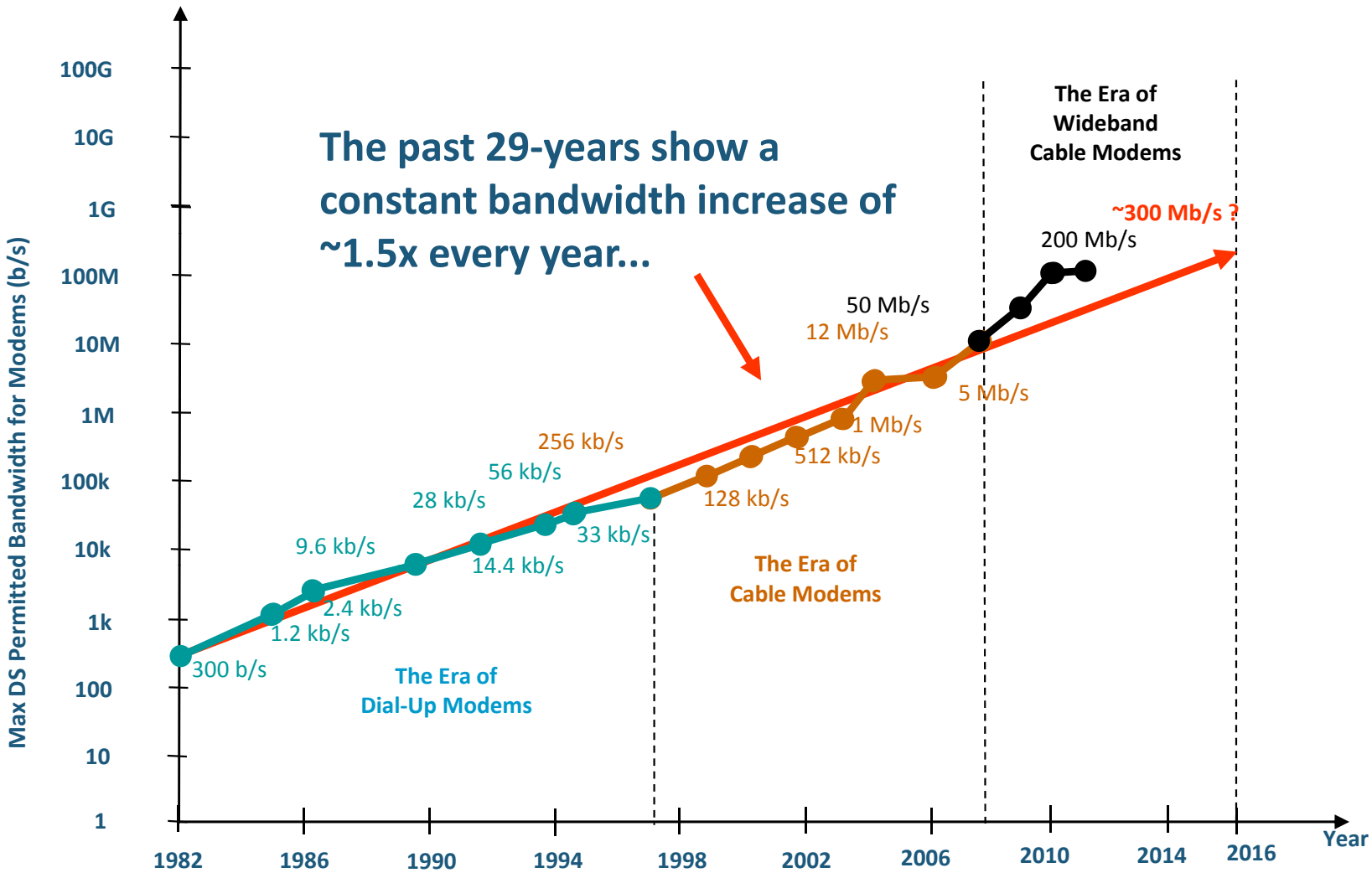
Global Broadband Speed 2010-2015

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



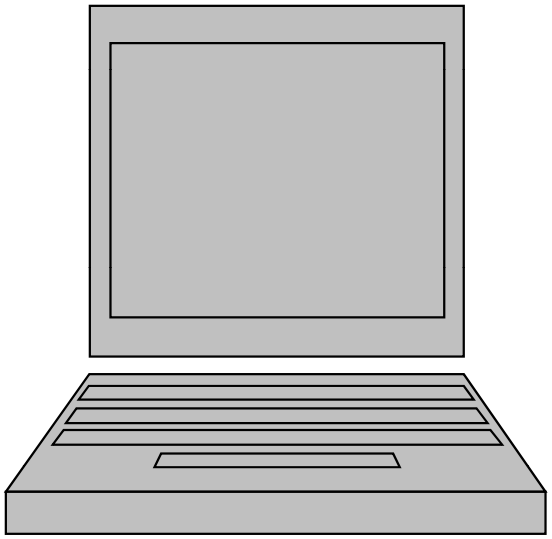
Source: nowell_01_0911.pdf citing Cisco Visual Networking Index (VNI) Global IP Traffic Forecast, 2010–2015, http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/nowell_01_0911.pdf

US Cable Industry: Maximum Permitted Bandwidth Trends (Downstream)



Source: http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/cloonan_01a_0911.pdf

Example: Traffic Generation Comparison

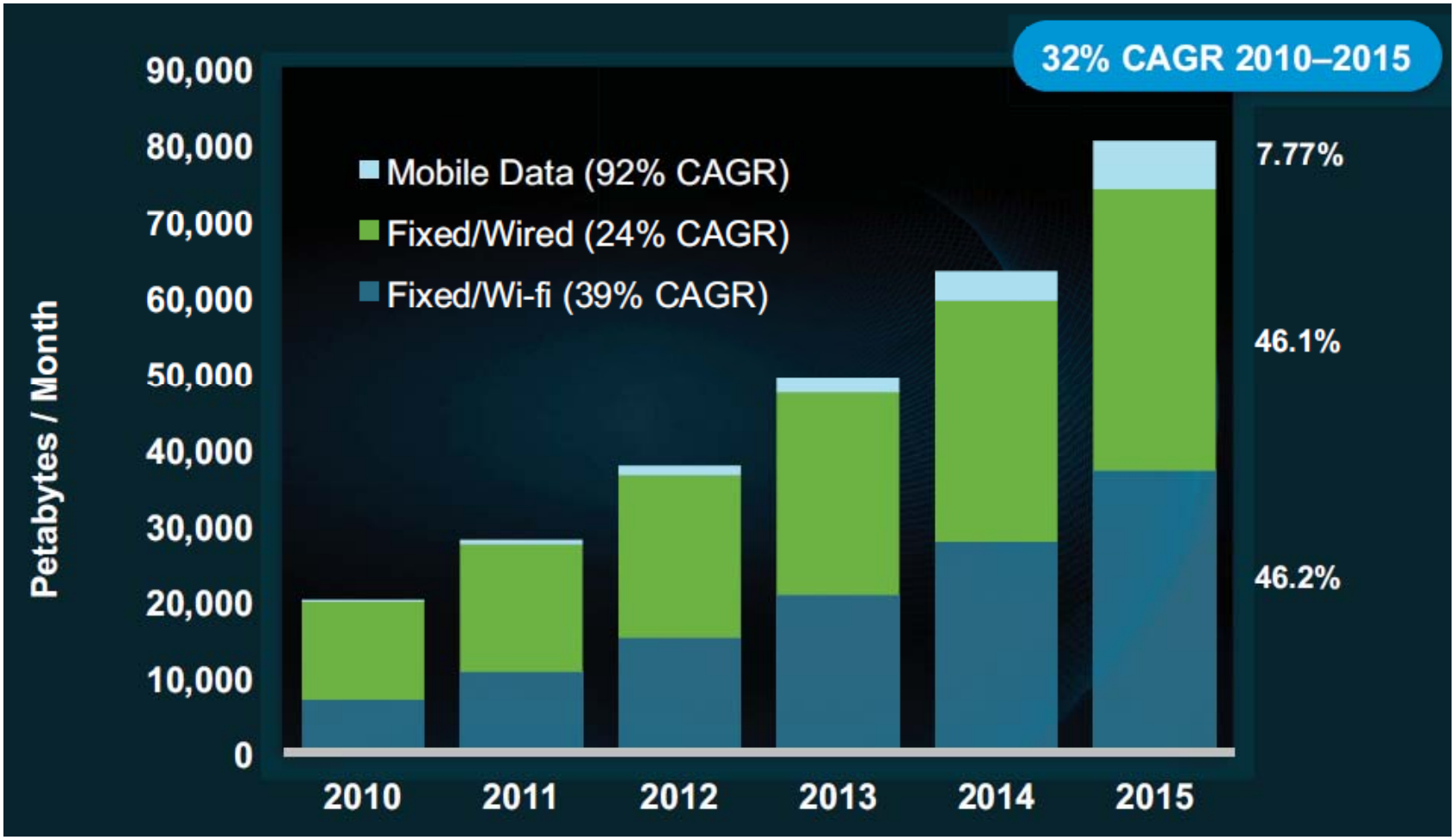


Bandwidth
Generation
Compared to a
32 bit based laptop

Device	Traffic multiplier
Tablet	1.1
64-bit Laptop/PC	1.9
Internet enabled HDTV	2.9
Gaming console	3.0
Internet enabled 3D TV	3.2

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

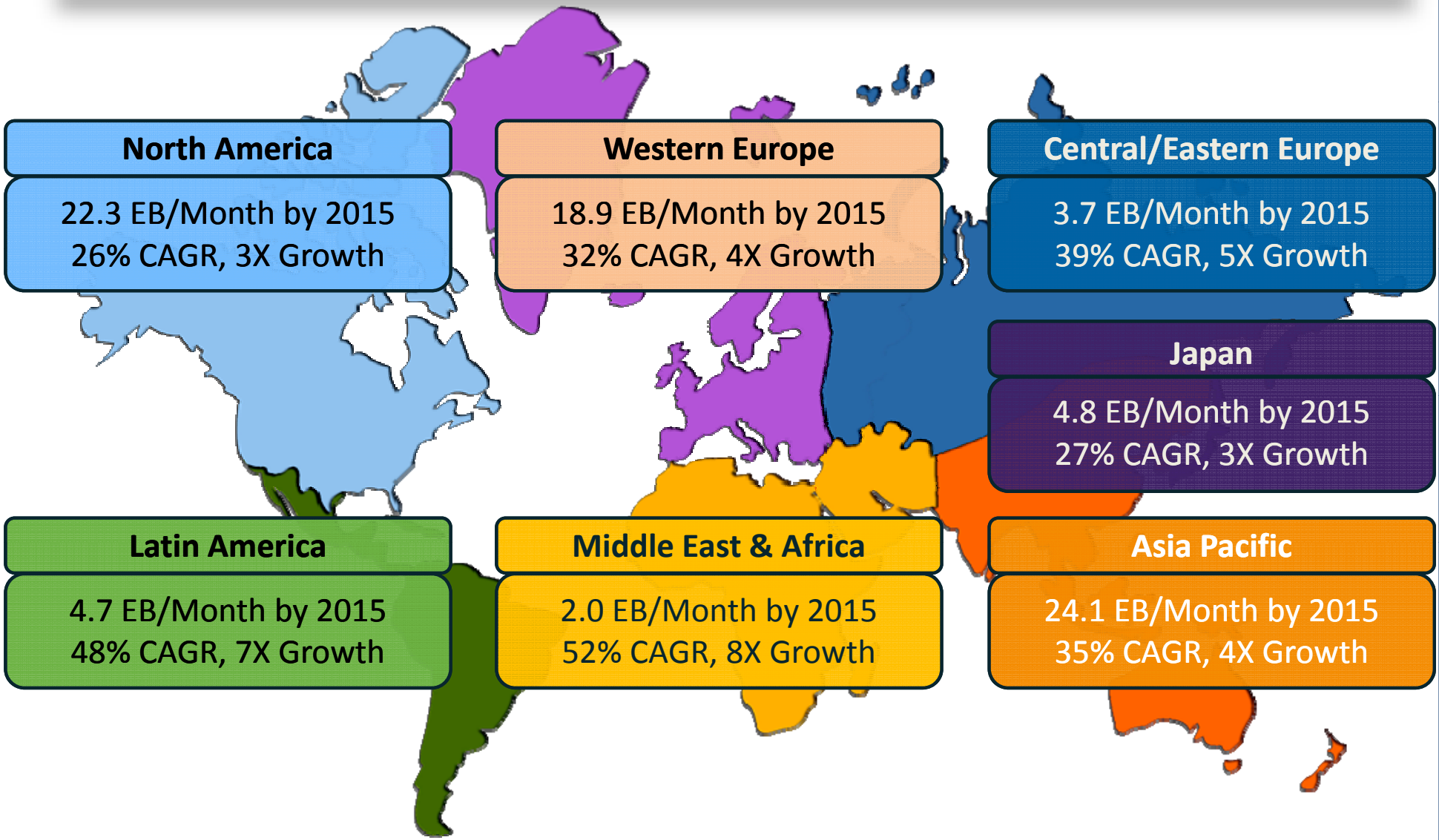
Global IP Traffic by Local Access Technology



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

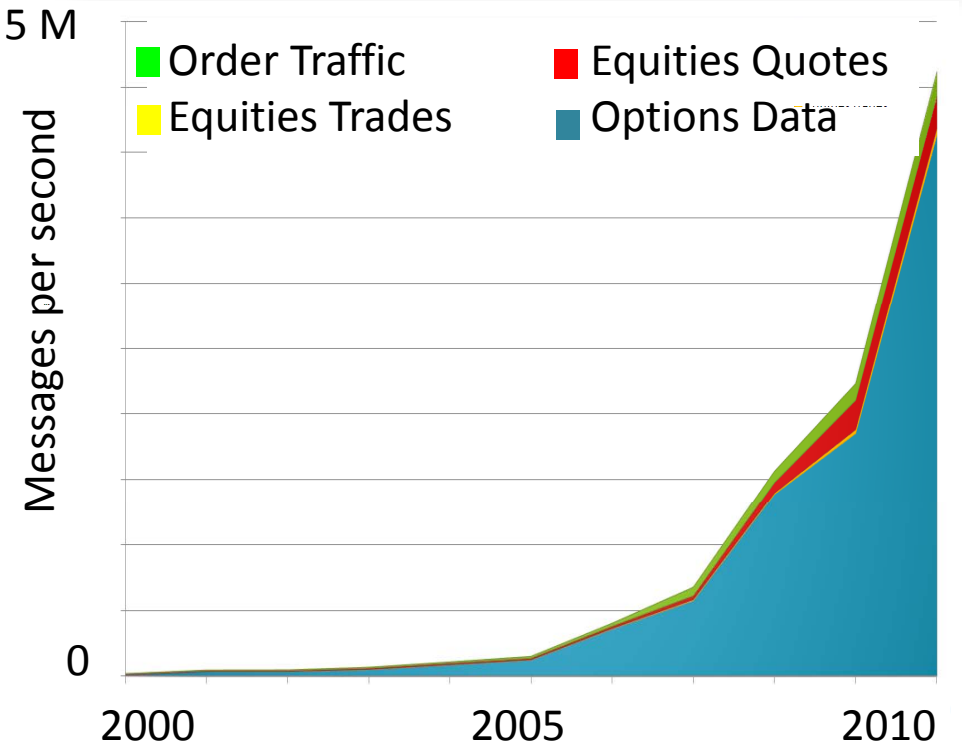
Global IP Traffic Growth, 2010–2015

Regional contributions to the Zettabyte journey



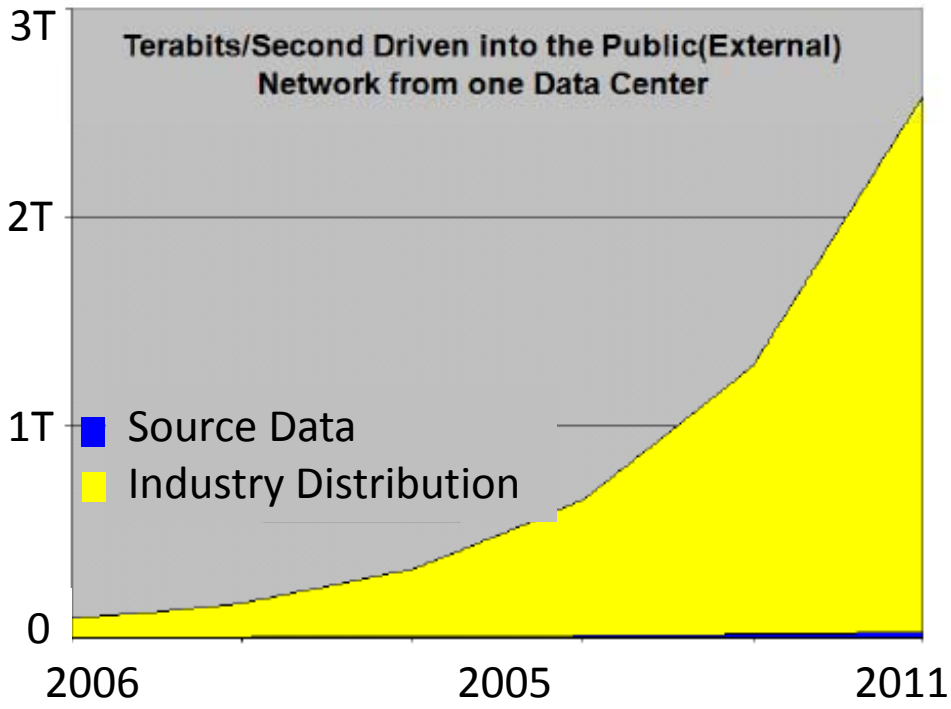
Source: nowell_01_0911.pdf citing Cisco Visual Networking Index (VNI) Global IP Traffic Forecast, 2010–2015, http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/nowell_01_0911.pdf

Example: Financial Sector



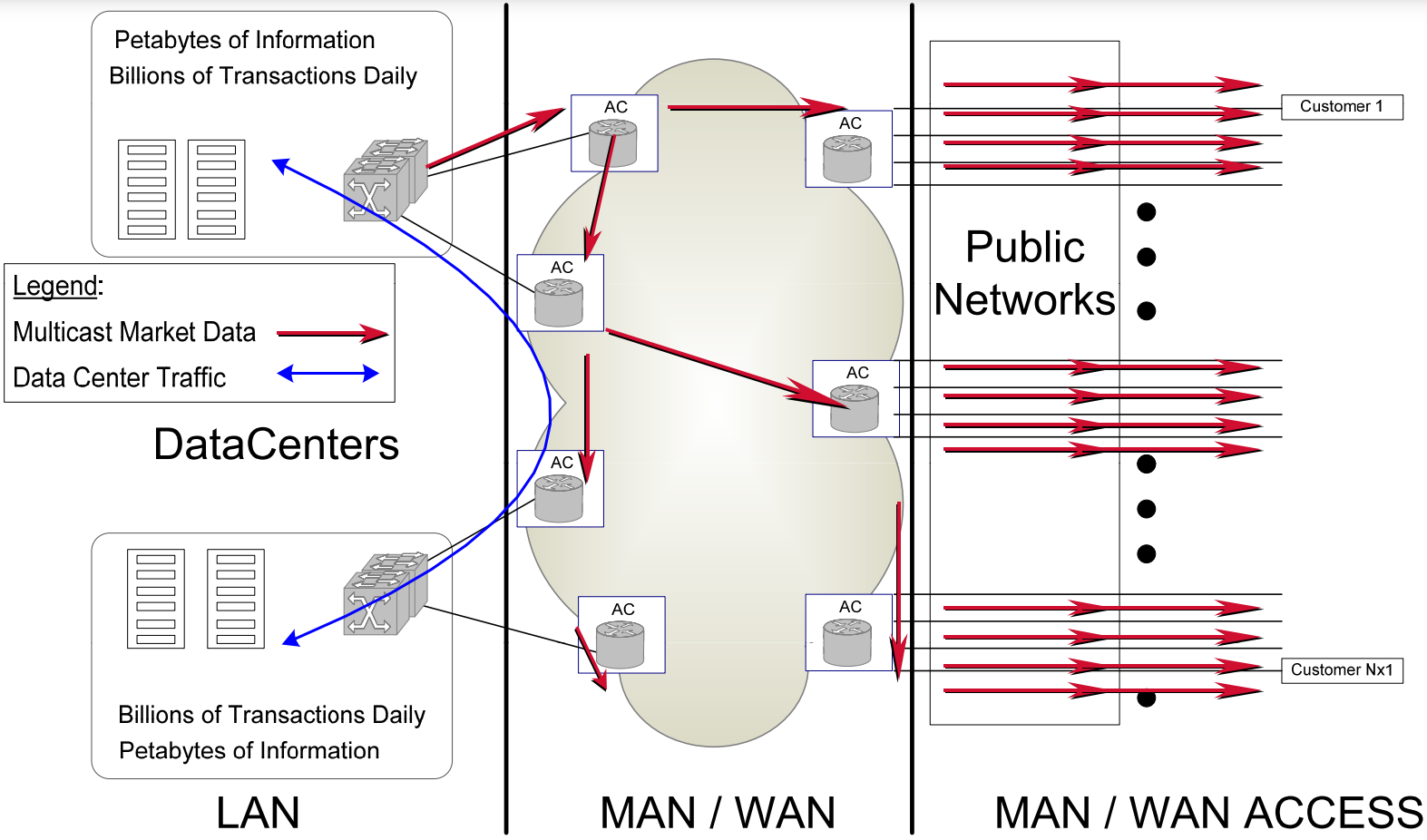
Usage growth

Bandwidth Growth



Source: http://www.ieee802.org/3/ad_hoc/bwa/public/jun11/bach_01a_0611.pdf

Bandwidth Growth is throughout the Eco-system



Networking equipment, compute (servers) equipment and storage equipment all required to scale to match application requirements

Source: http://www.ieee802.org/3/ad_hoc/bwa/public/jun11/bach_01a_0611.pdf

THE DATA CENTER

SCOTT KIPP, BROCADE

Data Center Growth

Increased Storage + **Increased Processing** + **Increased Bandwidth** = **Bandwidth Explosion**

Networking

Entered the 100GbE era in 2010
Individual switches have Tb/s of bandwidth

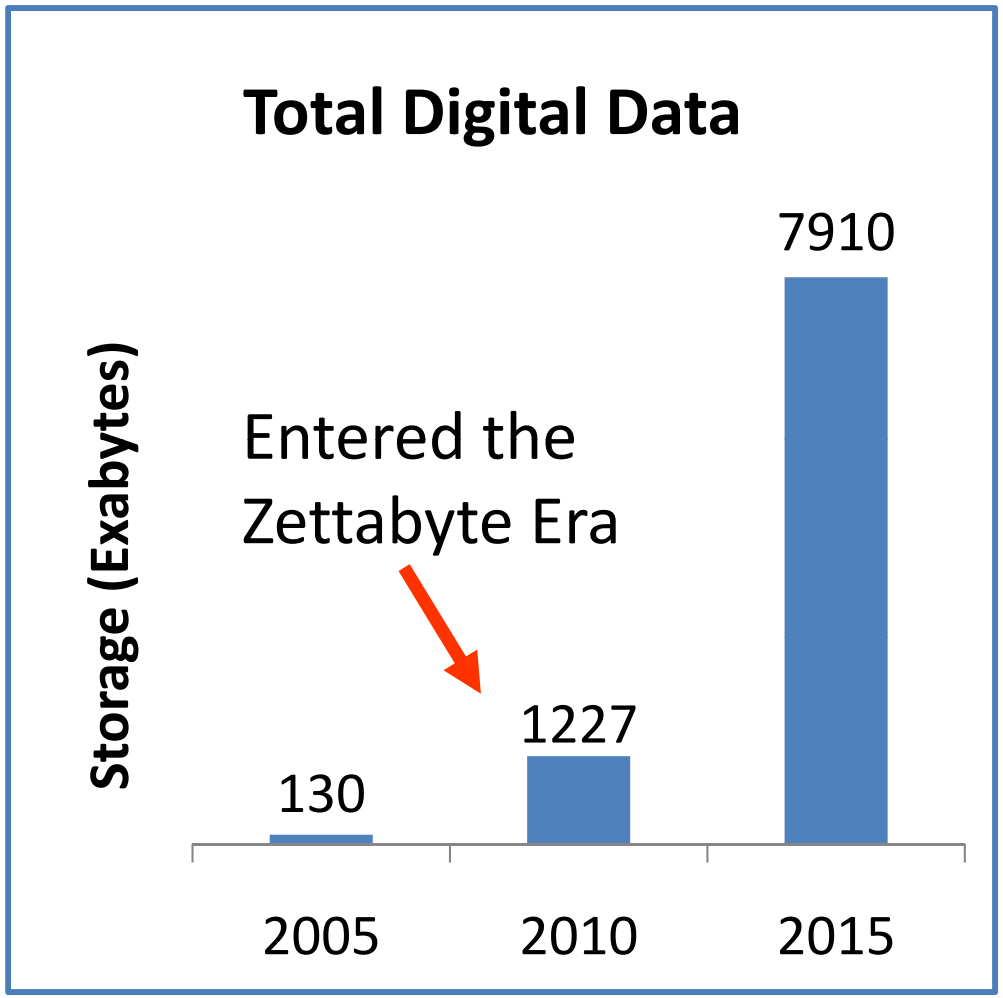
Compute

First petaflop supercomputers in 2011
Individual servers delivering 10s of Gb/s of I/O
PCIe 3.0 supports 2 x 40GbE NICs now

Storage

Entered the zettabyte (1 billion terabytes) era in 2010
Individual disk drives over 1 terabyte
1000 disk drive storage subsystem equals 1 Petabyte

Cumulative Data

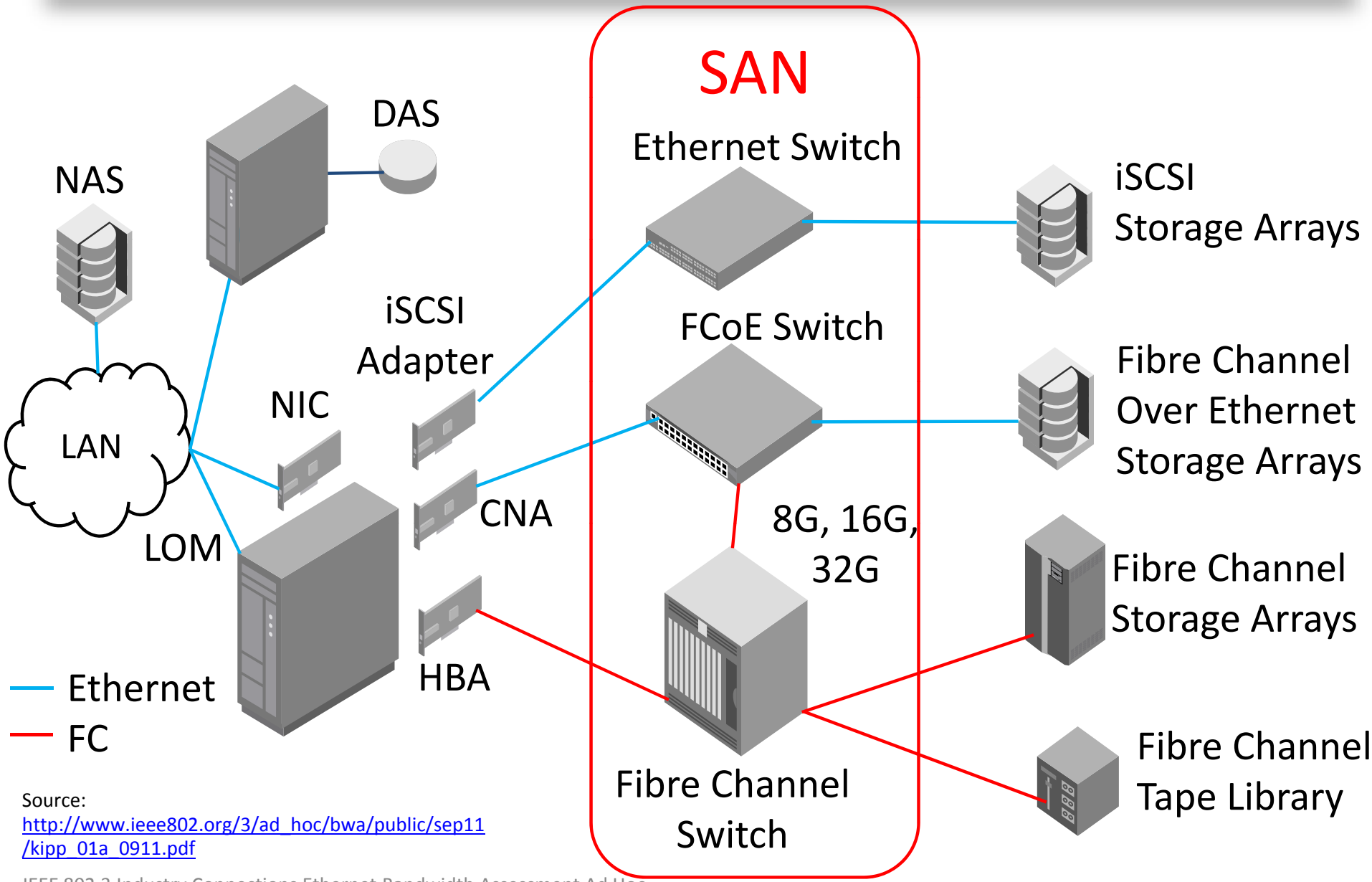


Growth over Next Decade	
# of Servers	x10
Storage	x50
# of Files	x75

Consider the implications!

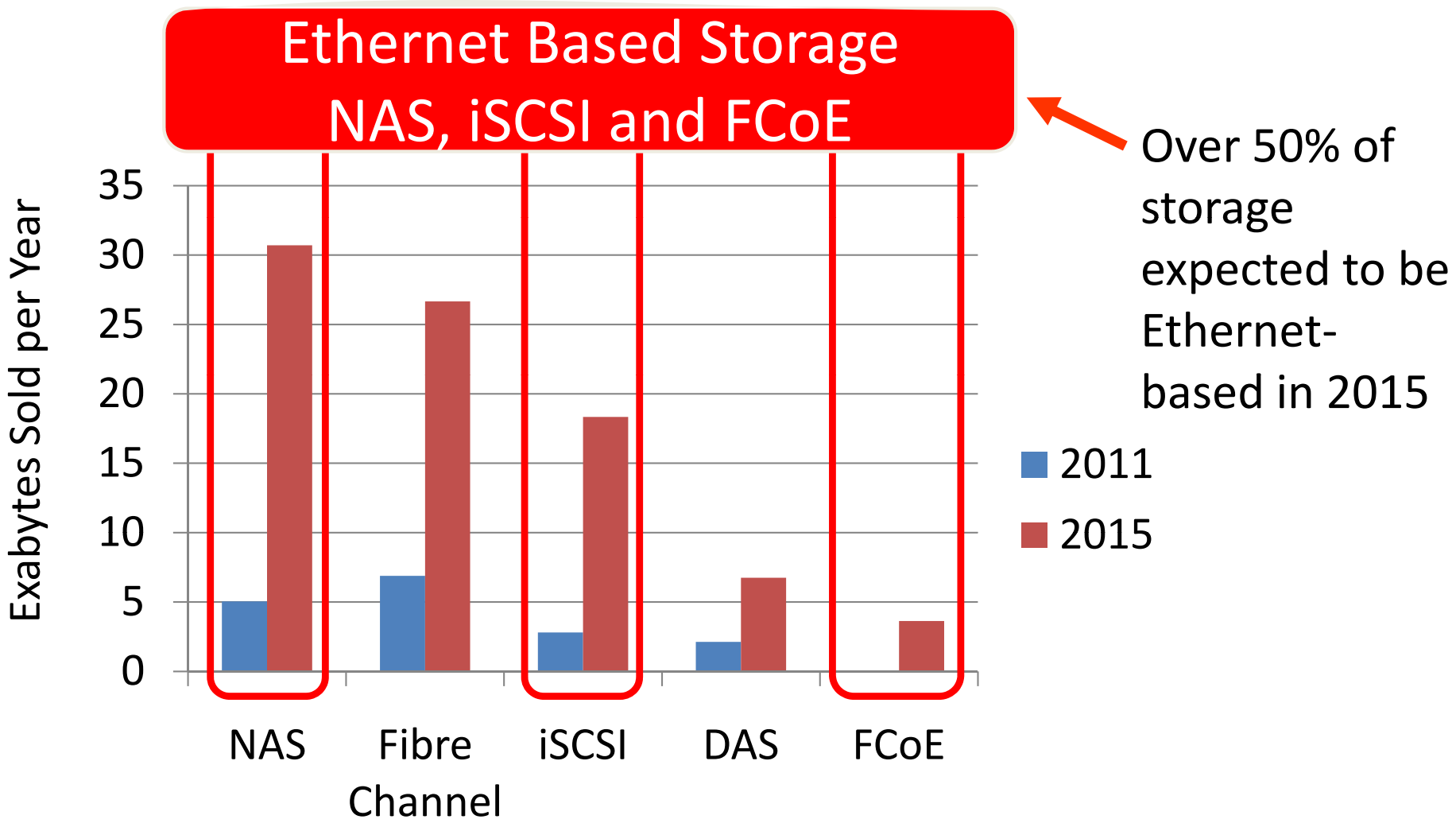
Source: http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/kipp_01a_0911.pdf

Storage Access Methods



Source:
http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/kipp_01a_0911.pdf

Growth in External Storage Subsystems



Source: kipp_01a_0911.pdf citing IDC Worldwide Enterprise Storage Systems 2011-2015
http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/kipp_01a_0911.pdf

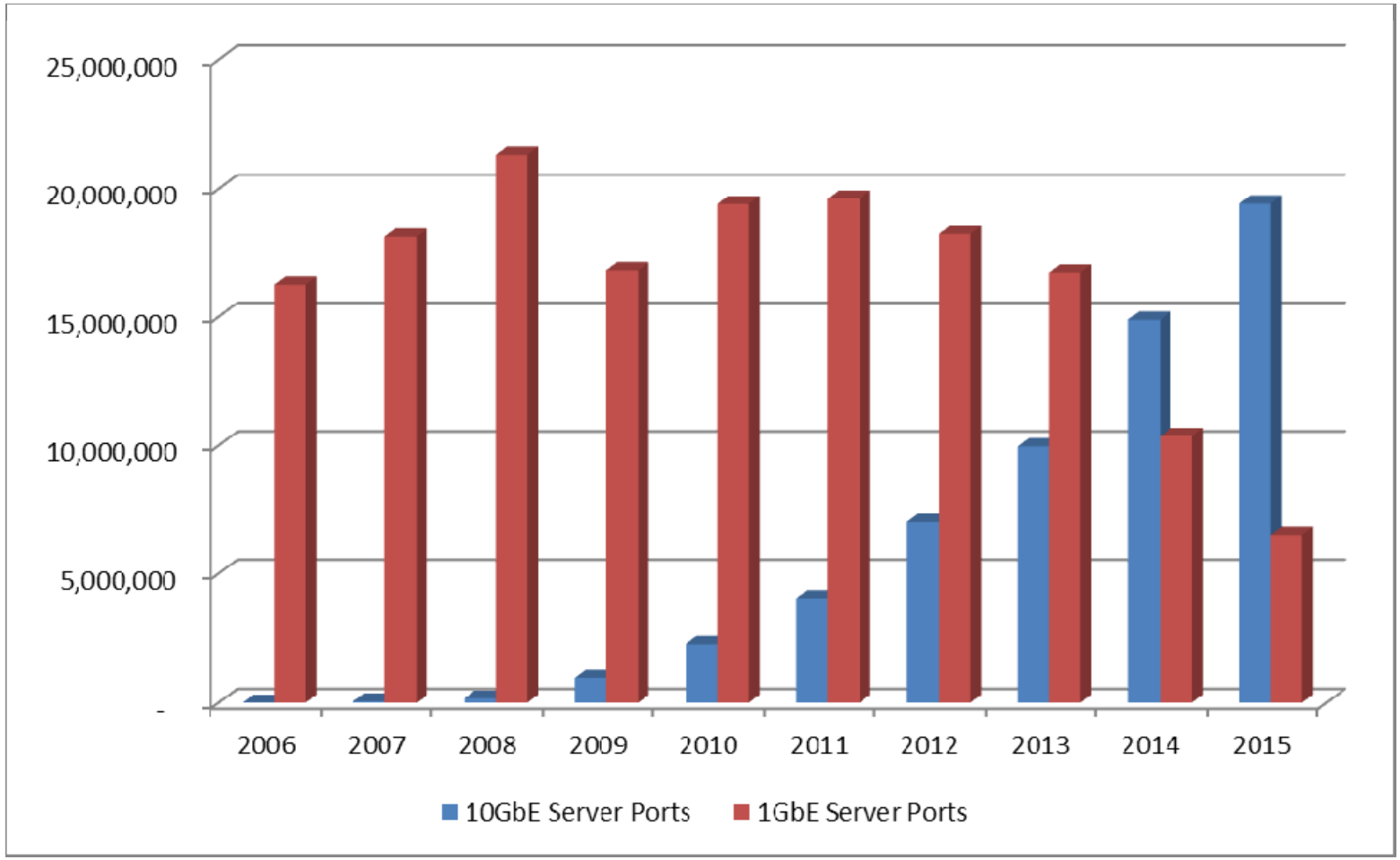
What's Happening With Servers?

- Moore's Law keeps increasing transistor count and improving performance
- 2012 servers begin using PCIe 3.0 to support 40GbE

	Year Released	Gigabyte/s	Speeds of Network Interface Cards (NICs)
PCIe 1.x	2003	4	2 x 10GbE
PCIe 2.x	2005	8	4 x 10GbE
PCIe 3.0	2012	16	2 x 40GbE
PCIe 4.0	2016	32	2 x 100GbE

Source: http://www.ieee802.org/3/ad_hoc/bwa/public/jul11/brown_01a_0711.pdf

Server Port Speed



Source: http://www.ieee802.org/3/ad_hoc/bwa/public/jul11/brown_01a_0711.pdf

Server Aggregation in a Cluster

Each server producing 10-80 Gb/s

Rack of 40 servers producing 0.4 -3.2 Tb/s

Each cluster of 25 racks producing 10-80 Tb/s

I/O per server (Gb/s)	5	10	20	40	80
Servers / rack	40	40	40	40	40
Bandwidth / rack (Gb/s)	200	400	800	1600	3200

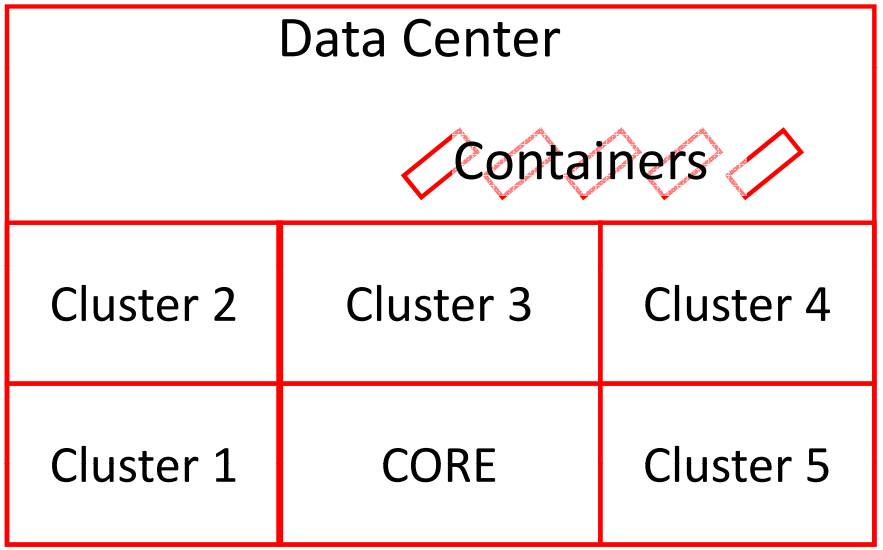
Each 1,000 server cluster sends fraction of possible bandwidth to Interconnect Fabric

Source: http://www.ieee802.org/3/ad_hoc/bwa/public/may11/kipp_01_0511.pdf

Cluster Aggregation in Data Center

Cluster or POD of Racks

Cluster Traffic to Core in Tb/s



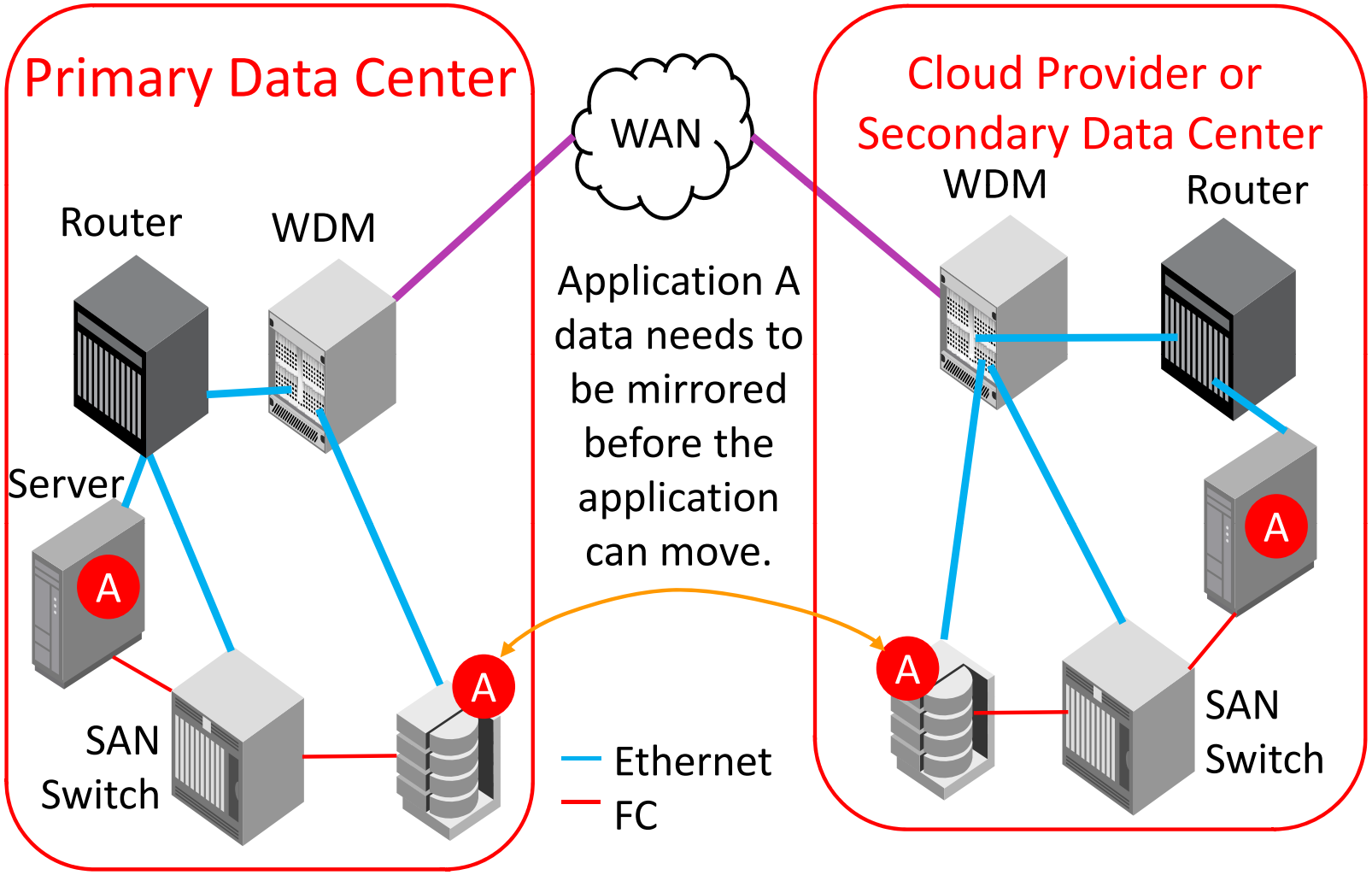
	Cluster bandwidth to core (Tb/s)			
	0.4	1	2	4
Clusters	10	10	10	10
Bandwidth to core (Tb/s)	4	10	20	40
Bandwidth to WAN (Gb/s)	20	40	200	400
Oversubscription to WAN	200	250	100	100

Core Switches/ Routers

Source: http://www.ieee802.org/3/ad_hoc/bwa/public/may11/kipp_01_0511.pdf

Example: Impact of Data Migration

Application migration are between servers, clusters or data centers

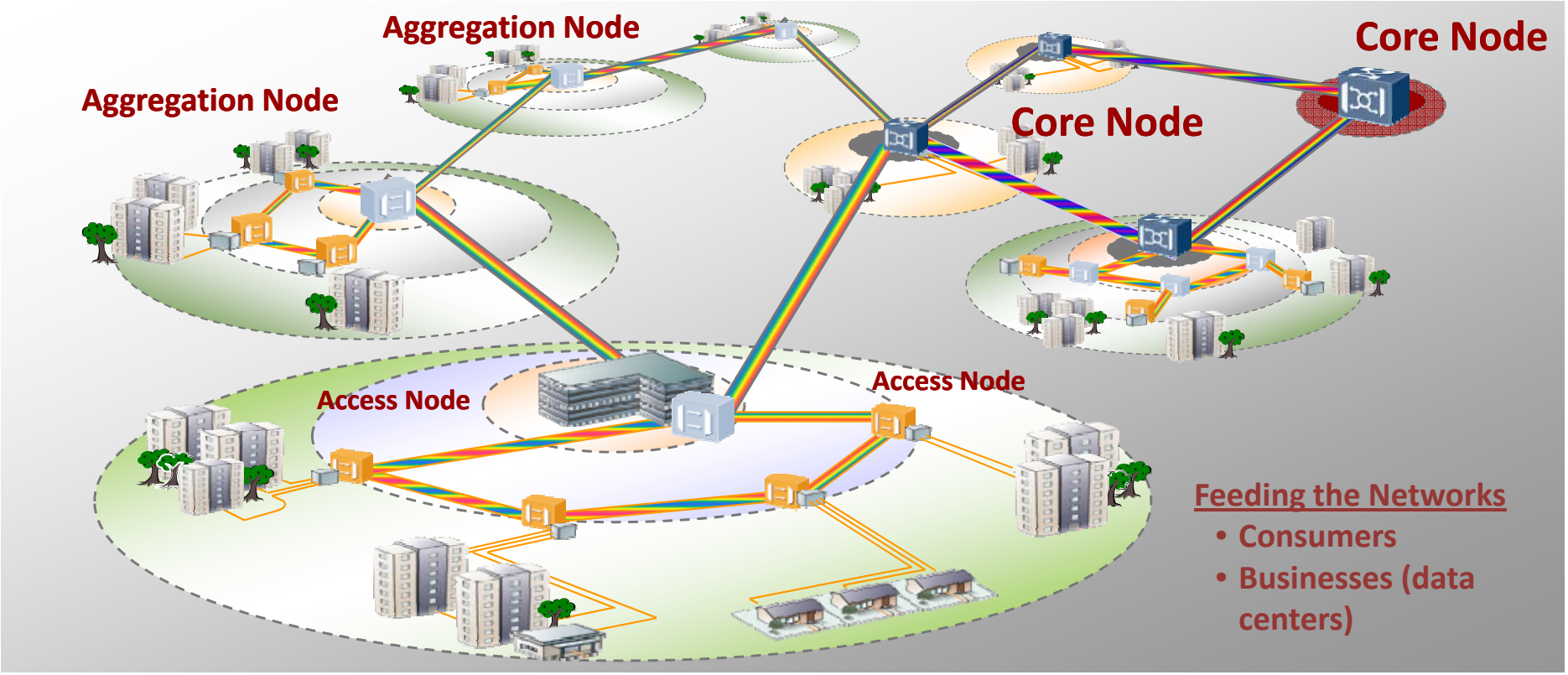


Source: http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/kipp_01a_0911.pdf

TRANSPORT NETWORKS

PETER STASSAR, HUAWEI

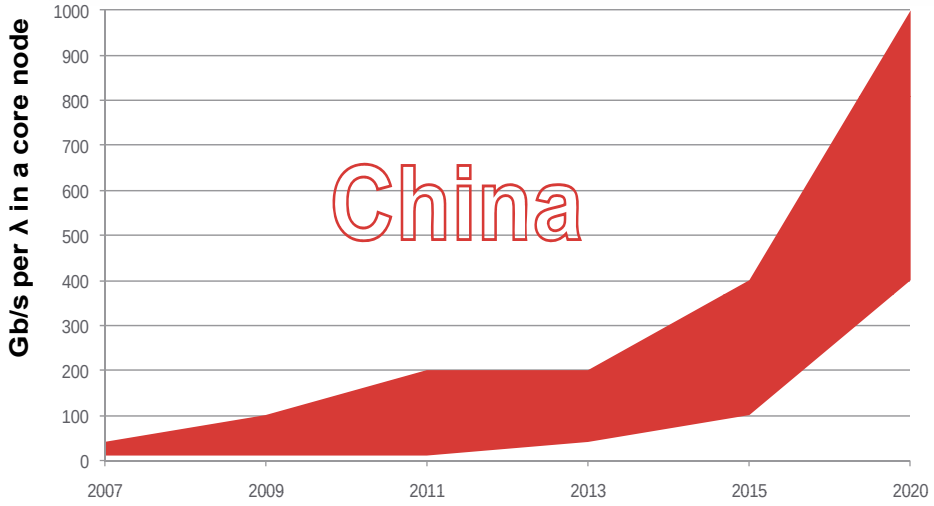
Transport Network Classifications

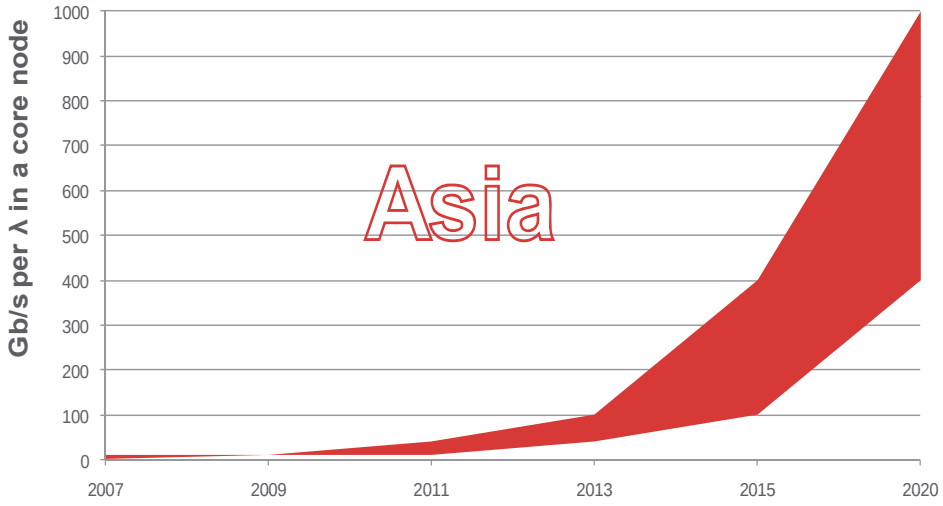


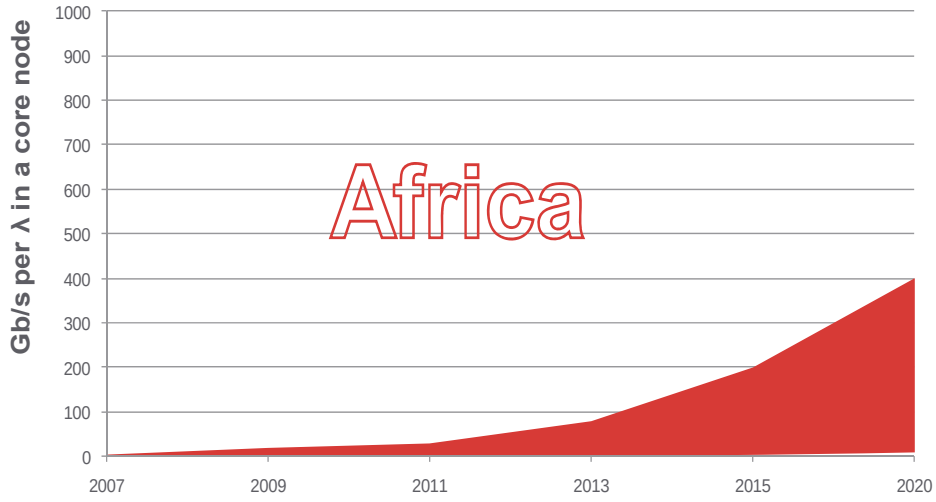
Note: "Aggregation Nodes" in single carrier networks are equivalent to "IXPs" in multi carrier networks

Source: http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/huang_01_1111.pdf

Bandwidth Needs Per Wavelength: Core Nodes in Single Carrier Networks





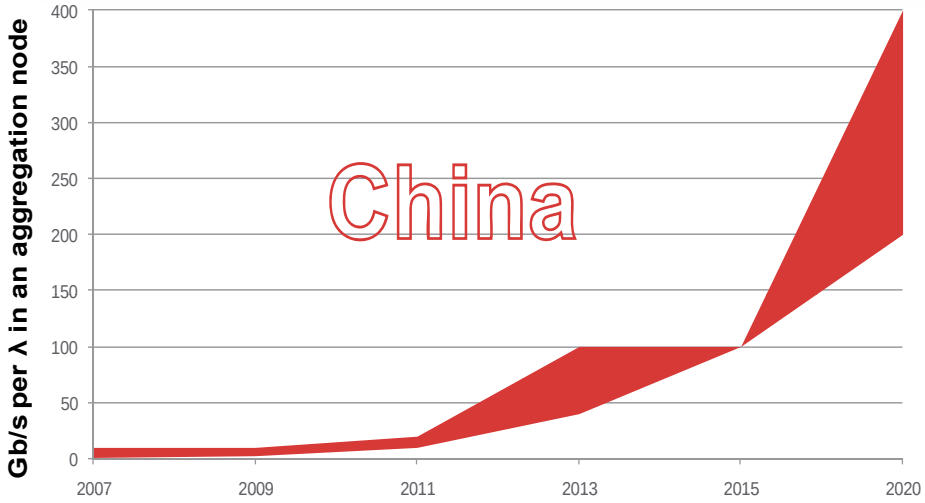


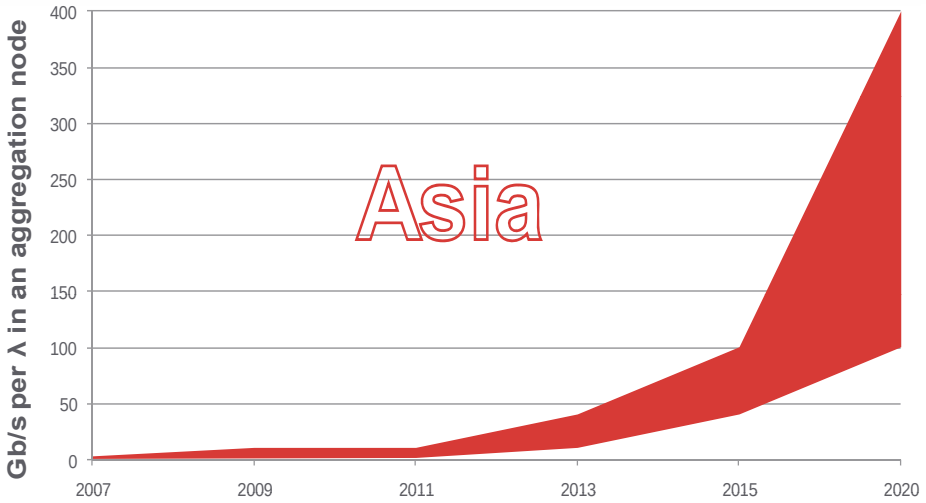
Notes:

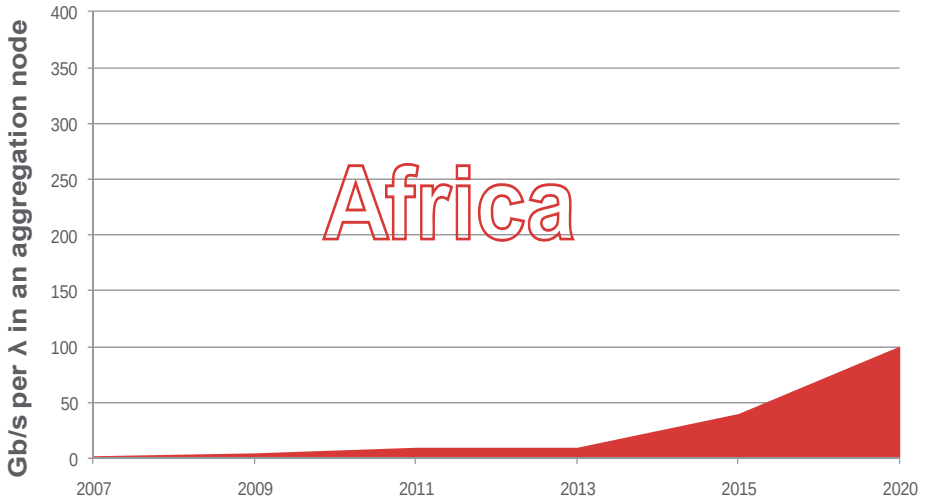
- Single carrier networks
- More than one carrier in overview
- 2015: Range 100G – 400G per λ
- 2020: Range 400G – 1T per λ
- Africa: same trend, a bit delayed

Source: http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/huang_01_1111.pdf

Bandwidth Needs Per Wavelength: Aggregation Nodes in Single Carrier Networks





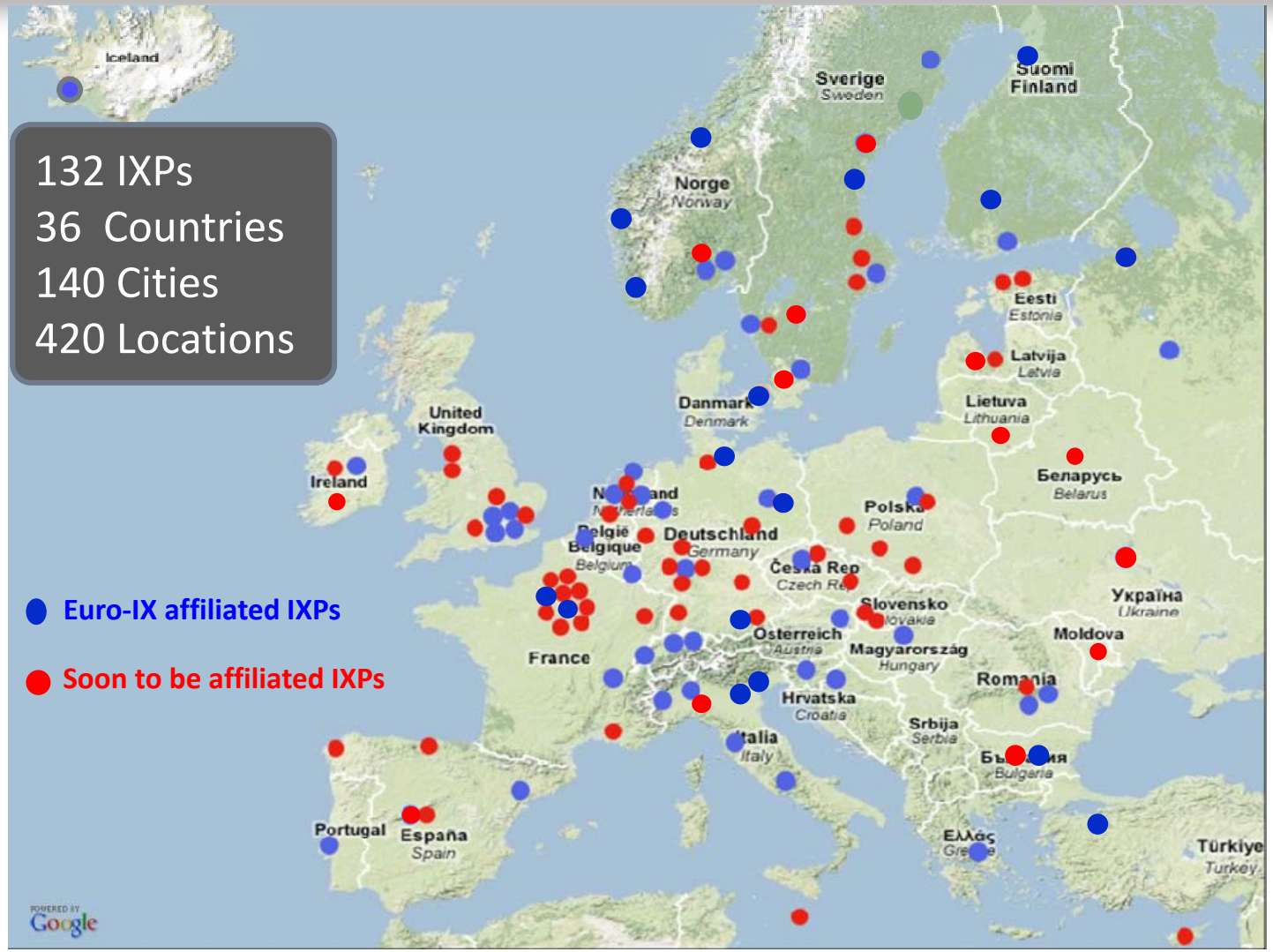


Notes:

- Single carrier networks
- More than one carrier in overview
- 2015: Range 50G – 100G per λ
- 2020: Range 100G – 400G per λ
- Africa: same trend, a bit delayed

Source: http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/huang_01_1111.pdf

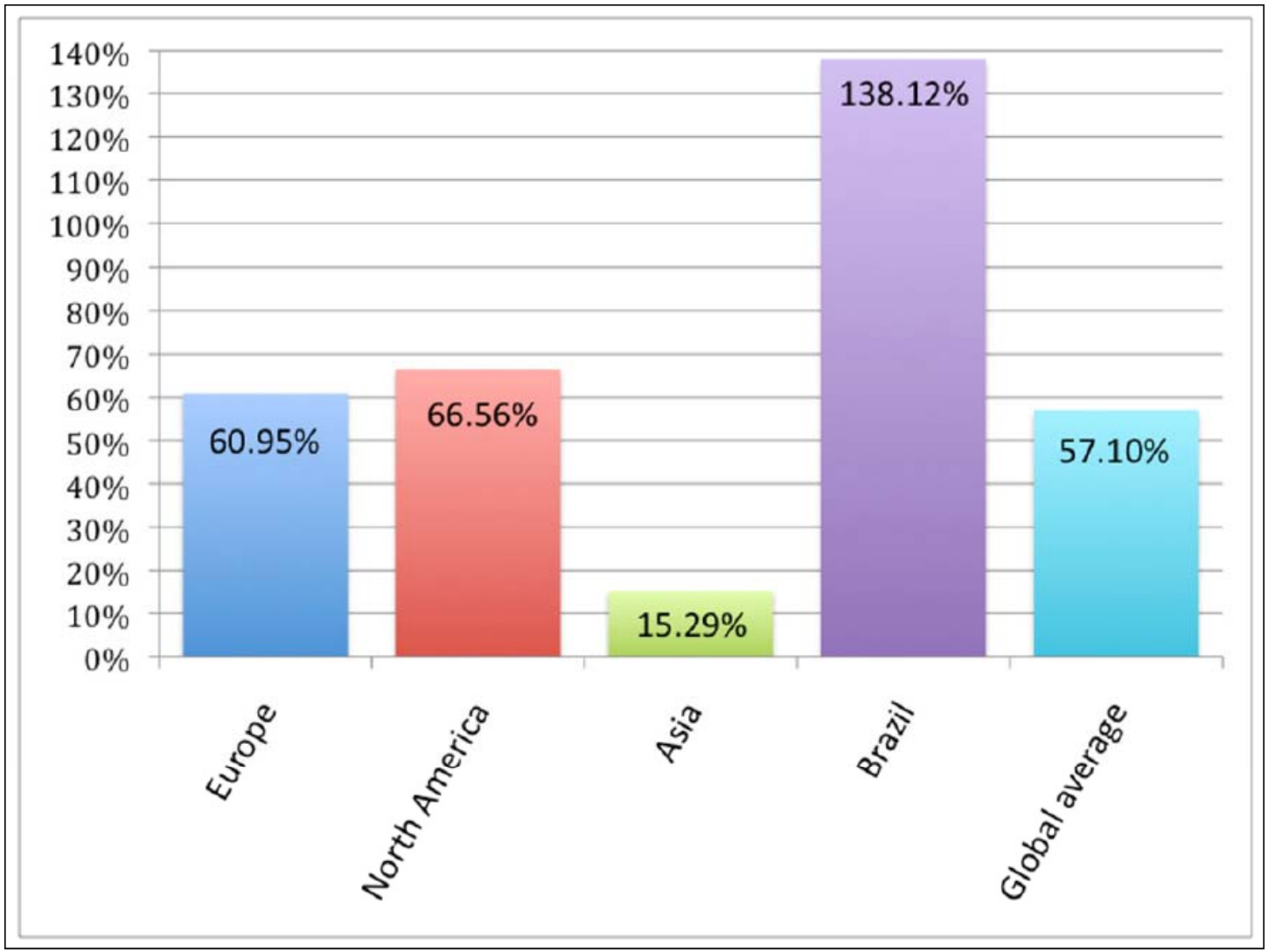
Euro-IX IXP Locations



Note: Global 321 IXP's (100%), EU 41%, America's 33%, Asia-Pacific 19%, Africa 7%

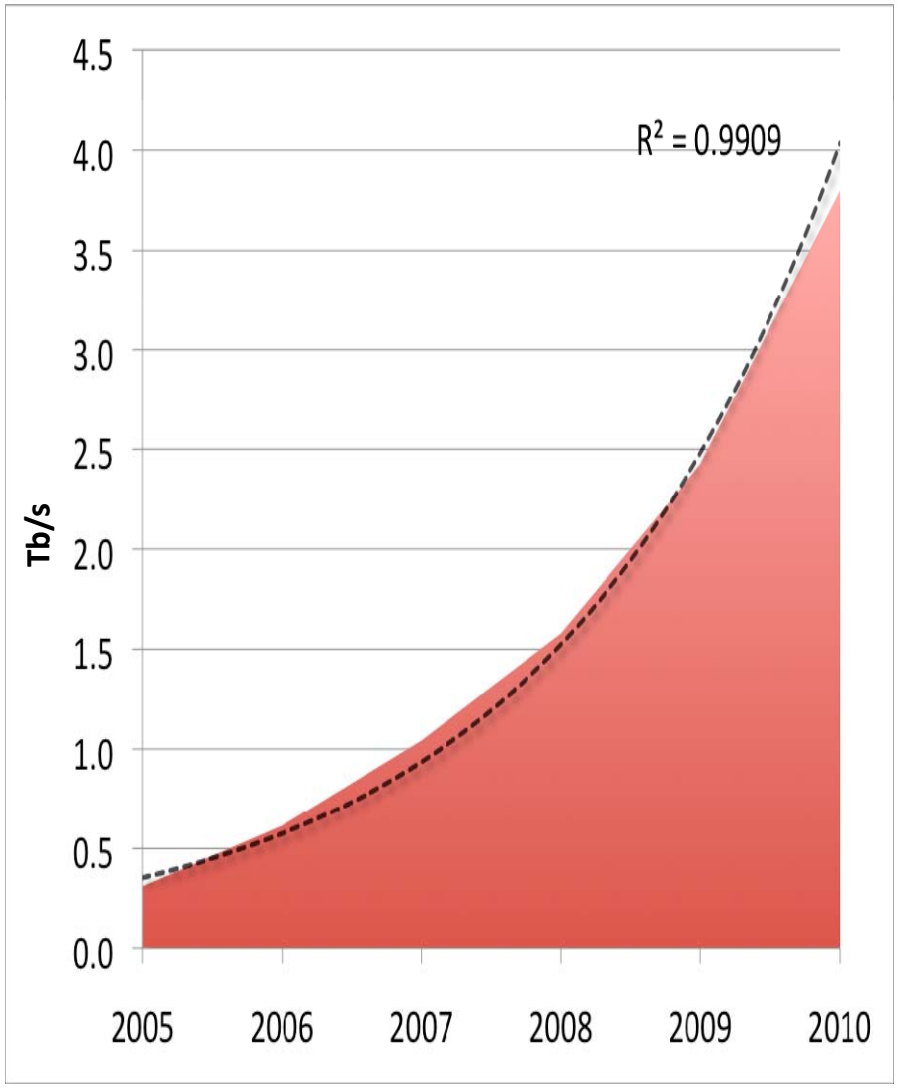
Source: http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/steenman_01_1111.pdf

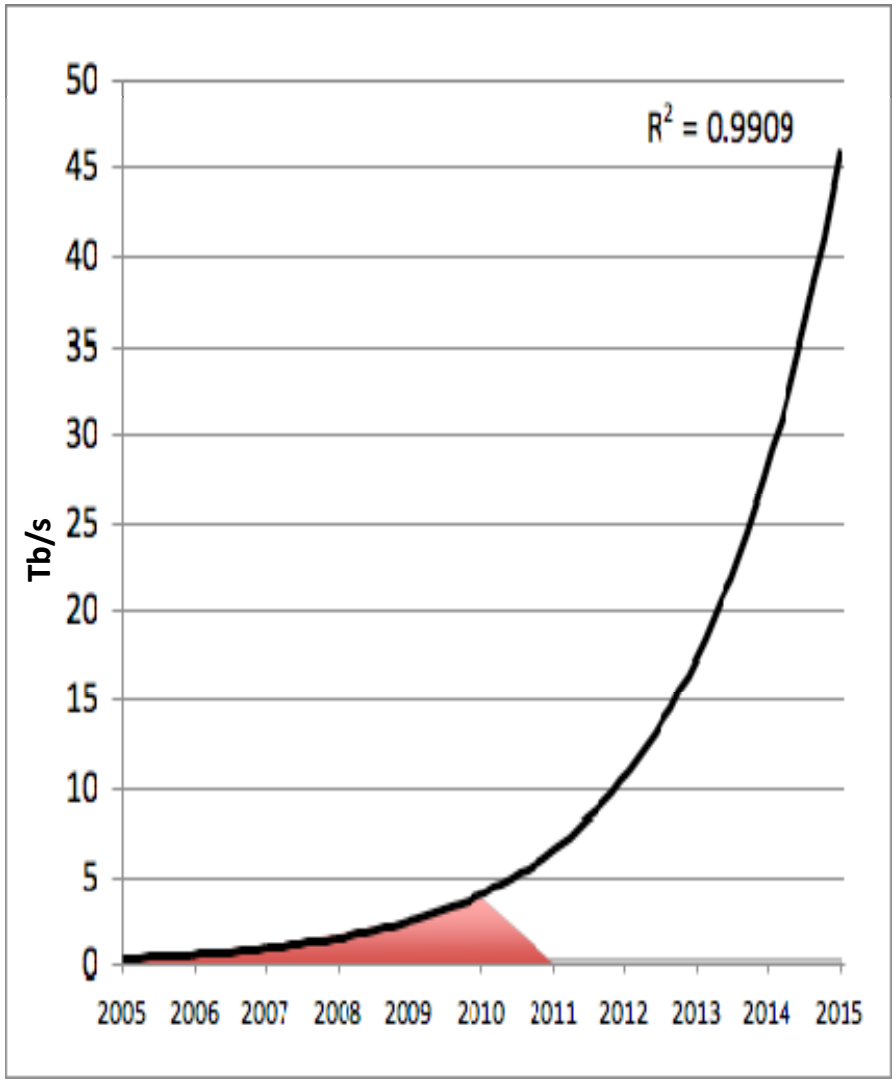
Global Annual IXP Peak Traffic Growth Rates: By Region (for 2010)



Source: http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/steenman_01_1111.pdf

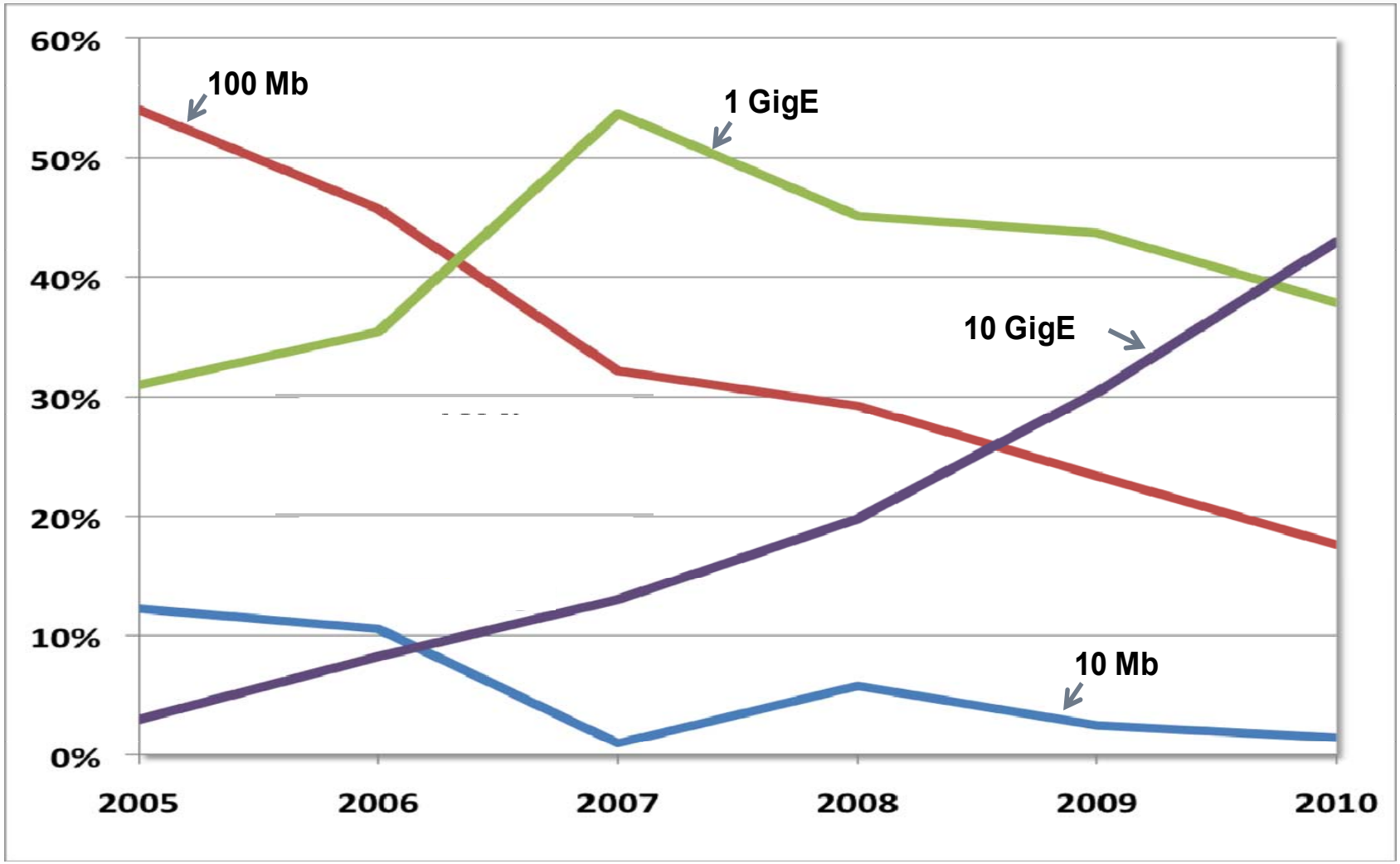
Five Year Peak European IXP Traffic Projections





Source: http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/steenman_01_1111.pdf

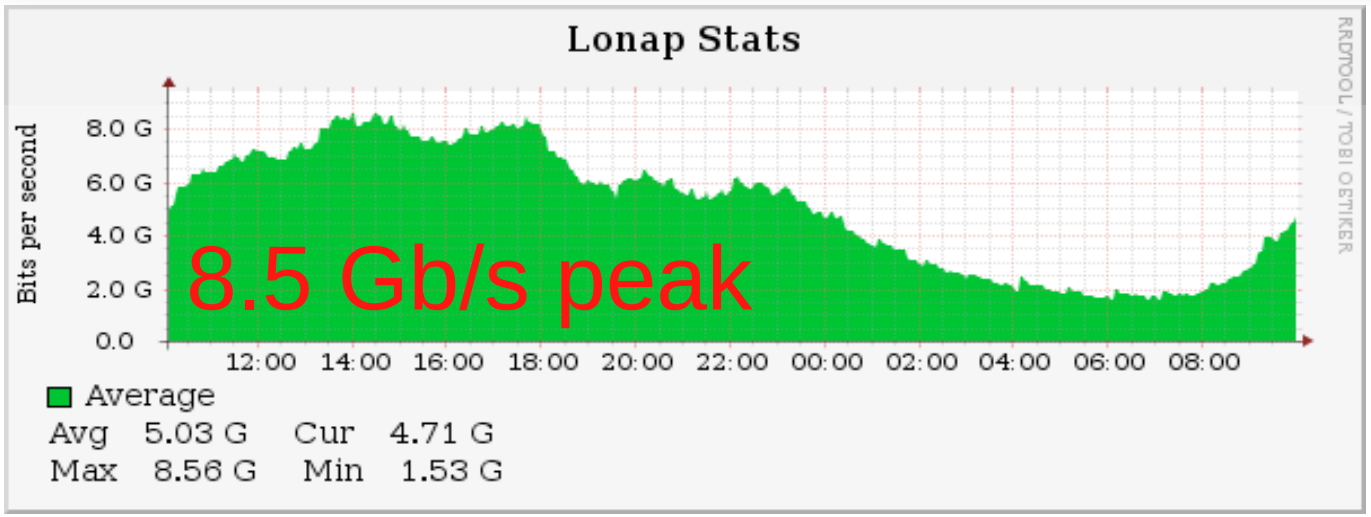
Euro-IX European Member Port Usage



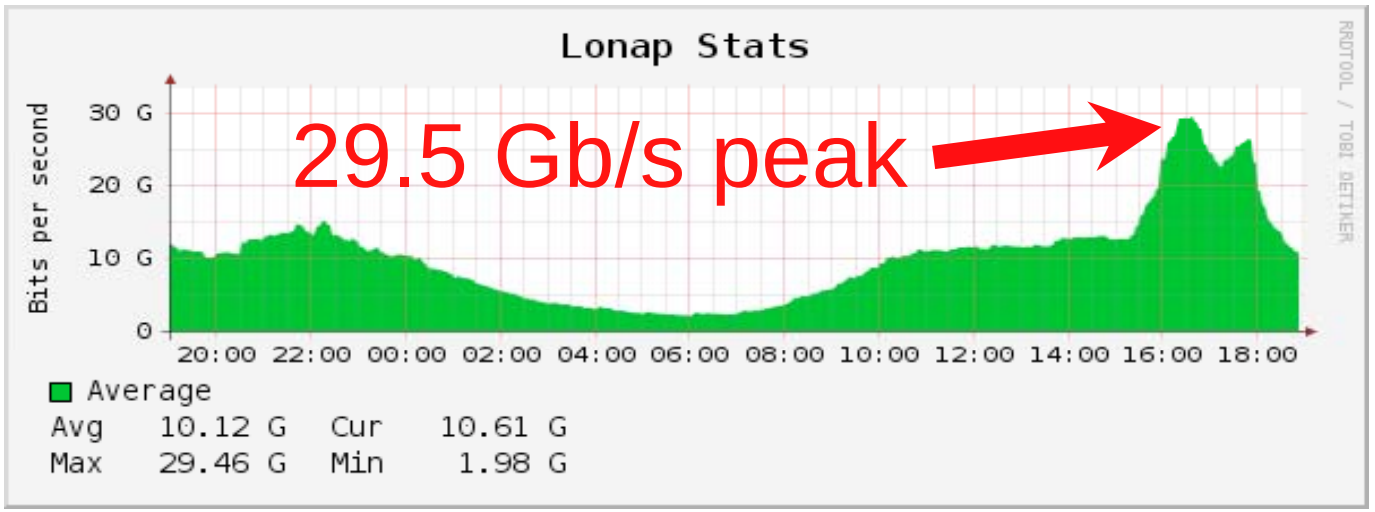
Notes: 10Mb almost zero, 100Mb strongly ↘, 1G starting ↘, 10G strongly ↗

Source: http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/steenman_01_1111.pdf

Example: Traffic Planning



LONAP (London) traffic on a typical 2010 weekday



Wed. afternoon World Cup 2010 England vs Slovenia

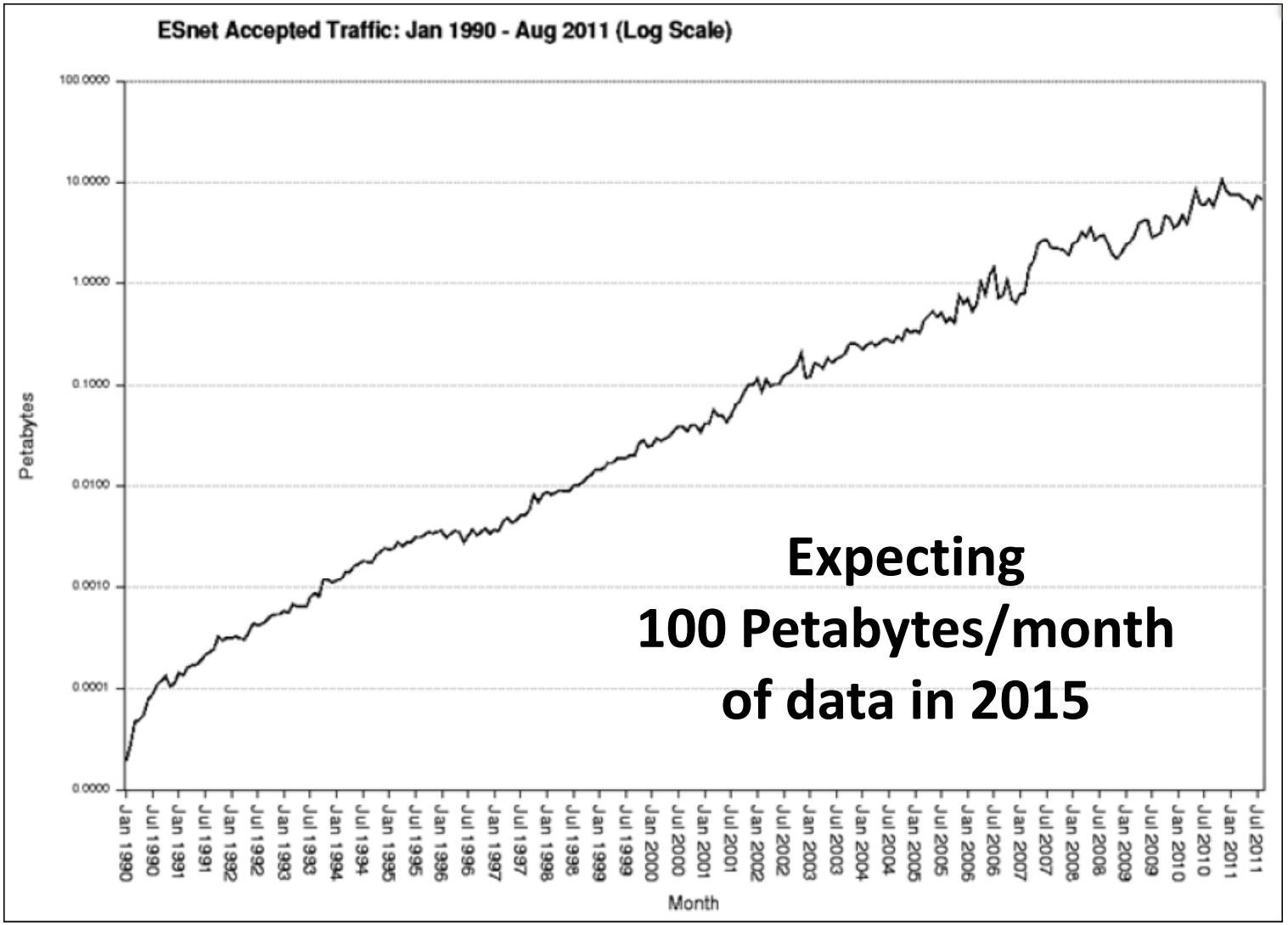
Source: http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/steenman_01_1111.pdf

Examples: Data Intensive Science

- CERN: Atlas detector in LHC (Large Hadron Collider) generates ~1 petabyte/sec
- Genome sequencing:
  - Per-instrument data rate strongly ↗ (~10x over 5 years)
  - Cost of sequencers strongly ↘ (10x over 5 years)
- Futures: 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

ESnet Accepted Traffic (Petabytes/month)



Source: http://www.ieee802.org/3/ad_hoc/bwa/public/dec11/dart_01_1211.pdf

SUMMARY

PETE ANSLOW, CIENA

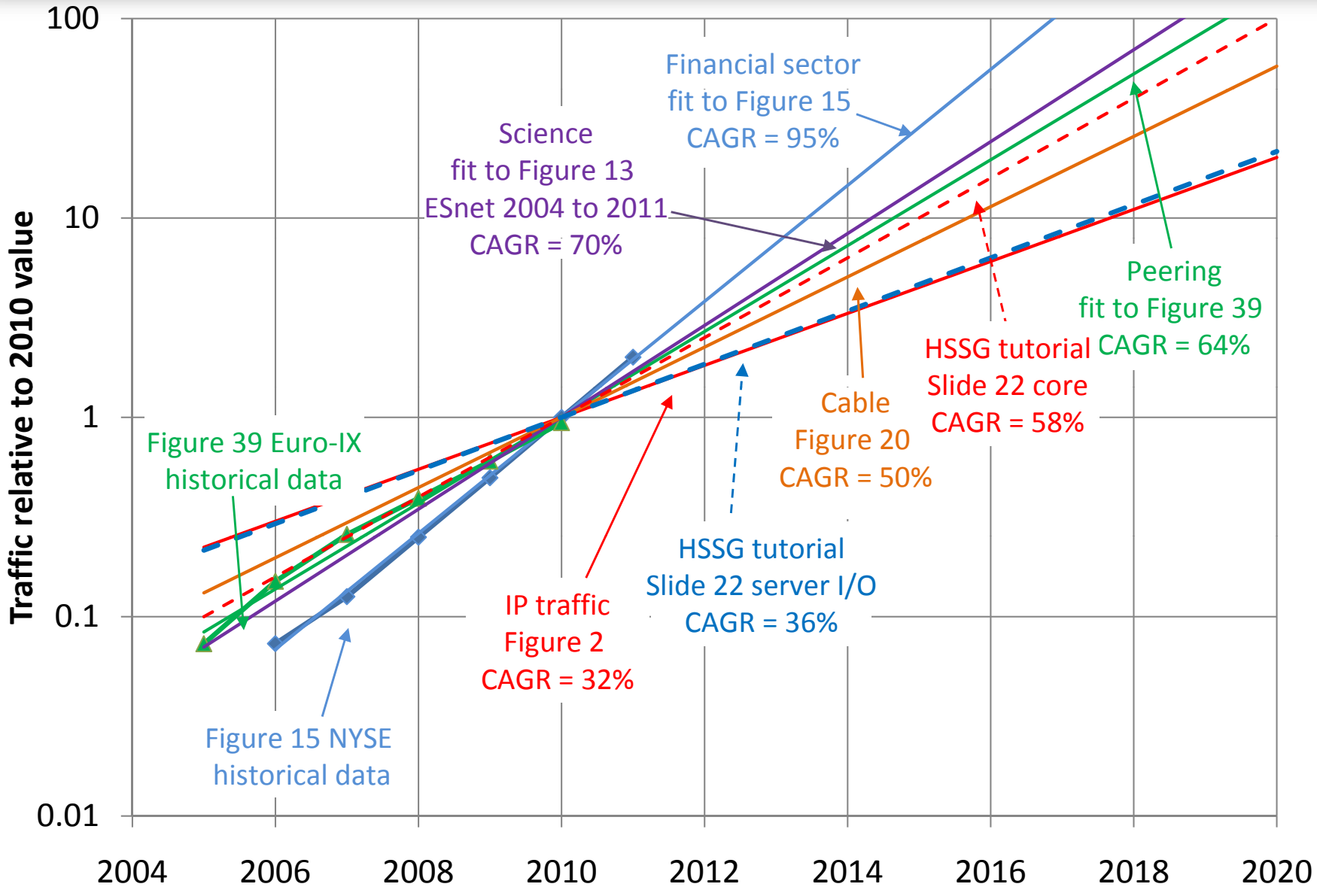
Summary method

- Relative growth of the various sectors plotted on a single chart
  - The growth of each sector was normalized to 2010 (the year IEEE Std 802.3ba was approved)
- This growth is a predictor of the future only if downward cost per bit trend is continued
  - Ethernet cost per bit has to fall with time or the predicted exponential rise in traffic will result in unsupportable costs

Variation Factors

- Reported growth trends are “sector” averages
  - There is considerable variation by region or by market segment
- Example 1 – Predicted Regional Growth of IP traffic (2010 – 2015)
  - Minimum: 26 % in North America
  - Maximum: 52 % in the Middle East and Africa
  - Reported average: 32%
- Example 2 – Growth in IXP peak Traffic in 2010
  - Minimum: 45.88 % in small IXPs
  - Maximum: 78.18 % in medium IXPs
  - Reported Average: 64%

Growth Rate Trends



Summary

- The exponential rise in traffic is predicted to continue

Increased # of users + **Increased access rates and methods** + **Increased services**

- Servicing demand with existing rates or new ones > 100 Gb/s will depend on the cost effectiveness of the solution

Publication Information

- IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc Report is pending final approval this week.

- Upon approval final report to be published:
  - http://www.ieee802.org/3/ad_hoc/bwa/BWA_Report.pdf

The IEEE 802.3 Ethernet Bandwidth Assessment Ad Hoc would like to thank all of the individuals who contributed data to this effort.

THANK YOU!
QUESTIONS?

ABBREVIATIONS

1GbE	1 Gb/s Ethernet
10GbE	10 Gb/s Ethernet
40GbE	40 Gb/s Ethernet
100GbE	100 Gb/s Ethernet
3D TV	three-dimensional television
BW	bandwidth
CAGR	compound annual growth rate
CMTS	cable modem termination system
CNA	Converged Network Adapter
DAC	digital-to-analog converter
DAS	direct attached storage
DOCSIS	Data Over Cable Service Interface Specification
DS	downstream
EPON	Ethernet passive optical network
FCoE	Fibre Channel over Ethernet
HBA	Host Bus Adapter
HHP	house-holds passed
HPC	high performance computing
HSSG	Higher Speed Study Group
I/O	input/output
IP	Internet Protocol
iSCSI	Internet small computer system interface
ISP	Internet service provider

IXP	Internet exchange point
LAN	local area network
LAG	link aggregation
LHC	Large Hadron Collider
LOM	LAN on motherboard
MAN	metropolitan area network
MSO	multi-system operator
NAS	network attached storage
NIC	network interface card
OEM	original equipment manufacturer
OTN	Optical Transport Network
P2P	peer-to-peer
PC	personal computer
PCIe	Peripheral Component Interconnect Express
QAM	quadrature amplitude modulation
RFOG	radio frequency over glass
SAN	storage area network
SMB	small and medium business
US	upstream
VOD	video on demand
WAN	wide area network
x86	a family of architectures based on the Intel 8086 CPU