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

- Consumer Broadband Access
- Broadband Access Networks
- Internet Backbone Networks
- Research Networks
- Enterprise Networks
- Corporate Data Centers and Enterprise (High Performance Computing)
- Internet eXchange and Interconnection Points
- Content Providers
- Research, Education and Government Facilities (High Performance Computing)
Why Higher Speed Ethernet?

Fundamental bottlenecks are happening everywhere

\[ \text{Increased # of users} + \text{Increased access rates and methods} + \text{Increased services} = \text{Bandwidth explosion everywhere} \]

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: The next speed of Ethernet must begin when 100GbE done!
- 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

Summary of Data Submissions

- Scott Kipp, Brocade, “Data Center Bandwidth Scenarios”
  - [Link](http://www.ieee802.org/3/ad_hoc/bwa/public/may11/kipp_01_0511.pdf)
  - [Link](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”
  - [Link](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”
  - [Link](http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/cloonan_01a_0911.pdf)
- Scott Kipp, Brocade, “Storage Growth and Ethernet”
  - [Link](http://www.ieee802.org/3/ad_hoc/bwa/public/sep11/kipp_01a_0911.pdf)
  - [Link](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”
  - [Link](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”
  - [Link](http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/huang_01_1111.pdf)
- Henk Steenman, AMS-IX / Euro-IX, “The European Peering Scene”
  - [Link](http://www.ieee802.org/3/ad_hoc/bwa/public/nov11/steenman_01_1111.pdf)
- Lone Hansen, BSRIA, “Global Data Centres Presentation IEEE”
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

More Devices

Key Growth Factors

More Internet Users

Speed Increasing

More Rich Media Content

Source: nowell_01_0911.pdf citing Cisco Visual Networking Index (VNI) Global IP Traffic Forecast, 2010–2015,
2015 Global Users and Network Connections

North America
288 Million Users
2.2 Billion Networked Devices

Western Europe
314 Million Users
2.3 Billion Networked Devices

Central/Eastern Europe
201 Million Users
902 Million Networked Devices

Japan
116 Million Users
727 Million Networked Devices

Latin America
260 Million Users
1.3 Billion Networked Devices

Middle East & Africa
495 Million Users
1.3 Billion Networked Devices

Asia Pacific
1330 Million Users
5.8 Billion Networked Devices


IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc
July 2012 IEEE 802 Plenary, San Diego, CA, USA
Global Broadband Speed 2010-2015
Average broadband speed will grow 4X; from 7 to 28 Mbps

IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc
July 2012 IEEE 802 Plenary, San Diego, CA, USA
US Cable Industry: Maximum Permitted Bandwidth Trends (Downstream)

The past 29-years show a constant bandwidth increase of ~1.5x every year...

Example: Traffic Generation Comparison

Bandwidth Generation Compared to a 32 bit based laptop

<table>
<thead>
<tr>
<th>Device</th>
<th>Traffic multiplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tablet</td>
<td>1.1</td>
</tr>
<tr>
<td>64-bit Laptop/PC</td>
<td>1.9</td>
</tr>
<tr>
<td>Internet enabled HDTV</td>
<td>2.9</td>
</tr>
<tr>
<td>Gaming console</td>
<td>3.0</td>
</tr>
<tr>
<td>Internet enabled 3D TV</td>
<td>3.2</td>
</tr>
</tbody>
</table>

Global IP Traffic by Local Access Technology


IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc
July 2012 IEEE 802 Plenary, San Diego, CA, USA
Regional contributions to the Zettabyte journey

North America
22.3 EB/Month by 2015
26% CAGR, 3X Growth

Western Europe
18.9 EB/Month by 2015
32% CAGR, 4X Growth

Central/Eastern Europe
3.7 EB/Month by 2015
39% CAGR, 5X Growth

Japan
4.8 EB/Month by 2015
27% CAGR, 3X Growth

Latin America
4.7 EB/Month by 2015
48% CAGR, 7X Growth

Middle East & Africa
2.0 EB/Month by 2015
52% CAGR, 8X Growth

Asia Pacific
24.1 EB/Month by 2015
35% CAGR, 4X Growth

Example: Financial Sector

Usage growth

Bandwidth Growth

Bandwidth Growth is throughout the Eco-system

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

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

Total Digital Data

Entered the Zettabyte Era

<table>
<thead>
<tr>
<th>Storage (Exabytes)</th>
<th>2005</th>
<th>2010</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>130</td>
<td>1227</td>
<td>7910</td>
</tr>
</tbody>
</table>

Growth over Next Decade

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td># of Servers</td>
<td>x10</td>
</tr>
<tr>
<td>Storage</td>
<td>x50</td>
</tr>
<tr>
<td># of Files</td>
<td>x75</td>
</tr>
</tbody>
</table>

Consider the implications!


IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc
July 2012 IEEE 802 Plenary, San Diego, CA, USA
Storage Access Methods

SAN
Ethernet Switch
FCoE Switch
8G, 16G, 32G
Fibre Channel Switch

iSCSI Storage Arrays
Fibre Channel Over Ethernet Storage Arrays
Fibre Channel Storage Arrays
Fibre Channel Tape Library

IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc July 2012 IEEE 802 Plenary, San Diego, CA, USA
Growth in External Storage Subsystems

Ethernet Based Storage NAS, iSCSI and FCoE

Over 50% of storage expected to be Ethernet-based in 2015


IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc
July 2012 IEEE 802 Plenary, San Diego, CA, USA
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

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

Server Port Speed

Annual Server Ports Shipped

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

<table>
<thead>
<tr>
<th>I/O per server (Gb/s)</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>40</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>Servers / rack</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Bandwidth / rack (Gb/s)</td>
<td>200</td>
<td>400</td>
<td>800</td>
<td>1600</td>
<td>3200</td>
</tr>
</tbody>
</table>

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

Cluster Aggregation in Data Center

Cluster or POD of Racks

Cluster Traffic to Core in Tb/s

Data Center

- Cluster 2
- Cluster 3
- Cluster 4
- Cluster 1
- CORE
- Cluster 5

<table>
<thead>
<tr>
<th>Cluster bandwidth to core (Tb/s)</th>
<th>0.4</th>
<th>1</th>
<th>2</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clusters</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Bandwidth to core (Tb/s)</td>
<td>4</td>
<td>10</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>Bandwidth to WAN (Gb/s)</td>
<td>20</td>
<td>40</td>
<td>200</td>
<td>400</td>
</tr>
<tr>
<td>Oversubscription to WAN</td>
<td>200</td>
<td>250</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Example: Impact of Data Migration

Application migration are between servers, clusters or data centers

Primary Data Center
- Router
- WDM
- Server
- SAN Switch

Cloud Provider or Secondary Data Center
- WDM
- Router
- SAN Switch

Application A data needs to be mirrored before the application can move.

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

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

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

Euro-IX IXP Locations

132 IXPs
36 Countries
140 Cities
420 Locations

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

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

Five Year Peak European IXP Traffic Projections


IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc
July 2012 IEEE 802 Plenary, San Diego, CA, USA
Euro-IX European Member Port Usage

Notes: 10Mb almost zero, 100Mb strongly ↓, 1G starting ↓, 10G strongly ↑


IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc
July 2012 IEEE 802 Plenary, San Diego, CA, USA
Example: Traffic Planning

LONAP (London) traffic on a typical 2010 weekday

Wed. afternoon World Cup 2010 England vs Slovenia

Examples: Data Intensive Science

• CERN: Atlas detector in LHC (Large Hadron Collider) generates \(\sim 1\) petabyte/sec

• Genome sequencing:
  – Per-instrument data rate strongly \(\uparrow\) (~10x over 5 years)
  – Cost of sequencers strongly \(\downarrow\) (10x over 5 years)

• Futures: Square Kilometer Array (SKA)
  – \(\sim 2800\) receivers in telescope array
  – 2 petabytes/sec to central correlator
    • sending \(\sim 100\) Gb/s to analysis centers

ESnet Accepted Traffic (Petabytes/month)

Expecting 100 Petabytes/month of data in 2015


IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc
July 2012 IEEE 802 Plenary, San Diego, CA, USA
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

- Financial sector: fit to Figure 15, CAGR = 95%
- Science: fit to Figure 13, ESnet 2004 to 2011, CAGR = 70%
- Peering: fit to Figure 39, CAGR = 64%
- HSSG tutorial: Slide 22 core, CAGR = 58%
- Cable: Figure 20, CAGR = 50%
- HSSG tutorial: Slide 22 server I/O, CAGR = 36%
- IP traffic: Figure 2, CAGR = 32%
- Figure 15 NYSE historical data
- Figure 39 Euro-IX historical data

IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad Hoc
July 2012 IEEE 802 Plenary, San Diego, CA, USA
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:
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
<table>
<thead>
<tr>
<th>Prefix</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1GbE</td>
<td>1 Gb/s Ethernet</td>
</tr>
<tr>
<td>10GbE</td>
<td>10 Gb/s Ethernet</td>
</tr>
<tr>
<td>40GbE</td>
<td>40 Gb/s Ethernet</td>
</tr>
<tr>
<td>100GbE</td>
<td>100 Gb/s Ethernet</td>
</tr>
</tbody>
</table>

**Abbreviations:**
- **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  
PCle 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