400 Gigabit Ethernet
Call-For-Interest

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 for *400 Gb/s 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
  – “A Higher Speed - Overview,”.
  – “The Need for 400 Gb/s Ethernet,”.
  – “The Technical Viability of 400 Gb/s Ethernet,”.

• Discussion
• Call for Interest
• Future Work
A HIGHER SPEED - OVERVIEW

Presented by
IEEE 802.3 Working Group
Orlando, FL, USA
March 19, 2013
The Ethernet Eco-System (2007 HSSG)
The Ethernet EcoSystem Now

• Update with a figure that shows today’s Applications
• Smart pad
• Smart phone
• HDTV
• Wi-Fi
• Euro-IX (JD)
• Data Center (40GBASE-T)
• OTN (Andy)
• R&D (JD>Bennett)
• Content Providers (Andy / Dave O)
• Cloud (Amazon, Google)

Today’s Trends
• Cloud slide – anytime anywhere anyhow (Business / consumer)
• Mobile
• Outsourcing of data center / IT
• Thin Client is here
• Flat Networks
• Data Centers
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

Global Broadband Speed 2010-2015
Average broadband speed will grow 4X; from 7 to 28 Mbps

North America
3.7-Fold growth
7.5 to 27 Mbps

Western Europe
3.9-Fold growth
9.2 to 36 Mbps

Central/Eastern Europe
3.3-Fold growth
6.1 to 20 Mbps

Japan
4.1-Fold growth
15.5 to 64 Mbps

Latin America
2.9-Fold growth
2.8 to 8 Mbps

Middle East & Africa
2.5-Fold growth
2.8 to 7 Mbps

Asia Pacific
4.6-Fold growth
5.5 to 25 Mbps

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

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

Source: nowell_01_0911.pdf citing Cisco Visual Networking Index (VNI) Global IP Traffic Forecast, 2010–2015,
Findings of IEEE 802.3 BWA Ad Hoc


Discuss
Limitations of analysis
Technical feasibility
Decreasing cost per bit
Technology Roadmap
THE NEED FOR 400 GIGABIT ETHERNET

Presented by
IEEE 802.3 Working Group
Orlando, FL, USA
March 19, 2013
10GbE Server Deployments

Can we turn this into bandwidth?
Update with latest numbers

All data used with permission Seamus Crehan, Crehan Research.
Examples - Events Drive Terabit Traffic

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.
400Gb/s vs. 4 x 100Gb/s Link Aggregation

- Traffic is often trunked into large tunneled flows
  - Insufficient entropy to do hashing efficiently
  - Link Aggregation (LAG) is inefficient
  - BW not considered which leads to flow imbalance
  - A faster interface provides predictable performance

- Sources of large flows:
  - Content distribution
  - Secure traffic

- Fewer items to manage provides operational efficiency
  - Bandwidth is growing exponentially
  - Without faster links, link count grows exponentially therefore management pain grows exponentially
Large flows result in individual links becoming congested and bundles losing efficiency.
Data Center Architecture Trend

Hierarchical Fat Tree architecture

Non-blocking architecture

400GbE need
400Gb/s Near-term Applications

- Core ⇔ Transport (400Gb/s Transport demonstrated)
- Core ⇔ Core
- Datacenter ⇔ Datacenter
- Datacenter upper layer switch interconnect (shown on previous slide)
THE TECHNICAL VIABILITY OF 400 GIGABIT ETHERNET

Presented by
IEEE 802.3 Working Group
Orlando, FL, USA
March 19, 2013
Matching Needs with Capabilities

“if you can’t always get what you want …”

“you find another way”
“Everything Axis”

Time Division Multiplexing
(i.e. Baud Rate)

Wavelength Division Multiplexing
(i.e. λs)

Modulation
(i.e. Bits per Hz)

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

100 Gbps
50 Gbps
25 Gbps
10 Gbps

10 Gbps
25 Gbps
50 Gbps
100 Gbps

8
16

4
8
12
16
Example: Anatomy of a 400GbE Optical PMD implementation

Definition of these interfaces…

… drives complexity of the module implementation
Electrical Interface Technology trade-offs

- Electrical width drives connector complexity
- Interface definition drives electronics complexity within module
- Advanced modulation now allows potential for reduced width

Region of Technical Feasibility

Module Implementation Complexity

Electronics implementation complexity 1b/sym
Electronics implementation complexity >1b/sym

400 Gb/s Ethernet Study Group CFI, V 0.2
Orlando, FL, USA
Optical Interface Technology trade-offs

- Packaging complexity drives cost
- Adoption of advanced modulation now allows trade-off of analog complexity vs. digital complexity (solid vs. dashed)
- Multiple technical feasibility data points exists within the green region
- Study Group/Task Force goal will be to converge towards optimum solution(s)

*Lane = wavelengths or fibers

**Region of Technical Feasibility**

- Module Implementation Complexity
- Optical Packaging complexity
- Electronics complexity 1b/sym
- Electronics complexity >1b/sym

400 Gb/s Ethernet Study Group CFI, V 0.2
Orlando, FL, USA
400Gb/s MAC Technical Feasibility

- 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
400Gb/s MAC Technical Feasibility, cont.

- Typical 100Gb/s MAC/PCS ASIC:
  - 45/40nm CMOS
  - 160b wide bus
  - 644MHz clock

- Potential 400Gb/s MAC/PCS ASIC:
  - 28/20nm CMOS
  - 400b wide bus
  - 1GHz clock

- 400Gb/s MAC/PCS FPGA will be feasible with wider buses and slower clocks
400 Gb/s Ethernet vs Terabit Ethernet
Technology Roadmap
Findings of IEEE 802.3 BWA Ad Hoc

The Need for Higher Speed

• Traffic is growing everywhere
  – The masses have more ways of faster access
  – Higher bandwidth content
  – New applications enabled

• Need for higher speed by multiple application
400Gb/s vs. Higher Rates

- Customers want parity in W/bit, $/bit, and bits/system
- Faster interface rates require exotic implementations
  - Not yet competitive per W, per $, or density
  - Higher R&D investment
  - Longer time to market
- 400GbE can reuse 100GbE building blocks
- 400GbE fits in the dense 100GbE system roadmap
- Data rates beyond 400Gb/s require an increasingly impractical number of lanes if 100GbE technology is reused
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
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 to form 400 Gb/s Ethernet SG on Thursday
• If approved
  – 802 EC informed of 400 Gb/s Ethernet SG on Friday
  – First 400 Gb/s Ethernet SG meeting, week of May 2013 IEEE 802.3 Interim.
THANK YOU!

March 19, 2013

DRAFT PRESENTATION
WORK IN PROGRESS
Contributors
Technology Axes

- Time Division Multiplexing (i.e. Baud Rate)
  - 100 Gbps
  - 50 Gbps
  - 25 Gbps
  - 10 Gbps

- Modulation (i.e. Bits per Hz)
  - 1 (e.g. NRZ)
  - 2 (e.g. PAM-4)
  - 4 (e.g. QAM-16)

- Space Division Multiplexing (i.e. Multiple Optical and Electrical Channels)
  - 4 optical channels
  - 8 electrical channels
  - 16 optical channels

- Wavelength Division Multiplexing (i.e. λs)
  - 4
  - 8
  - 12
  - 16

Graphically intense – need to find way to simplify
Time Division Axis

Time Division Multiplexing (i.e. Baud Rate)

100 Gbps
50 Gbps
25 Gbps
10 Gbps

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

Optical Channels
8
4
16

Electrical Channels
8
4
16

Wavelength Division Multiplexing (i.e. λs)

16
8
12
16

References /Demonstrations

- Blah
- Blah
- Blah
Modulation Axis

Time Division Multiplexing (i.e. Baud Rate)

Wavelength Division Multiplexing (i.e. λs)

Modulation (i.e. Bits per Hz)

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

4 8 16

References /Demonstrations
- Blah
- Blah
- Blah

400 Gb/s Ethernet Study Group CFI, V 0.2
Orlando, FL, USA
Wavelength Axis

- Time Division Multiplexing (i.e. Baud Rate)
  - 100 Gbps
  - 50 Gbps
  - 25 Gbps
  - 10 Gbps
- Modulation (i.e. Bits per Hz)
  - 1 (e.g. NRZ)
  - 2 (e.g. PAM-4)
  - 4 (e.g. PAM-16)
- Space Division Multiplexing (i.e. Multiple Optical and Electrical Channels)
  - 8 Channels
- Wavelength Division Multiplexing (i.e. λs)

References /Demonstrations
- Blah
- Blah
- Blah
Possible Approaches: “Add more Modules”

- Time Division Multiplexing (i.e. Baud Rate)
- Wavelength Division Multiplexing (i.e. λs)
- Modulation (i.e. Bits per Hz)
- Space Division Multiplexing (i.e. Multiple Optical and Electrical Channels)

- 100 Gbps
- 50 Gbps
- 25 Gbps
- 10 Gbps

- Electrical Channels
- Optical Channels

- 16
- 8
- 4

- 100GBASE-LR4
- CFP4

- Add Channels

- Modulation (i.e. Bits per Hz)
  - 4 (e.g. QAM-16)
  - 2 (e.g. PAM-4)

- 400G Interface

- DRAFT PRESENTATION
- WORK IN PROGRESS