100 GBE AND BEYOND

Greg Hankins <ghankins@brocade.com>

NANOG52
Agenda and What’s Covered in This Presentation

- Ethernet interface technology
  - Overview
  - 28 Gbps Common Electrical Interfaces (CEI)
  - New 100 Gbps Media Modules
  - 100 GbE Developments
  - Beyond 100 GbE...

- Optical technology developments are intentionally left out
  - Go see Drew Perkins’ talk tomorrow morning:
    “Dawn of the Terabit Age: Scaling Optical Capacity to Meet Internet Demand”

- Skipping router packet processing, lookup capabilities and memory architectures
  - Wire-speed 100 GbE is ~149 Mpps, or one packet every 6.7 ns at 64 byte frames
  - Maybe a topic for the next NANOG?
# Standards Organizations and You, Revisited

<table>
<thead>
<tr>
<th>Name</th>
<th>Primary Role (in Context of this Presentation)</th>
<th>Primary Players</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customers</td>
<td>Buy Your Services</td>
<td></td>
</tr>
<tr>
<td><strong>You</strong></td>
<td><em>Run Networks</em></td>
<td></td>
</tr>
<tr>
<td>Hardware Vendors</td>
<td>Make Equipment</td>
<td></td>
</tr>
<tr>
<td><strong>MEF</strong></td>
<td>Ethernet Service Definitions, Standards and Certification</td>
<td>Hardware Vendors, Network Operators</td>
</tr>
<tr>
<td><strong>IETF</strong></td>
<td>Higher Layer Protocol Standards</td>
<td>Hardware Vendors, Network Operators</td>
</tr>
<tr>
<td><strong>IEEE</strong></td>
<td>Ethernet Standards (802.1, 802.3)</td>
<td>Component and Hardware Vendors</td>
</tr>
<tr>
<td><strong>incits</strong></td>
<td>Fibre Channel Standards (T11)</td>
<td>Component and Hardware Vendors</td>
</tr>
<tr>
<td><strong>ITU</strong></td>
<td>Telecom Standards (SG15)</td>
<td>Component and Hardware Vendors</td>
</tr>
<tr>
<td><strong>10X10</strong></td>
<td>Optical Module Standards</td>
<td>Component and Hardware Vendors, Network Operators</td>
</tr>
<tr>
<td><strong>SFF Committee</strong></td>
<td>Media Module Standards</td>
<td>Component and Hardware Vendors</td>
</tr>
<tr>
<td><strong>OIF</strong></td>
<td>Component Interface Standards</td>
<td>Component and Hardware Vendors</td>
</tr>
</tbody>
</table>
Current State of the Industry

• There is already demand for other interfaces beyond the scope of IEEE 802.3ba (June 2010)

• Standard defines a flexible architecture that enables many implementations as technology changes

• New 40 GbE and 100 GbE standards are in progress
  • IEEE 802.3bg defined a 40 GbE serial interface to OTU3/STM-256/OC-768
  • The 2nd generation of 100 GbE will use 4 x 25 Gbps electrical and optical signaling
Current State of the Industry

• Fundamental 1\textsuperscript{st} generation technology constraints limits higher 100 GbE density and lower cost

• Electrical signaling inside the box
  • 100 Gbps Attachment Unit Interface (CAUI) uses 10 x 10 Gbps

• Optical signaling outside the box
  • 10x10 MSA: 10 x 10 Gbps
  • 100GBASE-LR4 and 100GBASE-ER4: 4 x 25 Gbps

• CFP module size and power consumption

CFP image courtesy of Finisar.
1st Generation vs 2nd Generation 100 GbE

2nd Generation 100 GbE Needs Faster Electrical Signaling

1st Generation 10 x 10 Gbps

2nd Generation 4 x 25 Gbps

10 Gbps Electrical Signaling and 10:4 Gearbox Adds Complexity, Cost, Space, Power...

25 Gbps Electrical and Optical Signaling

Source: http://grouper.ieee.org/groups/802/3/ba/public/jul08/cole_03_0708.pdf
## Front Panel Interface Density Trends

### Module Form Factor, Throughput and Power

<table>
<thead>
<tr>
<th>Year</th>
<th>Form Factor</th>
<th>Throughput</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010 SFP+ and QSFP</td>
<td>640 Gbps</td>
<td>60 W</td>
<td></td>
</tr>
<tr>
<td>2010 CFP</td>
<td>400 Gbps</td>
<td>80 W</td>
<td></td>
</tr>
<tr>
<td>2013+ CFP2</td>
<td>800 Gbps</td>
<td>80 W</td>
<td></td>
</tr>
<tr>
<td>2013+ 25 Gbps QSFP</td>
<td>2 Tbps</td>
<td>80+ W</td>
<td></td>
</tr>
<tr>
<td>2014+ CFP4</td>
<td>1.6 Tbps</td>
<td>80 W</td>
<td></td>
</tr>
<tr>
<td>2016+ ??</td>
<td>800 Gbps</td>
<td>80 W</td>
<td></td>
</tr>
</tbody>
</table>
# Key Industry Initiatives in 2011

Developing Technology for 2\textsuperscript{nd} Generation 100 GbE

<table>
<thead>
<tr>
<th>Initiative</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 Gbps Backplane and Copper Study Group</td>
</tr>
<tr>
<td>100 Gbps Interfaces Using 4 x 25 Gbps Electrical Signaling</td>
</tr>
<tr>
<td>Ethernet Bandwidth Assessment Ad Hoc</td>
</tr>
<tr>
<td>Lower Cost 10 x 10 Gbps Optical Modules</td>
</tr>
<tr>
<td>Next Generation Pluggable Media Module Form Factors</td>
</tr>
<tr>
<td>28 Gbps Electrical Interfaces</td>
</tr>
</tbody>
</table>
Agenda

• Overview
• 28 Gbps Common Electrical Interfaces (CEI)
• New 100 Gbps Media Modules
• 100 GbE Developments
• Beyond 100 GbE...
28 Gbps Common Electrical Interfaces (CEI)

• OIF is doing fundamental work on 28 Gbps electrical signaling which will make newer interfaces and pluggable media modules possible

• Lower power, Very Short Reach (VSR) 4” interfaces are being defined for several new applications
  • 1 lane for 32 Gbps Fibre Channel at 28.05 Gbps
  • 4 lanes for 100 GbE at 25.78125 Gbps
  • 16 lanes for 400 GbE at 25.78125 Gbps?

• CEI-28G-VSR is approaching technical stability, and is expected to be finished in January 2012
25 Gbps and 28 Gbps
Common Electrical Interfaces (CEI)

2. Chip to chip: CEI-28G-SR – 12”
3. Chip to module: CEI-28G-VSR – 4”
   (used by 2\textsuperscript{nd} generation 100 GbE media modules)
Agenda

• Overview
• 28 Gbps Common Electrical Interfaces (CEI)
• New 100 Gbps Media Modules
• 100 GbE Developments
• Beyond 100 GbE...
## 10 Gbps Module Review – 3 Generations of 10 GbE Over 7 Years

Each Module Increased Density, While Reducing Cost and Power

<table>
<thead>
<tr>
<th>Module Name (Images not to Scale)</th>
<th>1st Generation</th>
<th>2nd Generation</th>
<th>3rd Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>300PIN MSA</td>
<td>XENPAK</td>
<td>XPAK</td>
<td>X2</td>
</tr>
<tr>
<td>XFP</td>
<td>SFP+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Approximate Module Dimensions (Length x Width to Scale)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>300PIN MSA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Front Panel Density</th>
<th>1</th>
<th>4</th>
<th>8</th>
<th>8</th>
<th>16</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrical Interface</td>
<td>XSBI</td>
<td>XAUI</td>
<td>XAUI</td>
<td>XAUI</td>
<td>XFI</td>
<td>SFI</td>
</tr>
<tr>
<td>Electrical Signaling</td>
<td>16 x 644 Mbps</td>
<td>4 x 3.125 Gbps</td>
<td>4 x 3.125 Gbps</td>
<td>4 x 3.125 Gbps</td>
<td>1 x 10.3125 Gbps</td>
<td>1 x 10.3125 Gbps</td>
</tr>
</tbody>
</table>

Module images courtesy of Finisar.
# 100 Gbps Module Evolution

Two Generations of 100 GbE Expected to Take 5 Years

<table>
<thead>
<tr>
<th>Module Name (Images not to Scale)</th>
<th>1st Generation</th>
<th>2nd Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFP</td>
<td>CFP</td>
<td>CFP2</td>
</tr>
<tr>
<td>CXP</td>
<td>CXP</td>
<td>CFP4</td>
</tr>
<tr>
<td>25 Gbps QSFP</td>
<td>25 Gbps QSFP</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Approximate Module Dimensions (Length x Width to Scale)</th>
<th>1st Generation</th>
<th>2nd Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFP</td>
<td>4</td>
<td>22 - 44</td>
</tr>
<tr>
<td>CXP</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>25 Gbps QSFP</td>
<td>22 - 44</td>
<td>8</td>
</tr>
<tr>
<td>CFP2</td>
<td>22 - 44</td>
<td>16 - 32</td>
</tr>
<tr>
<td>CFP4</td>
<td>22 - 44</td>
<td>16 - 32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Front Panel Density</th>
<th>1st Generation</th>
<th>2nd Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>22 - 44</td>
<td>8</td>
</tr>
<tr>
<td>16</td>
<td>16 - 32</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Electrical Interface</th>
<th>1st Generation</th>
<th>2nd Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAUI</td>
<td>CAUI-4</td>
<td>CAUI-4</td>
</tr>
<tr>
<td>CPPI</td>
<td>CPPI-4</td>
<td>CPPI-4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Electrical Signaling (Gbps)</th>
<th>1st Generation</th>
<th>2nd Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 x 10</td>
<td>10 x 10</td>
<td>4 x 25</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>16 - 32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Media Type</th>
<th>1st Generation</th>
<th>2nd Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMF</td>
<td>Twinax, MMF</td>
<td>SMF</td>
</tr>
<tr>
<td>MMF/SMF?</td>
<td>SMF</td>
<td>SMF</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Advantages</th>
<th>1st Generation</th>
<th>2nd Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long Reach, High Power Dissipation</td>
<td>Long Reach, Higher Density</td>
<td></td>
</tr>
<tr>
<td>Small Size, Designed for Passive Cabling</td>
<td>Highest Density, Smaller Size,</td>
<td></td>
</tr>
<tr>
<td>Established Form Factor</td>
<td>Highest Density, Smaller Size,</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disadvantages</th>
<th>1st Generation</th>
<th>2nd Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Too Big</td>
<td>Short Reach, Too Small</td>
<td>Bigger Size</td>
</tr>
<tr>
<td>Limited Power Dissipation and Reach</td>
<td>Unproven Form Factor (vs. QSFP)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Availability (Subject to Change)</th>
<th>1st Generation</th>
<th>2nd Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>2010</td>
<td>2011 (InfiniBand)</td>
</tr>
<tr>
<td>2013+ (Ethernet)</td>
<td>2013+</td>
<td>2014+</td>
</tr>
</tbody>
</table>

CFP and CXP images courtesy of Finisar, QSFP image courtesy of the SFF Committee, SFP2 and SFP4 images courtesy of the CFP MSA.
100 Gbps CFP Module Evolution
Module Form Factor vs. Front Panel Density

100 Gbps Module Evolution
Graphical View of Module Form Factors

Diagram courtesy of Molex.
100 GbE Module Technologies Compared

Transmit Side of Module

100GBASE-LR4 10 km CFP

- Most expensive, complex and uses the most power
- Gearbox converts 10 x 10 Gbps electrical signaling into 4 x 25 Gbps signaling

CAUI 10 x 10 Gbps Electrical

10:4 Gearbox

1295 nm Laser
1300 nm Laser
1305 nm Laser
1310 nm Laser

MUX

1 TX Fiber With 4 Wavelengths of 25 Gbps

4 x 25 Gbps Optical

10x10 MSA 2 km, 10 km, 40 km CFP

- Less cost, complexity and power consumption
- Uses 10 x 10 Gbps electrical and optical signaling
- Doesn’t need the gearbox

CAUI 10 x 10 Gbps Electrical

Array of 10 Lasers in 1550 nm Range

MUX

1 TX Fiber With 10 Wavelengths of 10 Gbps

10 x 10 Gbps Optical

100GBASE-LR4 10 km CFP2

- Lower cost, complexity and power consumption
- Uses 25 Gbps electrical and optical signaling
- Doesn’t need the gearbox

CAUI-4 4 x 25 Gbps Electrical

10:4 Gearbox (Optional)

1295 nm Laser
1300 nm Laser
1305 nm Laser
1310 nm Laser

MUX

1 TX Fiber With 4 Wavelengths of 25 Gbps

CAUI-4 4 x 25 Gbps Electrical
4 x 25 Gbps Optical
Agenda

• Overview
• 28 Gbps Common Electrical Interfaces (CEI)
• New 100 Gbps Media Modules
• 100 GbE Developments
• Beyond 100 GbE...
Recent 100 GbE Developments

• 2nd generation projects based on 4 x 25 Gbps electrical signaling have started

• New IEEE Copper Study Group was approved in November, 2010
  • 100GBASE-KR4: 4 x 25 Gbps over backplane
  • 100GBASE-CR4: 4 x 25 Gbps over copper cable
  • http://www.ieee802.org/3/100GCU/index.html
Recent 100 GbE Developments

• 10x10 MSA is growing and working on several projects
  • Up to 25 members including AMS-IX, Facebook and Google
  • Finishing 10x10-10km and 10x10-40km standards, expected to be approved in July, 2011
  • Investigating muxing 8 bands of 40 km links to carry 8 x 100 Gbps over a single fiber pair

• IEEE is expected to start work in July, 2011 to define new interfaces that are expected to be available in 2013+
  • 100GBASE-SR4: 4 x 25 Gbps over OM3 MMF
  • 100GBASE-FR4: 4 x 25 Gbps over SMF for 500 m – 2 km
  • CAUI-4: electrical signaling to the CFP2
  • CPPI-4: electrical signaling to the 25 Gbps QSFP and CFP4
  • 25 Gbps QSFP and CFP2/4 will be competing for the highest front panel density
Putting it All Together – 100 GbE Line Card Architectures

1st Generation 10 Gbps and 25 Gbps Signaling

- 100GBASE-SR10
  - 100 m OM3
  - 150 m OM4 MMF
- 10x10-2km
  - 2 km SMF
- 100GBASE-LR4
  - 10x10-10km
  - 10 km SMF
- 100GBASE-ER4
  - 10x10-40km
  - 40 km SMF

2nd Generation 25 Gbps Signaling

- CFP SR Module
- CFP FR Module
- CFP LR Module
- CFP ER Module
- ASIC
- CAUI 10 x 10 Gbps Electrical

- CFP2 SR Module
- CFP2 FR Module
- CFP2 LR Module
- CFP2 ER Module
- ASIC
- CAUI-4 4 x 25 Gbps Electrical

- 100GBASE-SR4
  - 100 m OM3
- 100GBASE-FR4
  - 2 km SMF
- 100GBASE-LR4
  - 10 km SMF
Ethernet Average Selling Price (ASP) Trends

- Prices of previous Ethernet generations fell significantly during the first few years on the market.
- Already seeing a similar trend as 1\textsuperscript{st} generation 100 GbE volume increases, expect 2\textsuperscript{nd} generation 100 GbE to deliver significantly lower prices.

Data sources: Dell'Oro and Infonetics
## 100 GbE Technology Summary

<table>
<thead>
<tr>
<th>Physical Layer Reach</th>
<th>1? m Backplane</th>
<th>5+? m Copper Cable</th>
<th>7 m Copper Cable</th>
<th>100? m OM3 MMF</th>
<th>100 m OM3, 150 m OM4 MMF</th>
<th>2 km SMF</th>
<th>10 km SMF</th>
<th>40 km SMF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>100GBASE-KR4</td>
<td>100GBASE-CR4</td>
<td>100GBASE-CR10</td>
<td>100GBASE-SR4</td>
<td>100GBASE-SR10</td>
<td>10x10-2km</td>
<td>100GBASE-FR4</td>
<td>10x10-10km</td>
</tr>
<tr>
<td>Standard Status</td>
<td>Possible</td>
<td>Possible</td>
<td>2010 IEEE 802.3ba</td>
<td>Possible Future IEEE</td>
<td>2010 IEEE 802.3ba</td>
<td>2011 Future IEEE</td>
<td>Possible Future IEEE</td>
<td>2010 IEEE 802.3ba</td>
</tr>
<tr>
<td>Generation</td>
<td>2nd</td>
<td>2nd</td>
<td>1st</td>
<td>2nd</td>
<td>1st</td>
<td>2nd</td>
<td>1st</td>
<td>1st</td>
</tr>
<tr>
<td>Electrical Signaling (Gbps)</td>
<td>4 x 25</td>
<td>4 x 25</td>
<td>10 x 10</td>
<td>4 x 25</td>
<td>10 x 10</td>
<td>4 x 25</td>
<td>10 x 10</td>
<td>10 x 10</td>
</tr>
<tr>
<td>Media Signaling (Gbps)</td>
<td>4 x 25</td>
<td>4 x 25</td>
<td>10 x 10</td>
<td>4 x 25</td>
<td>10 x 10</td>
<td>4 x 25</td>
<td>10 x 10</td>
<td>10 x 10</td>
</tr>
<tr>
<td>Media Type</td>
<td>Backplane</td>
<td>Twinax</td>
<td>Twinax</td>
<td>MPO MMF</td>
<td>MPO MMF</td>
<td>Duplex SMF</td>
<td>Duplex SMF</td>
<td>Duplex SMF</td>
</tr>
<tr>
<td>Media Module</td>
<td>Backplane</td>
<td>25 Gbps QSFP, CFP2, CFP4</td>
<td>CXP</td>
<td>25 Gbps QSFP, CFP2, CFP4</td>
<td>CXP, CFP</td>
<td>25 Gbps QSFP, CFP2, CFP4</td>
<td>CFP</td>
<td>25 Gbps QSFP, CFP2, CFP4</td>
</tr>
</tbody>
</table>
Agenda

• Overview
• 28 Gbps Common Electrical Interfaces (CEI)
• New 100 Gbps Media Modules
• 100 GbE Developments
• Beyond 100 GbE...
Bandwidth Requirements Projection
All Solutions are Good, Fast, or Cheap – Pick Any Two

Beyond 100 GbE: Industry Challenges

2nd Generation 100 GbE and Higher Speeds

Technical Feasibility

Market Requirements

Economics Dictate the Solution

IEEE Provides an Open Industry Forum to Make Decisions
IEEE Ethernet Standards Timelines

- 8 years between 10 GbE and 100 GbE standards
- We need to start immediately in order to finish a new Ethernet speed standard by 2016

Diagram source: http://www.euro-ix.net/download/48/948
IEEE Ethernet Bandwidth Assessment
Ad Hoc

• Laying groundwork and investigating industry interest for the next Ethernet speed
  • Evaluate Ethernet wireline bandwidth requirements
  • Provide data and reference material to the IEEE
  • Gather information only, will not make a recommendation

• Web page:

• Mailing list:
IEEE Ethernet Bandwidth Assessment
Ad Hoc

• Network operator input is needed on future requirements
  • Speed, density, distance, cost, topology, anything really

• Presentations can be given on conference calls or at meetings, schedule is opportunistic

• Please get involved... this means **you**!!

• Request for data:

• Ad Hoc Chair contact:
  John D’Ambrosia, <jdambrosia@ieee.org>
Future 100 GbE Projects

• In the short term, 4 x 25 Gbps electrical and optical interfaces will keep the IEEE 802.3 Working Group busy for 2+ years

• 100 GbE serial is still not feasible in the near future
  • 25 Gbps signaling is challenging
  • We’ll get a better idea of what is possible as 25 Gbps technology matures

• 3rd generation 100 GbE is likely to be developed several years from now
Next Higher Speed Ethernet
250 GbE, 300 GbE, 400 GbE, or TbE?

• Using 10 x 25 Gbps signaling the next speed could be 250 GbE
  • The industry wants a larger jump

• 12 x 25 Gbps signaling matches the number of fibers in a high density MMF cable for 300 GbE
  • Unpopular too

• The likely candidate for the next speed is 400 GbE using 16 x 25 Gbps signaling
  • 16 x 25 Gbps wavelengths can be easily muxed/demuxed onto one SMF
  • MMF solutions would need 32 fibers in a high density cable MPO/MTP assembly
  • Evolution to 10 x 40 Gbps signaling

• TbE is simply impractical in the near future
  • 40 x 25 Gbps lanes in and 40 x 25 Gbps lanes out would make a gigantic media module
  • 40 Gbps serial lanes aren’t expected to be economical until after 2016, and will take considerable work as electrical losses grow exponentially with super high frequency signaling
400 GbE Module

- The 400 GbE module could be 16 channels wide and would be larger than the current 100 GbE CFP
Summary

• The 1st generation of 100 GbE uses 10x10 Gbps electrical lanes and large CFP media modules

• The 2nd generation of 100 GbE will use 4x25 Gbps electrical lanes and smaller CFP2/CFP4/25 Gbps QSFP modules

• Industry is working on 2nd generation 100 GbE for the next few years

• 400 GbE work may start in 2013+ and could finish by 2016+

• TbE is currently technically and economically unfeasible until 40 Gbps electrical lanes are defined after 2013 with a possible standard following many years later
Questions?