The 400GE (R)evolution
Ethernet for the next decade

September 2018
Neutral Peering Days, The Hague, The Netherlands

Dirk van den Borne
“What’s Driving Network Traffic Growth?

“Video killed the radio star”

We're getting addicting to the small screen... watching ever more (and higher definition) video

Daily hours spend with digital media per adult user

<table>
<thead>
<tr>
<th>Year</th>
<th>Mobile</th>
<th>Desktop / Laptop</th>
<th>Other Connected Devices</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>2.2</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>2009</td>
<td>3.0</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>2010</td>
<td>3.2</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>2011</td>
<td>3.7</td>
<td>0.8</td>
<td>0.3</td>
</tr>
<tr>
<td>2012</td>
<td>4.3</td>
<td>1.6</td>
<td>0.3</td>
</tr>
<tr>
<td>2013</td>
<td>4.9</td>
<td>2.3</td>
<td>2.3</td>
</tr>
<tr>
<td>2014</td>
<td>5.1</td>
<td>2.6</td>
<td>2.8</td>
</tr>
<tr>
<td>2015</td>
<td>5.4</td>
<td>5.4</td>
<td>5.4</td>
</tr>
<tr>
<td>2016</td>
<td>5.6</td>
<td>3.1</td>
<td>3.1</td>
</tr>
<tr>
<td>2017</td>
<td>5.9</td>
<td>3.3</td>
<td>3.3</td>
</tr>
</tbody>
</table>

From: Mary Meeker, Internet trends 2018, May 2018

Adapted from: Ericsson Mobility Report, June 2018
WHAT’S DRIVING DATA CENTER GROWTH?

The AI revolution

AI needs a lot of data for training and interference... vastly increasing networking bandwidth

Object Detection: Performance vs. Training Dataset Size

Google Research & Carnegie Mellon, 2017

From: Mary Meeker, Internet trends 2018, May 2018
400GE ADOPTION STARTING IN 2020 / 2021

400GE will be big, but it will take time

Adapted from Lightcounting, March 2018 High Speed Ethernet Optics Report
400GE ADOPTION STARTING IN 2020 / 2021

Ethernet transceiver sales across the industry

Adapted from Lightcounting, March 2018 High Speed Ethernet Optics Report
IEEE study group: August 2012

New group to set goals for next-gen Ethernet

Rick Merritt
8/20/2012 04:01 AM EDT
1 Comment

SAN JOSE – A new group will tackle an issue that’s been a hot debate in communications for at least two years—should the next big leap for Ethernet be to 400 Gbits/second or a Terabit/s. The IEEE 802.3 Industry Connections Higher Speed Ethernet Consensus group aims within a year to pick a target and kick off a formal standards effort.


IEEE 802.3bs standardization

802.3bs standard: December 6th, 2017

400G (and 200G) ETHERNET

400GE required fundamentally new technologies to enable a reduction of cost, power, and number of electrical/optical lanes
4-LEVEL PULSE AMPLITUDE MODULATION (PAM4)

Multi-level modulation

Electrical I/O:
- Electrical I/O has a strong frequency-dependent channel loss (“Skin effect”).
- Cost and power scale with trace length and baud rate.
- PAM4 enables 50G links with 25 Gbaud symbol rate, enabling longer and lower power electrical I/O.

Optical I/O:
- Pluggable optics cost and power scale with number of optical Tx/Rx.
- PAM4 halves the number of optical lanes required, enabling lower cost and power pluggable optics.

PAM4 needs (strong) FEC to close link budget on both electrical and optical I/O
IEEE KP4 FEC integrated into packet switching ASICs and/or pluggable optics.
### 400GE Optical Interfaces

Optical interfaces standardized as part of IEEE 802.3bs

<table>
<thead>
<tr>
<th>Reach</th>
<th>Media</th>
<th>16 x 25G optics</th>
<th>8 x 50G optics</th>
<th>4 x 100G optics</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR (100 m)</td>
<td>Parallel MMF</td>
<td>400GBASE-SR16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DR (500 m)</td>
<td>Parallel SMF</td>
<td>400GBASE-DR4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR (2 km)</td>
<td>Duplex SMF</td>
<td>400GBASE-FR8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR (10 km)</td>
<td>Duplex SMF</td>
<td>400GBASE-LR8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER (40 km)</td>
<td>Duplex SMF</td>
<td></td>
<td>400GBASE-ER8</td>
<td>-None-</td>
</tr>
</tbody>
</table>

**LAN:** 800-GHz spaced LAN-WDM grid

<table>
<thead>
<tr>
<th>Reach</th>
<th>Media</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR16</td>
<td>16 MMF x 25G NRZ</td>
<td>70-100 m</td>
</tr>
<tr>
<td>DR4</td>
<td>4 SMF x 100G PAM4</td>
<td>500 m</td>
</tr>
<tr>
<td>FR8</td>
<td>1 SMF x 8 λ LAN x 50G PAM4</td>
<td>2 km</td>
</tr>
<tr>
<td>LR8</td>
<td>1 SMF x 8 λ LAN x 50G PAM4</td>
<td>10 km</td>
</tr>
</tbody>
</table>
### POWER CONSUMPTION

400GE pluggable optics

<table>
<thead>
<tr>
<th>Optics</th>
<th>Estimated power consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>100GE</td>
<td>LR4</td>
</tr>
<tr>
<td></td>
<td>3.5 ~ 4W</td>
</tr>
<tr>
<td>400GE</td>
<td>SR16 / SR8 / SR4.2</td>
</tr>
<tr>
<td></td>
<td>8 ~ 10W</td>
</tr>
<tr>
<td>400GE</td>
<td>DR4 / FR4</td>
</tr>
<tr>
<td></td>
<td>10 ~ 12W (initially), 8 ~ 10W (longer-term)</td>
</tr>
<tr>
<td>400GE</td>
<td>FR8 / LR8</td>
</tr>
<tr>
<td></td>
<td>12 ~ 15W (initially), 10 ~ 12W (longer-term)</td>
</tr>
</tbody>
</table>

**Cooling and thermal line card design is critical for 400GE:**
- High quality heat sinks
- Optimize conductivity between module and cage.
- Optimized airflow

---

Pluggable optics power consumption is not reducing as fast as slot densities are increasing, making it critical for next-generation routers and switches.

© 2018 Juniper Networks

Packet-Optical evolution: Rise of the Pluggable, September 2018
100G PAM4 optical signaling will be the mainstream technology for 400GE:
- Fewer wavelengths reduces power consumption and cost.
- Need for stronger DSP equalization

IEEE specified the framework for 100G PAM4 optical signaling for use in 100GE and 400GE:
- 100GBASE-DR over 500 meters duplex SMF in 802.3cd.
- 400GBASE-DR4 over 500 meters parallel SMF in 802.3bs.

The 100G Lambda MSA extends the optical specification to 2 km and 10 km interfaces:
- 100G-FR over 2 km and 100G-LR over 10 km.
- 400G-FR4 over 2 km. 400G-LR4 over 10 km under discussion.

http://100glambda.com/

From: M. Li et al., Photonic Research, Feb. 2018
Form factor requirements for 400G pluggable optics have evolved as a result of 400GE standardization and industry development.

- Evolution of CFP2
  - 16-lane electrical interface
  - Supports >20W power profile

- Evolution of QSFP+ / QSFP28
  - 8-lane electrical interface
  - Supports 12~15W power profile

- New 400G-optimized design
  - 8-lane electrical interface
  - Supports 15W power profile

Top-to-bottom: CFP8, OSFP and QSFP-DD
400G PLUGGABLE FORM FACTOR
QSFP-DD (Double Density)

- 8 x 50G PAM4
- 400G AUI-8
- C2M interface

- 36 x QSFP-DD
- on a 1RU line card
- (14.4T / slot)

- Backwards-compatible to QSFP+ / QSFP28

- 8 x 50G PAM4 &
- 4 x 100G PAM4
- optics

- Power profile supports
- up to 12W / 15W*

- Front-to-back &
- side-to-side airflow
- /w riding heatsink

* With appropriate line card and cooling design
COHERENT DWDM PLUGGABLE OPTICS
for the QSFP ecosystem

QSFP28 is today the universal form factor of choice:
- QSFP28 ports are generally designed for a power consumption < 4~5 W.
- Coherent technology scales down in footprint and power over time, but it will still take many years to meet this target.

QSFP-DD is the next-generation universal form factor:
- Designed to support 200G / 400G pluggable optics, which have much higher power consumption (< 12~14W).
- Innovation in coherent DWDM technology is now focused on 400G DWDM small form factor pluggables → 400G DCI market.
- The same pluggables will also support 300G / 200G / 100G modes to address access, metro and regional markets.

Differentiation between client and line interfaces will largely disappear once the form factors become identical and you can mix-and-match on the same line card.
### EVOLUTION OF 400GE OPTICAL INTERFACES

**Industry initiatives to leverage higher baud rate optics**

<table>
<thead>
<tr>
<th>Reach</th>
<th>Media</th>
<th>16 x 25G optics</th>
<th>8 x 50G optics</th>
<th>4 x 100G optics</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR (100 m)</td>
<td>Parallel MMF</td>
<td>400GBASE-SR16 → SR8* / SR4.2*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DR (500 m)</td>
<td>Parallel SMF</td>
<td></td>
<td>400GBASE-DR4</td>
<td></td>
</tr>
<tr>
<td>FR (2 km)</td>
<td>Duplex SMF</td>
<td>400GBASE-FR8 → FR4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR (10 km)</td>
<td>Duplex SMF</td>
<td>400GBASE-LR8 → LR4 / 400ZR**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER (40 km)</td>
<td>Duplex SMF</td>
<td></td>
<td></td>
<td>-none-</td>
</tr>
</tbody>
</table>

---

**IEEE standard in progress**

**OIF 400ZR, 1 x 400G optics**

---

[Diagram showing 400GE optical interfaces with reach and media specifications.]

---

© 2018 Juniper Networks
400G PLATFORMS
Setting new benchmarks in density and power efficiency

Dense 400GE line cards and platforms across QFX, PTX and MX series will enable slot capacities up to 14.4T per slot/RU

CONCLUSIONS
The emerging 400GE ecosystem

400GE technology and building blocks are rapidly maturing:
- 1\textsuperscript{st}-gen 400G use 50G PAM4 electrical and optical signaling. Initial pluggable optics are now available, but CFP8 adoption is very limited.
- 2\textsuperscript{nd}-gen 400G using 100G PAM4 optical signaling will enable lower power and more future-proof 400G designs.

400GE increases routers and switches capacity up to 14.4T per slot/RU:
- QSFP-DD form factor extends successful QSFP ecosystem to 400G.
- ...but power consumption of optical modules becomes one of the main design challenges for dense 400G-optimized line cards and systems.
- 400GE interfaces with long-term relevance and widespread adoption are expected to be:
  - DR4 $\rightarrow$ 500 meters parallel SMF and 4 x 100GE break-out
  - FR4 $\rightarrow$ 2 km duplex SMF
  - 400ZR $\rightarrow$ beyond 2 km and DWDM transport
Q & A

Thank you for your time

Dirk van den Borne
dvandenborne@juniper.net