400G and 800G Ethernet and Optics

Andreas Bechtolsheim
Arista Networks, Inc
Founder, Chief Development Officer and Chairman
The upcoming 100G-400G Transition
The Easiest Way to go Faster is to go Faster

Ethernet Speed Transitions have been the primary driving force to improve the throughput and the price-performance of data center networks
For a new technology to ramp quickly, it must be more cost-effective than the previous technology it displaces.

Source: Brad Booth and Tom Issenhuth Microsoft, IEEE 802.3 400G
40G - 100G - 400G Switch Port Transition

Source: Dell'Oro Group July 2018 Ethernet Switching Forecast
40G to 100G Ethernet Transition [Ports]

100G went from < 10% to > 50% in one year

Source: Dell’Oro Market Research, Ethernet Switch Update, July 2018
400G ramp is slower than 100G for at least three key reasons:

1. Availability of new 400G optics
2. Availability of new 400G switches
3. Qualification of new systems

Source: Dell’Oro Market Research, Ethernet Switch Update, July 2018
Vast Majority of 400G Will be Deployed in Cloud
Expected 100G to 400G Bandwidth Cross-Over

Source: Dell’Oro Group July 2018 Ethernet Switching Forecast
Expected Transition from 50G to 100G SERDES

Source: 650 Group LLC, December 2018
Merchant Switch Silicon Bandwidth Growth

[Graph showing bandwidth growth from 2008 to 2019 with labels for 2008 (2.66X/3Y), 2011 (2X/2Y), 2013 (2.5X/2Y), 2015 (2X/2Y), 2017 (2X/2Y), and 2019.]
Jericcho VOQ Big Buffer Bandwidth Per Chip

5X Increase in Throughput per VOQ Chip

[Bar chart showing bandwidth from 2017 to 2019 with a 5X increase from J1+ to J2]
400G Datacenter Optics
### 400G Datacenter Optics Standards

<table>
<thead>
<tr>
<th>Name</th>
<th>Fiber</th>
<th>Reach</th>
<th>Modulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>400G-ZR/ZR+</td>
<td>Duplex SMF</td>
<td>10km-1000km</td>
<td>16-QAM</td>
</tr>
<tr>
<td>400G-FR4/LR4</td>
<td>Duplex SMF</td>
<td>2km/10km</td>
<td>100G-PAM4</td>
</tr>
<tr>
<td>400G-DR4</td>
<td>8xSMF</td>
<td>500m/2km</td>
<td>100G-PAM4</td>
</tr>
<tr>
<td>400G-SR8</td>
<td>16xMMF</td>
<td>50m</td>
<td>50G-PAM4</td>
</tr>
<tr>
<td>400G-CR8</td>
<td>copper</td>
<td>3m</td>
<td>50G-PAM4</td>
</tr>
</tbody>
</table>

Arista will support all 400G Optics that are relevant in market
Transition of Cloud Networks from 100G to 400G

- 100G-DCO → 400G-ZR
- 100G-LR4 → 400G-LR4
- 100G-CWDM4 → 400G-FR4
- 100G-PSMF4 → 400G-DR4
- 100G-AOC → 400G-AOC
- 100G-CR4 → 400G-CR8
- 4X the Network Capacity with same network topology
The most interesting new optics: 400G-ZR and 400G-ZR+
What is 400G-ZR/ZR+?

- Industry’s First Multi-vendor DWDM Standard
- Coherent, Tunable, Pluggable DCO Module
- 400G, 300G, 200G and 100G speeds
- Dense Client Optics Formfactor
- Supports 14.4 Tbps per 1U
- Max 20W power for 400G-ZR+
Order of Magnitude Cost Reduction

Source of Chart:: Dell’Oro DWDM Update July 2018

400G: $100/Gbps
400G-ZR: $10/Gbps
## 400G-ZR Standards and Reach

<table>
<thead>
<tr>
<th>Client Interface</th>
<th>Framing/FEC</th>
<th>Modulation</th>
<th>Reach</th>
</tr>
</thead>
<tbody>
<tr>
<td>400GE</td>
<td>OIF 400ZR</td>
<td>400G-16QAM</td>
<td>Up to 1000km</td>
</tr>
<tr>
<td>2x200GE</td>
<td>OpenROADM 2.1</td>
<td>300G-8QAM</td>
<td>Up to 2000km</td>
</tr>
<tr>
<td>4x100GE</td>
<td>OpenROADM 3.1</td>
<td>200G-QPSK</td>
<td>Up to 4000km</td>
</tr>
<tr>
<td>OTU4</td>
<td>ITU G709.2</td>
<td>100G-QPSK</td>
<td>Up to 8000km</td>
</tr>
<tr>
<td></td>
<td>ITU G709.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IEEE 802.3ck</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Use Cases for 400G-ZR/ZR+

- DCI (Datacenter Interconnect)
- Metro-Reach DWDM Networks
- Long-Reach DWDM Networks
- 5G Aggregation
- Cable R-PHY Aggregation
400G-ZR+ Covers Most of USA with 400G DWDM

Internet-2 100G Research Network (May 2017)
400G-ZR+ Covers all of Europe with 400G-DWDM

400G-ZR+ Covers Most of Asia
Why Pluggable Form Factor?

Customers Can Source 400G-ZR Modules Directly
  -> Avoids Margin Stacking

System Vendor Can Build One System Design
  -> No extra Investment Required to Deliver DCO

Customers Can Mix and Match DCO and Client Optics
  -> Easy configurability and easy field replacement

Multiple SKUs Expected
  -> 10km-100km-300km-1000km, high-output-power, etc
Pluggable DCO Form Factor Transition to OSFP

- **200G-16QAM 1000km**
  - 16 per 1U (3.2T)

- **100G-2PAM56 100km**
  - 36 per 1U (3.6T)

- **400G-16QAM 1000km**
  - 36 per 1U (14.4T)
Roadmap to 800G-ZR

- 800G-16QAM Feasible with 120 Gbaud
- Same Pluggable Formfactor (800G-OSFP)
- Targeting Same Power Envelope ~ 20W
- Double the Power Efficiency per bandwidth
- Significantly improved price-performance
- Backward compatible with 400G-ZR/ZR+
400G-ZR/ZR+ Summary

- First True Multi-vendor Interoperable DCO Standard
- Revolutionary Price-Performance
- Very High Density: 14.4T per 1U
- Very Low Power: 20W for 400G 1000km Reach
- High-density Pluggable Formfactor
- Eliminates Separate Transport Shelf
- Eliminates Special DCO System Designs
- Roadmap to 800G-ZR/ZR+ in 2022/2023
400G and 800G Optics Module Form Factors
The OSFP (Octal Small Form Factor Pluggable)

**High Port Density:** Up to 36 per 1U  
28.8T with 8x100G SerDes

**High Thermal Capability**  
Up to 20W Power Capability

**Backward Compatible with QSFP**  
With Simple OSFP-QSFP Adaptor
The QSFP-DD (QSFP Double Density)

- Eight Lanes at 56G-PAM4
  - Supports 400G with 8x50G lanes

- Port Density: 36 per 1U
  - 14.4 Tbps per 1U

- Dual Row Connector Design
  - Challenging to support 112G

- Thermal Limitations
  - Difficult to support > 15W

QSFP-DD Type 1 and Type 2 form factors.
Pluggable Form Factors Comparison

36 Port Density per 1U

20W Thermal Capacity for 400G-ZR+ and 800G

High Signal Integrity for 112G-PAM4 SerDes

OSFP is the right good choice for ZR+ and 800G (Dual 400G)
112G-PAM4 SerDes Demonstration (OFC 2018)

10 inch overall channel

10” Trace Channel plus OSFP Connector
24db Insertion Loss Die to Die, 16db Ball-to-Ball
Measured BER= approx. 6x10^{-7}

>25dB S/N @ 28Ghz
Thermal Requirement for 400G-ZR+: 20W

- 400G-ZR
  100km Reach
  15W Power

- 400G-ZR+
  up to 1000km Reach
  20W Power

400G-ZR+ Optics Approaching the Performance of Traditional High-end DWDM Optics
Thermal Requirement 800G Optics: 20W

400G-FR4/LR4 Optics
10-12W Thermal Envelope

Dual 400G/800G Optics
Need 20W Thermal Envelope

No Significant Power Reduction going from 400G to 800G
The Biggest Challenge for Operators

How to Increase Bandwidth for next-gen Applications while simultaneously lowering CAPEX and OPEX
400G Router Price per Port

10X Improvement in Price-Per Port with Merchant Silicon Routers compared to legacy Router Price Points

Legacy Router

Merchant Silicon
400G DWDM Price Per Bandwidth

Order of Magnitude Cost-Reduction with 400G-ZR/ZR+ compared to legacy Optical Transport Price Points
Fatter Pipes are Easier to Manage

Fatter Pipes are more efficient and easier to manage than equivalent bandwidth with smaller pipes.

One 400G Pipe

100G Pipe
100G Pipe
100G Pipe
Fatter Pipes are Lower Cost per Bandwidth

400G is fundamentally lower cost than 4x100G
400G Summary

Large Improvement in Bandwidth Price-Performance

Enables Fundamentally more cost-effective Networks that are also more efficient and easier to manage

Timeline: Field Trials in 2019, Production in 2020