The Key to 400G
Technology, infrastructure, and design strategies for making 400 Gb/s Ethernet a reality.

Gary Bernstein, RCDD
Leviton Network Solutions
Outline

• The growth of cloud computing
• Market estimates for 100G/200G/400G Ethernet
• New 400G active equipment
• Latest 100, 200 and 400G transceivers options
• Trends with multimode vs. single-mode optics
• 400G use cases and cabling strategies
  – Switch to Switch
  – Switch to Server breakout applications
Global Cloud Traffic Growth

• Cloud traffic will grow 3.3-fold from 2016 to 2021
• Cloud accounts for 95% of traffic by 2021 up from 82% in 2015

Source: Cisco Global Cloud Index, 2016-2021
Hyperscale Data Center Growth

• Quantity will increase from **338 to 620** by 2021

• They will represent **53%** of all installed servers by 2021

• In 2018, there were **24 companies** considered hyperscale

Source: Cisco Global Cloud Index, 2016-2021
The Need for Speed
Ethernet Speed Market Forecast

Transceiver modules by speed, percent of total

100G Ethernet Growth QSFP Modules

• Growing much faster than expected

• 100G-CWDM4 is highest volume (SM-2F)

• 100G-SR4 is second highest volume

Source: LightCounting, Sept. 2018
400G Transceiver Forecast

40/100/400 Gb/s Ports Shipments

- 40 Gb/s
- 100 Gb/s
- 400 Gb/s

Port Shipments in Millions

Source: 650 Group
IEEE802.3bm defines 25 Gb/s lanes

IEEE802.3cd defines 50 Gb/s lanes

100G Lambda MSA defines 100 Gb/s lanes

- Uses an SFP28 form factor
- Uses an SFP56 form factor
- Uses a QSFP28 form factor
- IEEE Study Group formed in January 2019 (802.3cu)
## Enterprise vs. Cloud Network Speeds

### Current vs. Future Network Configurations

<table>
<thead>
<tr>
<th></th>
<th>Enterprise Data Centers</th>
<th>Cloud Data Centers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SERVER</td>
<td>UPLINKS</td>
</tr>
<tr>
<td><strong>Current Network</strong></td>
<td>1G</td>
<td>10G</td>
</tr>
<tr>
<td><strong>Speeds</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Future Network</strong></td>
<td>10G</td>
<td>40G</td>
</tr>
<tr>
<td><strong>Speed Options</strong></td>
<td></td>
<td><strong>OR</strong></td>
</tr>
<tr>
<td></td>
<td>10G</td>
<td>25G</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>OR</strong></td>
</tr>
<tr>
<td></td>
<td>100G</td>
<td>400G</td>
</tr>
</tbody>
</table>
MM vs. SM Transceiver Estimated Volumes

40G to 400G Speed Modules

Source: LightCounting, Sept. 2017

SM will account for 68% of volume in 2022
100G Optical Transceivers & Cabling

100GbE Channel Cost Comparison

- Optics
- Cabling

Transceivers + Cabling Price

$0 $1,000 $2,000 $3,000 $4,000 $5,000 $6,000 $7,000 $8,000 $9,000 $10,000

- MM Parallel SR4
- SM Duplex CWDM4
- SM Parallel PSM4
- MM Duplex SWDM4
- MM Duplex SR-BD

PSM4 and SR4 are at same price

* Transceiver Prices are estimated List Prices thru public sources.
* OM4 cabling is used for SR4 & SR-BD, OM5 cabling is used for SWDM4
NEW 400G SWITCHES ON THE MARKET
Terabit-Scale Switch Bandwidth

Key Benefits to Customers:

- 4X bandwidth increase per RU: from 3.2T to 12.8T
- Enables higher density 100G ports using breakouts
- Enables 2-4X lower cost per port
Switch Form Factor Evolution

- SFP+ 10G
- QSFP+ 40G
- QSFP28 100G
- QSFP-DD/OSFP 400G
New 400G Transceiver Form Factors

**QSFP-DD (QSFP Double Density)**
- Can support up to 32 ports in 1U
- Backwards compatible to QSFP+/28
- Accepts LC, MPO (12, 16, 24), CS connectors

**OSFP (Octal Small Form Pluggable)**
- Can support up to 32 ports in 1U
- Adapter to convert QSFP to OSFP for backwards compatibility
- Accepts LC, MPO (12, 16, 24), CS connectors
New Arista 400G Switches 7060 Series

- 1RU TOR with 32 ports
- 100G, 200G or 400G
- Max of 12.8 Terabits
- Low latency of 700ns
- 128 x 100G ports
- 32 x 400G ports
- **Offer both QSFP-DD and OSFP Options**
New Arista 400G Switches 7368X4 Series

• 4RU, 8-slot chassis w/max of 25.6 Terabits
• 32 x 400G ports
• 128 x 100G ports
• 16 ports/100G line card
• 4 ports/400G line card in QSFP-DD & OSFP
• Low latency of 700ns
New Arista 400G Switches 7280R3 Series

- 1RU TOR
- 32 x 100G ports
- 4 x 400G uplink ports
- Max of 4.8 Terabits
- Offer both QSFP-DD and OSFP Options
New Cisco 400G Switches

- 1RU ToR with 32 ports
- 25/50/100/200 or 400Gb/s
- Max of 12.8Terabits
- Low latency of 700ns
- 128 x 100G ports
- 32 x 400G ports
- All 400G ports use QSFP-DD
New Cisco 400G Switches

- 4RU, 8-slot chassis
- 32 x 400G ports
- 128 x 100G ports
- 16 ports/100G line card
- 4 ports/400G line card
- Super low latency of 370ns
- **All 400G ports use QSFP-DD**
New Cisco 400G Switches 9300 Series

9316
• 16 x 400G ports
• 40, 100 or 400G

93600
• 28 x QSFP28 40/100G ports
• 8 x 400G ports

All 400G ports use QSFP-DD
New Juniper 400G Switch QFX5220-32CD

- 1RU ToR with 32 ports
- 25/50/100/200 or 400Gb/s
- Max of 12.8Terabits
- 128 x 100G ports
- 32 x 400G ports
- All 400G ports use QSFP-DD
# 100G Optical Transceivers Q4-2019

<table>
<thead>
<tr>
<th>Transceiver</th>
<th>Switch Vendor</th>
<th>Form Factor</th>
<th>IEEE Compliant</th>
<th>Fiber Type</th>
<th>Distance (meters)</th>
<th># of fibers</th>
<th>Connector</th>
</tr>
</thead>
<tbody>
<tr>
<td>100G-SR10</td>
<td>All</td>
<td>CFP/CFP2/CPAK</td>
<td>Yes</td>
<td>OM3/OM4</td>
<td>100/150</td>
<td>20</td>
<td>24F MTP</td>
</tr>
<tr>
<td>100G-SR10 MXP</td>
<td>Arista</td>
<td>Embed. Optics</td>
<td>No</td>
<td>OM3/OM4</td>
<td>100/150</td>
<td>24</td>
<td>24F MTP</td>
</tr>
<tr>
<td>100G-SR4</td>
<td>All</td>
<td>QSFP28</td>
<td>Yes</td>
<td>OM3/OM4</td>
<td>70/100</td>
<td>8</td>
<td>24F MTP</td>
</tr>
<tr>
<td>100G-XSR4</td>
<td>Arista, Juniper</td>
<td>QSFP28</td>
<td>Yes</td>
<td>OM3/OM4</td>
<td>300</td>
<td>8</td>
<td>12F MTP</td>
</tr>
<tr>
<td>100G-LRL4</td>
<td>Arista</td>
<td>QSFP28</td>
<td>Yes</td>
<td>OS2</td>
<td>2,000</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>100G-CWDM4</td>
<td>All</td>
<td>QSFP28</td>
<td>No</td>
<td>OS2</td>
<td>2,000</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>100G-LR4</td>
<td>All</td>
<td>CFP2/CPAK/ QSFP28</td>
<td>Yes</td>
<td>OS2</td>
<td>10,000</td>
<td>2</td>
<td>LC/SC</td>
</tr>
<tr>
<td>10x10-LR</td>
<td>Cisco</td>
<td>CPAK</td>
<td>No</td>
<td>OS2</td>
<td>1,000</td>
<td>20</td>
<td>24F MTP</td>
</tr>
<tr>
<td>100G-PSM4</td>
<td>All</td>
<td>QSFP28</td>
<td>No</td>
<td>OS2</td>
<td>500</td>
<td>8</td>
<td>12F MTP</td>
</tr>
<tr>
<td>100G-SWDM4</td>
<td>Arista</td>
<td>QSFP28</td>
<td>No</td>
<td>OM3/OM4</td>
<td>70/100</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>100G-SR-BD</td>
<td>Cisco (40/100), Arista</td>
<td>QSFP28</td>
<td>No</td>
<td>OM3/OM4</td>
<td>70/100</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>100G-FR</td>
<td>Cisco, Arista</td>
<td>QSFP28</td>
<td>No</td>
<td>OS2</td>
<td>2,000</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>100G-DR</td>
<td>Arista</td>
<td>QSFP28</td>
<td>No</td>
<td>OS2</td>
<td>500</td>
<td>2</td>
<td>LC</td>
</tr>
</tbody>
</table>
## 100G Optical Transceivers Q4-2019

<table>
<thead>
<tr>
<th>Transceiver</th>
<th>Switch Vendor</th>
<th>Form Factor</th>
<th>IEEE Compliant</th>
<th>Fiber Type</th>
<th>Distance (meters)</th>
<th># of fibers</th>
<th>Connector</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 100G-SR10</td>
<td>All</td>
<td>CFP/CFP2/CPAK</td>
<td>Yes</td>
<td>OM3/OM4</td>
<td>100/150</td>
<td>20</td>
<td>24F MTP</td>
</tr>
<tr>
<td>2 100G-SR10 MXP</td>
<td>Arista</td>
<td>Embed. Optics</td>
<td>No</td>
<td>OM3/OM4</td>
<td>300</td>
<td>24</td>
<td>24F MTP</td>
</tr>
<tr>
<td>3 100G-SR4</td>
<td>All</td>
<td>QSFP28</td>
<td>Yes</td>
<td>OM3/OM4</td>
<td>70/100</td>
<td>8</td>
<td>12F MTP</td>
</tr>
<tr>
<td>4 100G-XSR4</td>
<td>Arista, Juniper</td>
<td>QSFP28</td>
<td>Yes</td>
<td>OS2</td>
<td>2,000</td>
<td>8</td>
<td>LC</td>
</tr>
<tr>
<td>5 100G-LRL4</td>
<td>Arista</td>
<td>QSFP28</td>
<td>No</td>
<td>OS2</td>
<td>2,000</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>6 100G-CWDM4</td>
<td>All</td>
<td>QSFP28</td>
<td>No</td>
<td>OS2</td>
<td>10,000</td>
<td>2</td>
<td>LC/SC</td>
</tr>
<tr>
<td>7 100G-LR4</td>
<td>All</td>
<td></td>
<td>No</td>
<td>OS2</td>
<td>1,000</td>
<td>20</td>
<td>24F MTP</td>
</tr>
<tr>
<td>8 10x10-LR</td>
<td>Cisco</td>
<td></td>
<td>No</td>
<td>OS2</td>
<td>500</td>
<td>8</td>
<td>12F MTP</td>
</tr>
<tr>
<td>9 100G-PSM4</td>
<td>All</td>
<td>QSFP28</td>
<td>No</td>
<td>OS2</td>
<td>70/100</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>10 100G-SWDM4</td>
<td>Arista</td>
<td>QSFP28</td>
<td>No</td>
<td>OM3/OM4</td>
<td>70/100</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>11 100G-SR4-3</td>
<td>Arista, Juniper</td>
<td>QSFP28</td>
<td>No</td>
<td>OM3/OM4</td>
<td>2,000</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>12 100G-FR</td>
<td>Arista</td>
<td>QSFP28</td>
<td>No</td>
<td>OS2</td>
<td>2,000</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>13 100G-DR</td>
<td>Arista</td>
<td>QSFP28</td>
<td>No</td>
<td>OS2</td>
<td>500</td>
<td>2</td>
<td>LC</td>
</tr>
</tbody>
</table>

*Many 100G Options Available*
New Ethernet Terminologies for 200/400G

**SR8** 100m reach
- 8 optical channels (16 total MM fibers-MPO). Each channel operates at 50Gb/s.

**DR4** 500m reach
- 4 optical channels (8 total SM fibers-MPO). Each channel operates at 100Gb/s.

**FR4** 2km reach
- 4 optical channels multiplexed onto 1 pair (2 total SM fibers-LC). Each channel operates at 100Gb/s.
# 200G Optical Transceivers Q4-2019

<table>
<thead>
<tr>
<th>200G Transceiver</th>
<th>STD</th>
<th>MFR</th>
<th>Form Factor</th>
<th>Breakout Option</th>
<th>Fiber Type</th>
<th>Distance (meters)</th>
<th># of fibers</th>
<th>Connector</th>
</tr>
</thead>
<tbody>
<tr>
<td>200G-FR4</td>
<td>IEEE</td>
<td>None</td>
<td>?</td>
<td>No</td>
<td>OS2</td>
<td>2,000</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>200G-DR4</td>
<td>IEEE</td>
<td>None</td>
<td>?</td>
<td>Yes</td>
<td>OS2</td>
<td>500</td>
<td>8</td>
<td>12F MTP</td>
</tr>
<tr>
<td>2X100G-PSM4</td>
<td>Prop.</td>
<td>Cisco</td>
<td>QSFP-DD</td>
<td>Yes</td>
<td>OS2</td>
<td>500</td>
<td>24</td>
<td>24F MTP</td>
</tr>
<tr>
<td>200G-SR4</td>
<td>IEEE</td>
<td>None</td>
<td>?</td>
<td>Yes</td>
<td>OM3/OM4/OM5</td>
<td>70/100/100</td>
<td>8</td>
<td>12F MTP</td>
</tr>
<tr>
<td>2X100G-SR4</td>
<td>Prop.</td>
<td>Cisco</td>
<td>QSFP-DD</td>
<td>Yes</td>
<td>OM3/OM4/OM5</td>
<td>70/100/100</td>
<td>24</td>
<td>24F MTP</td>
</tr>
<tr>
<td>2x100G-CWDM4</td>
<td>Prop.</td>
<td>Cisco</td>
<td>QSFP-DD</td>
<td>Yes</td>
<td>OS2</td>
<td>2,000</td>
<td>4</td>
<td>CS</td>
</tr>
</tbody>
</table>

= Expected High Volume Options
# 400G Optical Transceivers Q4-2019

<table>
<thead>
<tr>
<th>400G Transceiver</th>
<th>STD</th>
<th>MFR</th>
<th>Form Factor</th>
<th>Breakout Option</th>
<th>Fiber Type</th>
<th>Distance (meters)</th>
<th># of fibers</th>
<th>Connector</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 400G-FR4*</td>
<td>IEEE/MSA</td>
<td>Arista, Cisco, Juniper</td>
<td>QSFP-DD, OSFP</td>
<td>No</td>
<td>OS2</td>
<td>2,000</td>
<td>2</td>
<td>LC</td>
</tr>
<tr>
<td>2 400G-DR4</td>
<td>IEEE</td>
<td>Arista, Cisco, Juniper</td>
<td>QSFP-DD, OSFP</td>
<td>Yes</td>
<td>OS2</td>
<td>500</td>
<td>8</td>
<td>12F MTP</td>
</tr>
<tr>
<td>3 400G-XDR4 (DR4+)</td>
<td>Prop.</td>
<td>Arista, Juniper</td>
<td>QSFP-DD, OSFP</td>
<td>Yes</td>
<td>OS2</td>
<td>2,000</td>
<td>8</td>
<td>12F MTP</td>
</tr>
<tr>
<td>4 400G-SR8*</td>
<td>IEEE</td>
<td>Arista</td>
<td>OSFP</td>
<td>Yes</td>
<td>OM3/OM4/OM5</td>
<td>70/100/100</td>
<td>16</td>
<td>16F/24F MTP</td>
</tr>
<tr>
<td>5 400G-SR4.2(BD)*</td>
<td>IEEE/MSA</td>
<td>Cisco</td>
<td>QSFP-DD</td>
<td>Yes</td>
<td>OM3/OM4/OM5</td>
<td>70/100/150</td>
<td>8</td>
<td>12F MTP</td>
</tr>
<tr>
<td>6 400G-2FR4</td>
<td>Prop.</td>
<td>Arista</td>
<td>OSFP</td>
<td>Yes</td>
<td>OS2</td>
<td>2,000</td>
<td>4</td>
<td>CS</td>
</tr>
</tbody>
</table>

* For SR4.2, IEEE802.3cm estimated to be published in Q1-2020. For FR4, IEEE802.3cu estimated to be published in Q4-2020.

Note: Does not include long reach 10Km options. “SR16” or “FR8” options not available on market.
400G Multi-Source Agreements (MSAs)

- **400G bidi**: 8x50G, MM parallel
- **100G Lambda**: Created 400G-FR4, 4x100G, SM duplex
- **CWDM8**: 8x50G, SM duplex

Logos of Alibaba, Arista, Cisco, Finisar, Intel, and Microsoft.
200G & 400G USE CASES
200/400G Cabling Designs Factors

- **Enterprise vs. cloud provider company**
- **Current server speeds**: 1G, 10G or 25G
- **Reach requirements**
- **Power levels**: 10-12W (3x100G)
- **Channel insertion loss budgets**
- **Cost per channel**
SWITCH-TO-SWITCH USE CASES
400G: Switch-to-Switch Use Case

2-fiber Single-Mode Configuration (10/40/100/200/400G) with 400G FR4/FR8

- Base24 Fiber Backbone Cabling
- Provides Duplex (2-fiber) connections at equipment
- Supports 400G-FR4 and 400G-FR8
400G: Switch-to-Switch Use Case
8-fiber MM Configuration (10/40/100/200/400G) with SR4.2 (BiDi)

OM4

- Base24 Fiber Backbone Cabling
- Provides Parallel (8-fiber) connections at equipment
- Supports 400G-SR4.2 over OM4
400G: Switch-to-Switch Use Case

8-fiber SM Configuration (10/40/100/200/400G) with DR4

- Base24 Fiber Backbone Cabling
- Provides Parallel (8-fiber) connections at equipment
- Supports 400G-DR4 over OS2
200 or 400G: Switch-to-Switch Use Case

16/24-fiber SM or MM Channel Configuration (200/400G) with SR8 or 2x100

- Base24 Fiber Backbone Cabling
- Provides Parallel (16/24-fiber) connections at equipment
- Supports 400G-SR8 or 2x100-SR4 over OM4
- Supports 2x100-PSM4 over OS2
BREAKOUT USE CASES
400G to 4x100G Breakout Use Case
8-fiber Single-Mode Channel B/O configuration with 400G-DR4

• Convert 8-fiber channel into 4 x duplex channels
• No need to replace Base24 backbone cabling
• Creates 4 x 100G channels/module or 128 x 100G ports/1U Switch
Patching Solution – ToR w/100G Breakout

128 x 100G Ports, 2RU of Rack Space

32-Port TOR Switch

8-F MTP Array Cords

MTP-LC Conversion Cassette

LC Patch Cord

QSFP28 100G-DR
200G to 2x100G Breakout Use Case

16-fiber OM4 Configuration B/O to 8-fiber channels with 400G-SR8

- OSFP must be configured in 200G mode
- Convert 16-fiber channel into 2 x 8-fiber channels
- No need to replace Base24 backbone cabling
- Creates **2 x 100G** channels/OSFP or **64 x 100G** ports/1U Switch
- Other potential breakout options will become available with conversion cassettes
200G to 8x25G Breakout Use Case

24-fiber Single-Mode Channel B/O configuration with 2x100-PSM4

- Convert 24-fiber channel into 8 x duplex channels
- No need to replace Base24 backbone cabling
- Creates 8 x 25G channels/module or 256 x 25G ports/1U switch
200G to 8x25G Breakout Use Case

24-fiber Single-Mode Channel B/O configuration with 2x100-SR4

- Convert 24-fiber channel into 8 x duplex channels
- No need to replace Base24 backbone cabling
- Creates 8 x 25G channels/module or 256 x 25G ports/1U switch
200G to 2x100G Breakout Use Case

4-fiber Single-Mode Channel B/O configuration with 2x100G-CWDM4/400G-2FR4

- OSFP 400G must be configured in 200G mode
- Uses new “CS” connector interface – must convert to LC
- Base24 Fiber Backbone Cabling
- Creates 2 x 100G channels/OSFP or 64 x 100G ports/1U Switch
The Key to 400G – Summary

- **400G switches** have recently been introduced to the market
- There are many transceiver choices for 200 and 400G applications
- Customers will be interested in various breakout scenarios
- Networks designs must have flexibility to allow migration to 200/400G
- Majority of the 200/400G options require single-mode cabling
- 100/200/400G cabling designs are very complicated
The Key to 400G
Technology, infrastructure, and design strategies for making 400 Gb/s Ethernet a reality.

Gary Bernstein, RCDD
Leviton Network Solutions