Next-generation Interconnects: The Critical Importance of Connectors and Cables

Live Webcast
October 19, 2021
10:00 am PT / 1:00 pm ET
Today’s Presenters

Tim Lustig
Moderator
Director, Corporate Ethernet Marketing
NVIDIA

Kent Lusted
Principal Engineer
Intel

Sam Kocsis
Director of Standards and Technology
Amphenol

Brad Smith
Director of Marketing, Networking
NVIDIA
SNIA-at-a-Glance

180
industry leading organizations

2,500
active contributing members

50,000
IT end users & storage pros worldwide

Learn more: snia.org/technical  🚀 @SNIA
Technologies We Cover

- Ethernet, Fibre Channel, InfiniBand®
- iSCSI, NVMe-oF™, NFS, SMB
- Virtualized, HCI, Software-defined Storage
- Storage Protocols (block, file, object)
- Securing Data
The material contained in this presentation is copyrighted by the SNIA unless otherwise noted.

Member companies and individual members may use this material in presentations and literature under the following conditions:

- Any slide or slides used must be reproduced in their entirety without modification.
- The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.

This presentation is a project of the SNIA.

Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.

The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.
Agenda

- Common pluggable connectors and media types
- Copper and optical media
- Development of 25/50GbE
- Why cable types matter
- Copper cabling and transceivers
- Signal modulations
- Optical connectors
- Why the variety
- Real world use cases
Common Pluggable Connectors and Media Types

Kent Lusted
Introduction

- Pluggable connectors come in a variety of sizes and lane widths
  - One-lane and four-lane are the most common. Eight-lane and two-lane are being deployed now. Sixteen-lane are on the horizon
- Both copper and optical media are commonly used in networks
  - Copper for cost optimized, short reach
  - Optical for medium to longer reach
- 100 Gb and 25 Gb Ethernet
  - Three types of 25 Gb Ethernet exist
Pluggable Networking Connector Form Factor Lineage

- 16 lanes
- 8 lanes
- 4 lanes
- 2 lanes
- 1 lane

# lanes

- SFP
- DSFP
- microQSFP
- QSFP
- OSFP
- QSFP-DD
- SFP-DD
- OSFP-XD

Expand
- to 16 lanes
- to 8 lanes
- to 4 lanes
- Fill in 2 lanes
- Expand to 4 lanes

2001
2006
time
today

2001
2006
time
today

<table>
<thead>
<tr>
<th># lanes</th>
<th>1 lane</th>
<th>2 lanes</th>
<th>4 lanes</th>
<th>8 lanes</th>
<th>16 lanes</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>2001</td>
<td>2006</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Pluggable Example - SFP

- **SFP** is a 1-lane connector.
- Used for copper cables or optical or BASE-T transceivers.
- **SFP** is a 1G connector. **SFP+** is a 10G/lane rated connector. **SFP28** is a 28Gbps rated SFP connector.
Quad Small Form factor Pluggable (QSFP) is a 4-lane connector type.
QSFP is used for Ethernet, InfiniBand, Fibre Channel, SAS, Omni Path, and various proprietary protocols
It is not a “CR4 connector” nor is it a “40G or 100G connector”
Twin-axial “Direct Attach Copper (DAC)” Cables

- **Twin-axial Cable**
  - Point-to-point

- **Twin-axial Breakout Cable**
  - Point to multi-point

### Passive DAC (Direct Attach Copper) cable
Passive DAC (Direct Attach Copper) cable is a high-speed twinaxial conductor electrical cable with connectors at each end, but no active components in the assembly.

### ACC/AEC (Active Copper or Active Electric) cables
ACC/AEC (Active Copper or Active Electric) cables have active devices such as redrivers or retimers inside the connector.
Pluggable Optical Transceiver Modules

**Pluggable Transceiver** is a module with an electrical connector in one end (to host) and optical connector in the other end (to the medium/fiber). It can have one or more parallel lanes in each direction (transmit and receive).

**Active Optical Cable** (AOC) has a non-removable optical cable fixed to the module.

- **Separable interface to fiber medium**
- **Electrical interface to host**
- **Optical fiber cable**
- **Captive interface to fiber medium**
Optical Fiber Types

SINGLE MODE AND MULTIMODE OPTICAL FIBERS

Optical Fibers
8μm to 60μm core; 125μm cladding

Yellow Color Fiber

Aqua Color Fiber (OM3)
Development of 100Gb and 25Gb Ethernet

- Around 2012, Industry recognized the need to move to 25G-class SERDES.
- Four lane media was already common due to 40G-CR4 and 40G-KR4. Therefore, the path to 100G = 4 lanes of 25G SERDES.
- As 100GbE was deploying, there was strong interest to use 1-2 lanes of 25G for server connections.
  - Then came the IEEE 802.3 versions of 25GbE: 25GBASE-CR-S, 25G-BASE-CR.
- There are three versions of 25GbE; they are different from each other.
A Result: Three 25G Cable PHY Types

- **IEEE 802.3 25GBASE-CR**
  - Supports three FEC modes (RS(528)-FEC, BASE-R FEC, no FEC)
  - Cables up to 5m (when using the RS-FEC)
- **IEEE 802.3 25GBASE-CR-S**
  - Subset of 25GBASE-CR.
  - Does not support RS-FEC. Only BASE-R FEC and no-FEC
    - Compatible with Consortium’s 25GBASE-CR1 when no-FEC mode is chosen
    - Up to 3m cable
- **Ethernet Technology Consortium 25GBASE-CR1**
  - Compatible with IEEE’s 25GBASE-CR-S
  - Up to 3m cable
Three 25G CR PHY to Cable Mapping

<table>
<thead>
<tr>
<th>Cable Type</th>
<th>Reach</th>
<th>FEC Modes</th>
<th>Supported PHY types</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA-25G-L</td>
<td>5m</td>
<td>RS(528)</td>
<td>YES</td>
</tr>
<tr>
<td>CA-25G-S</td>
<td>3m</td>
<td>RS(528), BASE-R</td>
<td>YES</td>
</tr>
<tr>
<td>CA-25G-N</td>
<td>3m</td>
<td>RS(528), BASE-R, No FEC</td>
<td>YES</td>
</tr>
</tbody>
</table>
Why Does Cable Type Matter?

- Cable type affects Auto-Negotiation advertisement and FEC choices
- Cable type may be interpreted differently by each vendor and could result in link up issues.
  - Know what cable type you have; visual identification is insufficient
Summary Points

- Pluggable connectors support copper and optical media
- Copper media can be passive or active
- Optical media uses different types of fiber
  - Transceiver may be separable or captive of the fiber
- 25 Gb Ethernet and copper cables have a special relationship
  - Three types of Ethernet PHYs, Three types of cables!
Copper Cabling and Transceivers 101

Brad Smith
4 Types of Cables and Transceivers Products

10G/25G/100G/200G/400G

Direct Attach Copper (DAC)
Copper wires “Directly Attaches” system together
Lowest Priced, Zero-Latency, Zero-Power
Up to 2,3 and 5m Reaches

Active Optical Cables (AOC)
2 Transceivers w/optical fibers bonded inside.
Lowest-Priced Optical Link
Up to 100m Reaches

Optical Transceivers
Electrical signals pulse laser light sent into optical fibers
Optical Connectors & Long Reaches
Up to 100m/2km/10km/40km ... Reaches

Passive Copper Cables

Active Copper Cables (ACC)
DAC Copper with a Signal Booster IC
Priced between DACs & AOCs, Latency + Power
Up to 2,3 and 7m Reaches

Transceivers with Integrated Fibers

Transceivers with Detachable Optical Fiber Connectors

Pre-emphasis
Or DSP IC

AKA: “Active DAC”
“Active Electric Cables (AEC)”
25G/100 Gb Ethernet Cables & Transceivers

1 & 4-channels, 25G-NRZ, in SFP28/QSFP28

Direct Attach Copper (DAC)

- **Up to 5 meters**
  - 25GbE SFP28
  - 100GbE QSFP28
  - 100GbE-to-2x50GbE QSFP28-QSFP28
  - 100GbE-to-4x25GbE QSFP28-SFP28

Active Optical Cables (AOC)

- **Up to 100 meters**

Optical Transceivers

- **Up to 10k meters**

- **Multimode**
  - 25G SR
    - Up to 100 meters
  - 100G SR4
    - Up to 100 meters

- **Single mode**
  - 25G LR
    - Up to 10k meters
  - 100G PSM4
    - Up to 500 meters
  - 100G CWDM4
    - Up to 2k meters
  - 100G LR4
    - Up to 10k meters

Transceivers with Detachable Connectors

- **Up to 5 meters**
  - 25GbE SFP28 (Multimode)
  - 100GbE QSFP28 (Multi/Single mode)
  - 100GbE-to-2x50GbE QSFP28-QSFP28 (Multi/Single mode)
  - 100GbE-to-4x25GbE QSFP28-SFP28 (Multi/Single mode)
Inside DAC Cables

Simplest high-speed interconnect

Wires, shielding, PCB, soldier, shell, label

Active DAC or Active Copper Cable includes a signal boosting IC in the ends to extend the reach

2x400G DAC has 48 wires
3 wires/direction (2+ground)
8-channels x 2 directions x 3 wires = 48 wires
Typical DAC-in-the-Rack Use

Top of Rack Switches

- (16) 100G QSFP28 ports
- (18) 25G SFP28 & (4) 100G QSFP28 ports
- (48) 25G SFP28 & (8) 100G QSFP28 ports
- (32) 100G QSFP28 ports
- (64) 100G QSFP28 ports

DACs = Wires, shielding, PCB, soldier, shell, label

Advantages:
- Zero-latency
- Zero-power consumption
- Lowest-cost
- Simplicity
- High-reliability
- Long reach

Disadvantages:
- Heavy
- Bulky
- Shorter with faster line rates

25G-NRZ = up to 5-meters

Passive Copper DACs

- 25G-NRZ up to 5-meters
- 50G-PAM4 up to 3-meters
- 100G-PAM4 up to 2-meters

25G-NRZ = up to 5-meters
50G-PAM4 = up to 3-meters
100G-PAM4 = up to 2-meters
DACs and Active Copper Cables

**Passive Copper DACs**
- 25G-NRZ up to 5-meters
- 50G-PAM4 up to 3-meters
- 100G-PAM4 up to 2-meters

**Active Copper DACs**
- 25G-NRZ up to 7-meters
- 50G-PAM4 up to 5-meters
- 100G-PAM4 up to 3-meters

+ Slightly longer reaches
+ Thinner cables
- Consumer power
- Induces latency

Active DACs = Wires, shielding, PCB, soldier, shell, label + ICs
As Line Rate Increases
Cable Thickness Increases + Lengths Become Shorter

Shorter due to EMI and difficulty detecting data signal
Thicker cable due to additional cable shielding needed

1G/10G DACs
- 7-10 meters
- 10G/40GbE + QDR/FDR

25G/100G DACs
- 5 meters
- 25G/100GbE + EDR

50G/200G DACs
- 2-3 meters
- 200GbE + HDR

100G/400G DACs
- 2-3 meters
- 400GbE

2x400G/800G
- 1-2 meter
- 2x400G NDR

Faster line rates turn wires into radio antenna’s
Wires become very short

Transition to ACCs and Transceivers
Signal Modulation: NRZ and PAM4

NRZ and PAM4 are different forms of amplitude (intensity) modulation of data signals

- **NRZ** (Non-return to Zero)
- **PAM4** (Pulse Amplitude Modulation 4-levels)

**25G-NRZ**
- Traditional one data bit (1,0)
- 1-bit x 25GHz clock = 25Gb/s
- (Non-return to Zero)

**50G-PAM4**
- Stacks 2 data bits on top of each other, same 25GHz clock; 2-bits x 25GHz = 50Gb/s
- (Pulse Amplitude Modulation 4-levels)

**100G-PAM4**
- Twice faster clock than 50G-PAM4, 2-bits x 50GHz = 100Gb/s
- (Non-return to Zero)

**EDR** has 1-bit (1,0)
**HDR** stacks 2-bits on top of each other
**NDR** runs twice as fast as HDR

Applies to both Electrical and Optical Modulation
Form-factor Buzzword-ology

Why all the new form-factors?

- **SFP** = Small Form-factor Plug (1-channel)
- **Quad SFP** (4-channels)
- **QSFP-DD** = Double Density (8-channels)
- **OSFP** = Octal SFP (8-channels)

Max line rates

- SFP+ = 1x10G
- QSFP+ = 4x10G
- SFP28 = 1x25G
- QSFP28 = 4x25G
- SFP56 = 1x50G
- QSFP56 = 4x50G
- SFP112 = 1x100G
- QSFP112 = 4x100G
- SFP-DD = 2x50G, 2x100G
- QSFP56-DD = 8x50G,
- QSFP112-DD = 8x100G
- OSFP = 8x50G, 8x100G
- OSFP-XD = 16x100G, 16x200G

MSA: Multiple Source Agreement
Industry standard. enclosures for transceiver
Holds wires and transceiver components.

Zinc/Copper Shells
Can be used with
DAC, AOCS, MM or SM transceivers.
Speed rated (EFI limit)
Data center: QSFP + OSFP Transceivers and Form-factors

QSFP28
QSFP56
QSFP112
QSFP-DD
OSFP
OSFP-XD

EDR / 100GbE
4x25G NRZ
4 channels

HDR / 200GbE
4x50G PAM4

NDR / 400GbE
4x100G PAM4

400GbE/800GbE
8x50G-PAM4
8x100G PAM4

Twin-NDR 2x400G
2x400GbE
8x100G PAM4

1.6Tb + 3.2Tb
16x100G-PAM4
16x200G-PAM4

QSFP

Optical connectors

Multiple-Push-On
Ultra Polished Connector (flat)
MPO/UPC

Multiple-Push-On
Angled Polished Connector
MPO/APC

Lucent Connector
LC

1-Channel

2-Channel

Single mode = yellow
Multimode = Aqua + Tan

NEW
OSFP, the "Monster-truck" of Transceivers

**OSFP & QSFP112-DD enables:**
- 32-ports of 800Gb/Transceiver = 51T switch
- 64-ports of 800Gb/Transceiver= 102T switch

**OSFP-XD enables:**
- 32-ports of 1.6Tb/Transceiver = 102T switch
- 32-ports of 3.2Tb/Transceiver= 204T switch
- 64-ports of 3.2Tb/Transceiver= 408T switch

*Summary: Pluggable optics will be around a LONG time*

1G SFP+ ...to ... 3.2T QSFP-XD
Multi-mode & Single-mode Fibers

**Multi-mode**

- Large light carrying core
- Easy and low-cost to interface with lasers and detectors
- Signal bounces down fiber and degrades; limiting lengths to ~100-meters
  - Used with only one signal at 850nm wavelength
  - Used with SR, SR4, SR8 transceivers

**Single-mode**

- Tiny light carrying core
- Difficult and higher-cost to interface with lasers and detectors
- Signal stays together over hundreds of kilometers - used in to 40km InfiniBand
  - Can multiplex 4 or 8 signals simultaneously into 1 fiber at 1310nm wavelength
  - Used with DR1, DR4, FR4, LR4 transceivers

---

**Aqua jacket signifies Multi-mode OM4**

**Short Reach**
- SR
- SR4
- SR8
- 100-meters max (OM4)

**Long Reach**
- PSM4 / DR4
- CWDM4 / FR4
- LR4 / ER
- 100-Kilo-meters max

---

**Yellow jacket signifies Single-mode SMF28 OS1**

**Multiple paths = Multi-mode**

Pulse light rays scatter & arrive at different times, stretching out signal & reducing intensity. Multiple pulses overlap. Limits max length to ~100-meters

---

Pulse light rays stay concentrated, intense & arrive at same time. Multiple colored wavelength signals can be transmitted simultaneously. Max lengths up to 48km and 4-8 channels.
Optical Connectors

25G-NRZ, 50G-PAM4

- MPO/UPC - 12-fiber connector
- Aqua fiber = Multi-mode
- -55dB back reflection
- 50-um core
- Reflected laser energy is low - Flat polish OK
- Used with:
  - 100GbE/EDR SR4 transceivers
  - 200GbE/HDR SR4 transceivers

400G SR8 uses MPO/APC-16 with 16 multi-mode fibers,
MPO/UPC-12, MPO/APC-16 and MPO/APC are not interoperable with each other

100G-PAM4

- MPO/APC - 12-fiber connector
- Green shell = APC
- -65dB back reflection
- 7-um core
- Concentrated laser light reflected off axis to protect laser source
- Used with:
  - 400G/800GbE SR8, DR4 transceivers
  - NDR transceivers both single-mode & multi-mode

Multiple-Push-On
Ultra Polished Connector (flat)
MPO/UPC

Multiple-Push-On
Angled Polished Connector
MPO/APC
> 500 Different Cables & Transceivers Products

10G/25G/100G/200G/400G Ethernet, SAS, InfiniBand, Fibre Channel

**Direct Attach Copper (DAC)**

Copper wires "Directly Attaches" system together
Lowest Priced, Zero-Latency, Zero-Power
Up to 2, 3 and 5m Reaches

- 25G, 50G
- 100G
- 200G
- 400G

**Active Optical Cables (AOC)**

2 Transceivers w/optical fibers bonded inside.
Lowest-Priced Optical Link
Up to 30m and 150m Reaches

- 25G
- 50G
- 100G
- 200G
- 400G

**Optical Transceivers**

Electrical signals pulse laser light sent into optical fibers
Optical Connectors & Long Reaches
Up to 100m/2km/10km/40km Reaches

- SR1
- SR2
- SR4
- SR8
- DR1
- DR4
- FR4
- LR4
- ER

Copper Cables
Active Copper Cables

Transceivers with
Integrated Fibers

Transceivers with Detachable
Optical Connectors

©2021 Storage Networking Industry Association. All Rights Reserved.
Interconnects are **Cost Optimized for Each Reach**

- **DAC Copper Cables (DACs)**
- **Active Copper Cables (ACCs)**
- **Active Optical Cables (AOCs)**
- **Multi-Mode Transceivers**
- **Single-Mode Transceivers**

**100G–to-400G Cables and Transceivers**
Price increases as reach extends:

- **Parallel Optical**
  - 100m
  - 500m
  - 2km
  - 10km

- **Serial Optical**
  - 100G LR4
  - 400G LR4

**DAC-In-the Rack**

**ACCs-extend reach of copper DACs**

**AOCs –Across-the-Top**

**Multi-mode-for-Structured Cabling**

**Single-mode-for-Across the Data center**

**Parallel Optical**
- 100m
- 500m
- 2km Serial Optical 10km

**Multi-Mode Transceivers**
- **Single-Mode Transceivers**
- **DAC Copper Cables (DACs)**
- **Active Copper Cables (ACCs)**
- **Active Optical Cables (AOCs)**
- **Multi-Mode Transceivers**
- **Single-Mode Transceivers**
Optical Descriptors and Codes

Multi-mode transceivers up to 100-meter reach
- SR: Short-Reach 1-channel  
- SR4: Short-Reach 4-channel  
- SR8: Short-Reach 8-channel  

NDR current 30-meters

Single-mode transceivers 500m, 2km, 10km reach
- DR1: Datacenter-Reach 1-channel (500m)  
- DR4: Datacenter-Reach 4-channel (500m)  
- DR8: Datacenter-Reach 8-channel (500m)

New PSM4

- FR4: Far-Reach 4-channel (2km)(multiplexed)  
- LR4: Long-Reach 4-channel (10km)(multiplexed)

New CWDM4

Optical connectors:
- LC: Lucent Connector 2 fibers
- MPO-12: Multiple Push On -12 fiber (on 8 used)
- MPO-16: Multiple Push On -16 fiber (for SR8)
- /UPC: Ultra Physical Contact flat fiber polish (EDR, HDR)
- /APC: Angled Polished Connector (NDR)

All can be used with multi-mode or single-mode fibers

How it all goes together

Optical fibers:
- OS1: Optical Single-mode -1 (inside)
- OM4: Optical Multi-mode -type 4  
  (100m at 25G-NRZ)
  (100m at 50G-PAM4)
  (30-50m at 100G-PAM4)

OM3: Optical Multi-mode -type 3 (old)  
(100m at 10G-NRZ)
(70m at 25G-NRZ)
(70m at 50G-PAM4)
Multimode Optics: Transceivers or AOCs

Top of Rack Switches
(16) 100G QSFP28 ports
(18) 25G SFP28 + (4) 100G QSFP28 ports
(48) 25G SFP28 & (8) 100G QSFP28 ports
(32) 100G QSFP28 ports
(64) 10G QSFPP28 ports

Across Rows & Data Centers
In Structured Cabling Pipes
Switch to Switch
Up to 100m

To Remote & Switches Routers
Up to 100m

Leaf-Spine-Super Spine Switches

Linking Within System Racks

Short Reach Transceivers
25G SR 100m
100G SR4 100m
100G SWDM4 100m
200G SR4 100m
400G S8 100m
2x400G S8 50m

Up to 100m
Single mode Optics: Transceivers

Top of Rack Switches

(16) 100G QSFP28 ports
(18) 25G SFP28 + (6) 100G QSFP28 ports
(48) 25G SFP28 & (8) 100G QSFP28 ports
(32) 100G QSFP28 ports
(64) 100G QSFP28 ports

Linking within System Racks

100G QSFP28
25G LR SFP28
50G QSFP28
4x25G
10G QSFP28
50G QSFP28

CPU Server Arrays

Transceivers
- 25G LR 10km
- 100G PSM4 500m
- 100G CWDM4 2km
- 100G LR4 10km
- 200G SR4 100m
- 200G FR4 2km
- 400G SR8 100m
- 400G DR4 500m
- 400G FR4 2km
- 400G LR4 10km

Across Rows & Data Centers
In Structured Cabling Pipes
Up to 10km

To Remote & Switches Routers
Up to 10km

Up to 10km

Linking Between Sub-Systems

System Rack

Leaf-Spine-Super Spine Switches

Flash Arrays
Cables & Transceivers for 100g/400GbE Networks

FOR SHORT REACH, LONG REACH, STRUCTURED & BREAKOUT CABLEING

- **DAC**
  - ToR-to-GPU/CPU Between Switches
  - Within Racks
  - Up to 2.5m

- **AOC**
  - Between Switches
  - Within Racks
  - 3-30m...100m
  - Splitter Cables & Transceivers
  - For Inside & Between Racks
  - AOC: 3-100m

- **SR/SR4/SR8**
  - For structured cabling
  - Between Racks
  - Up to 100m
  - 8, 16- fiber Parallel

- **DR4**
  - Between Racks & Breakouts
  - Up to 500m

- **FR4**
  - Leaf-Spine Links
  - Spine-Super-Spine
  - Up to 2km

- **LR/LR4**
  - Linking to Routers
  - Up to 10km

- **SR/SR4/SR8**
  - Single-mode For Structured Cabling
  - Long Reaches
  - Leaf-Spine Links

- **400G DR4 to FR1/DR1 with Splitters**
  - Fiber Splitters

- **AOC: 3-100m**
  - Single-mode Optics For Structured Cabling
  - Up to 10Km

- **Multi-Mode Optics**
  - Up to 10Km

- **Top-of-Rack Switches**
  - CPU Server Arrays
  - SSD Arrays
  - CPU Server Arrays
  - HDD Arrays
  - NVMe Arrays
  - SSD Arrays
  - HDD Arrays
  - NVMe Arrays

- **In-the-Rack**
  - Storage Industry Association

- **NSF STORAGE**

- **1X**
  - 2x
  - 4X
  - 8X

- **In-the-Rack**
  - Multi-Mode Optics
  - Up to 10Km

- **Single-Mode Optics**
  - Up to 10Km

- **3m - 100m**

- **1X**
  - 2x
  - 4X
  - 8X

- **Quad 25G SFP Breakout**

- **Splitter Cables & Transceivers**
  - For Inside & Between Racks
  - 1X
  - 2x
  - 4X
  - 8X

- **Spine/Super-Spine**
  - Switch Rack
  - Leaf-Spine Links

- **Top-of-Rack Switches**
  - CPU Server Arrays
  - SSD Arrays
  - CPU Server Arrays
  - HDD Arrays
  - NVMe Arrays
  - SSD Arrays
  - HDD Arrays
  - NVMe Arrays

- **HR.AI**

- **3m - 100m**

- **1X**
  - 2x
  - 4X
  - 8X

- **Quad 25G SFP Breakout**
Typical Leaf-Spine Network

Any node can reach any other node with ≤ 3 switch hops

Core Router link
FR4/LR4 Transceivers
2km-10km

Spine-Super Spine Network
Layer 3
DAC, ACC or AOC cables
1m-20m

GPU/Server/DPU
Storage Network
Layer 1
Multi-mode Transceivers:
SR4/SR8 & AOCs

Leaf Network
Layer 2
Single-mode Transceivers:
CWDM4/FR4
PSM4/DR4
500m-2km

Super Highway

1X
2x
4X
8X

Compute/Storage
DACs, ACCs

Typical Leaf-Spine Network Super Highway

Any node can reach any other node with ≤ 3 switch hops
Where is Most of the Ethernet Data Traffic?

Ans: Between CPUs – Switches – Storage

70%-80% data traffic is inside the rack
Between CPUs – Switches – Storage

25%-15% of data Traffic is in the Leaf-Spine Switch
Middle/End-of-Row Switch network

<5% of data traffic reaches Core Router

The most sensitive area for reducing latency delays and costs

70%-80% data traffic is inside the rack
Between CPUs – Switches – Storage
Typical Data Center Cluster

DACs-in-the-Rack

100GbE

100GbE-to-2x50GbE

100GbE-to-4x25GbE

DAC
Up & down the Rack.
Linking Servers & storage to Switches.
Hybrid breakout cables

Server/GPU Storage Rack

NIC
CPU, GPU FLASH Storage

Spine Rack

100GbE

200

100GbE-to-2x50GbE

100GbE-to-4x25GbE
Where AOCs Are Used?

- Used where access is easy (cable trays)
- Installation is an issue with AOC transceiver ends which are easily damaged.
- Hard to install under floors or in structured cabling pipes

HPCs may use AOCs everywhere including up & down the rack
Where Connectorized Transceivers Used

Multi-mode connectorized transceivers to optical patch panels and in Structured Cabling Pipes & under raised floors where connectors are needed.
Single-mode transceivers spanning across data centers, campuses and to telco links

*NEEDS OPTICAL CONNECTORS*

- 25G SR 100m
- 100G SR4 100m

- 500m DR4 (PSM4)
- 2km FR4 (CWDM4)
  - Connecting clusters together
  - Leaf-spine network over long reaches up to 500m
  - Cross campus networks at 2km FR4
  - To routers up to 10km using LR4
Fujitsu Riken FUGAKU HPC - #1 in Top500

>100 thousand AOCs
Hyper Scale Data Centers

~750,000 sq. feet QTC Data center.

Largest Data center being built by China Mobile

11.3M sq. feet!

~15 of these!

- All about scale-out and low-power, low-cost
- Extensive use of single-mode transceivers for 4.7dB link budgets (2km)
  - PSM4, CWDM4, FR4, DR4 @500m-2km
  - AOCs & SR4s @<100m across system rows
  - DACs in the Server/GPU racks @3-5m
Summary

- About: DACs, ACCs, AOCs, multimode + single mode transceivers
- Why so many different cables and transceivers
- Quick look at what’s coming 800GbE..3.2T
- How they are all used in various systems
Real World Use Cases

Sam Kocsis
Critical Importance by way of Functional Design

Pluggable IO solutions are mechanical interfaces that must meet electrical performance targets.

Insert too little → poor mechanically
Insert too much → poor electrically

The tolerance stack-up of the contact beam on the edge pad often defines the capabilities of the interface.
Pluggable IO Examples

Denser form factors attractive from an application perspective, but present new design challenges.
Evolution of the Performance of Pluggable IO

Pluggable IO Interface Insertion Loss (SDD21)

Evolutionary progress of the pluggable IO interface is not limited to specific form factors.

New form factors attempt to correct the shortcomings of existing interfaces, but generally focus on implementation advantages.
Cables and IO Solutions “Inside the Box”

“Inside the Box” solutions are focused on chip-to-chip links

Form factors tend to be a little bit different compared to other pluggable IO solutions discussed today

The same tolerance stack-up applies to these IO solutions as well

Interaction of the fixed and free sides of the interface significantly impact the capabilities of the interface
The interconnect ecosystem has done an amazing job continuing to advance technology and innovate with new system architectures

What is next...?
Introduction of Cabled-Hosts: Add MORE Interconnect!

Take advantage of favorable insertion loss of cable compared to PCB

Eliminate some of the painful interface points in a traditional link

Enable a new paradigm in design methods

Cabled-Hosts highlight the critical importance of interconnect solutions from end-to-end
Summary

- Pluggable connectors come in a variety of sizes and lane widths
- Interconnects are cost optimized for reach
  - Copper for cost optimized, short reach
  - Optical for medium to longer reach
- There are a variety of interconnects and signal modulation
- How to decipher cabling and transceivers
- Typical use cases
After this Webcast

- Please rate this webcast and provide us with your feedback
- This webcast and a copy of the slides will be available at the SNIA Educational Library [https://www.snia.org/educational-library](https://www.snia.org/educational-library)
- A Q&A from this webcast, including answers to questions we couldn’t get to today, will be posted on our blog at [https://sniansfblog.org/](https://sniansfblog.org/)
- Follow us on Twitter [@SNIANSF](https://twitter.com/SNIANSF)
Thank You