SNIA NSF | NETWORKING STORAGE

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

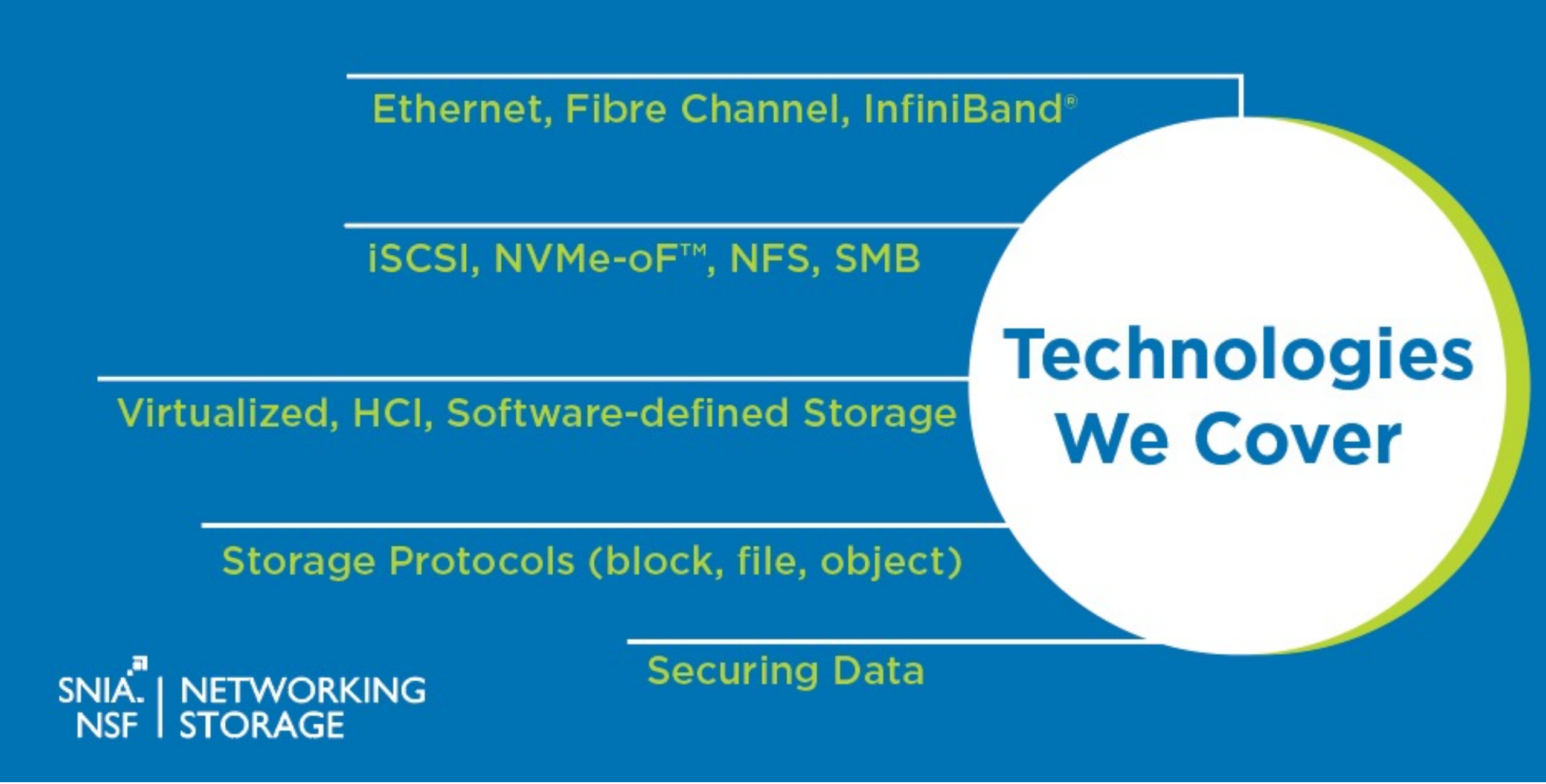
SNIA-at-a-Glance

180
industry leading organizations

2,500
active contributing members

50,000
IT end users & storage pros worldwide

Learn more: **snia.org/technical** **@SNIA**



SNIA Legal Notice

- 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

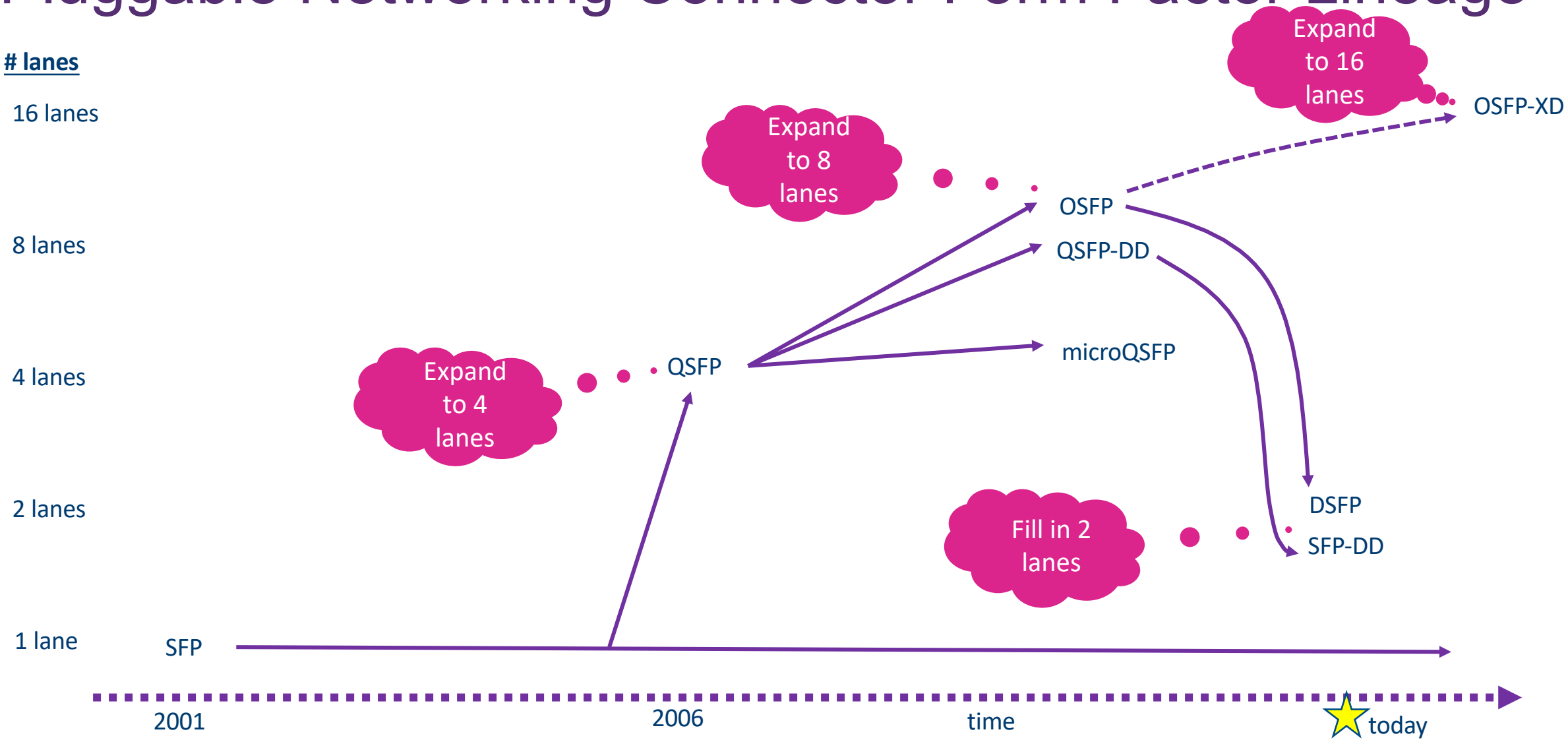
Common Pluggable Connectors and Media Types

Kent Lusted

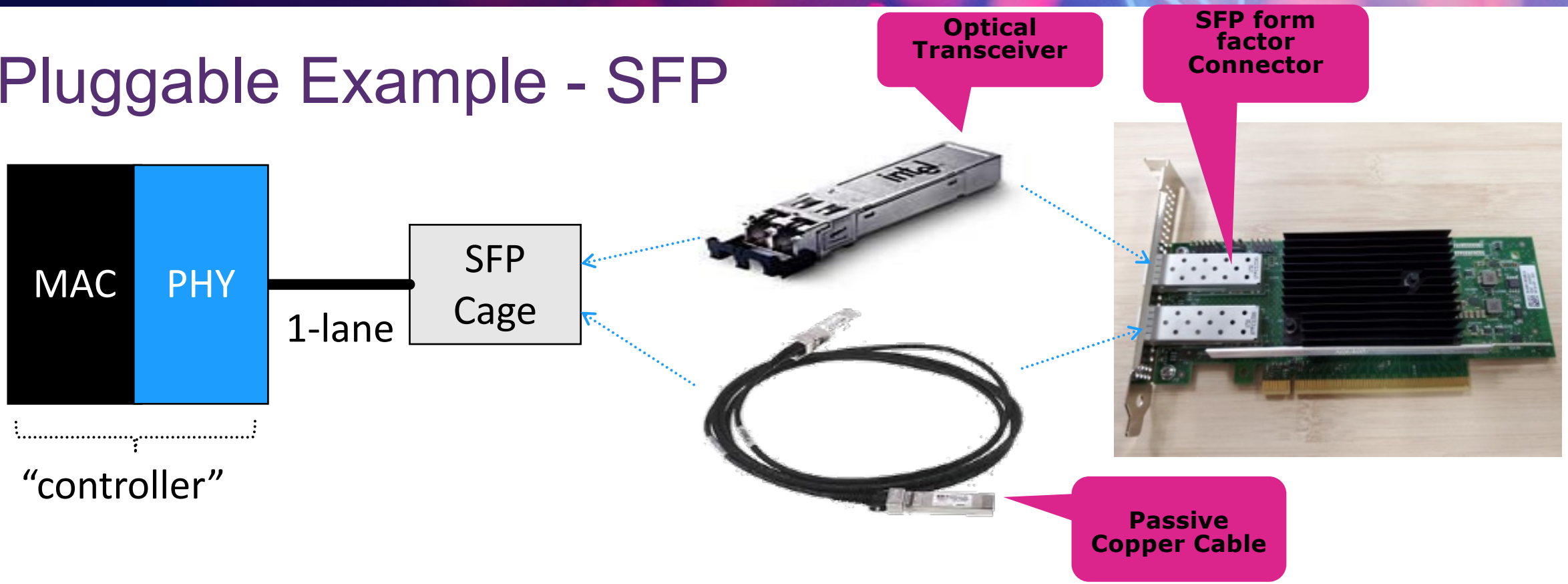
Introduction

- 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

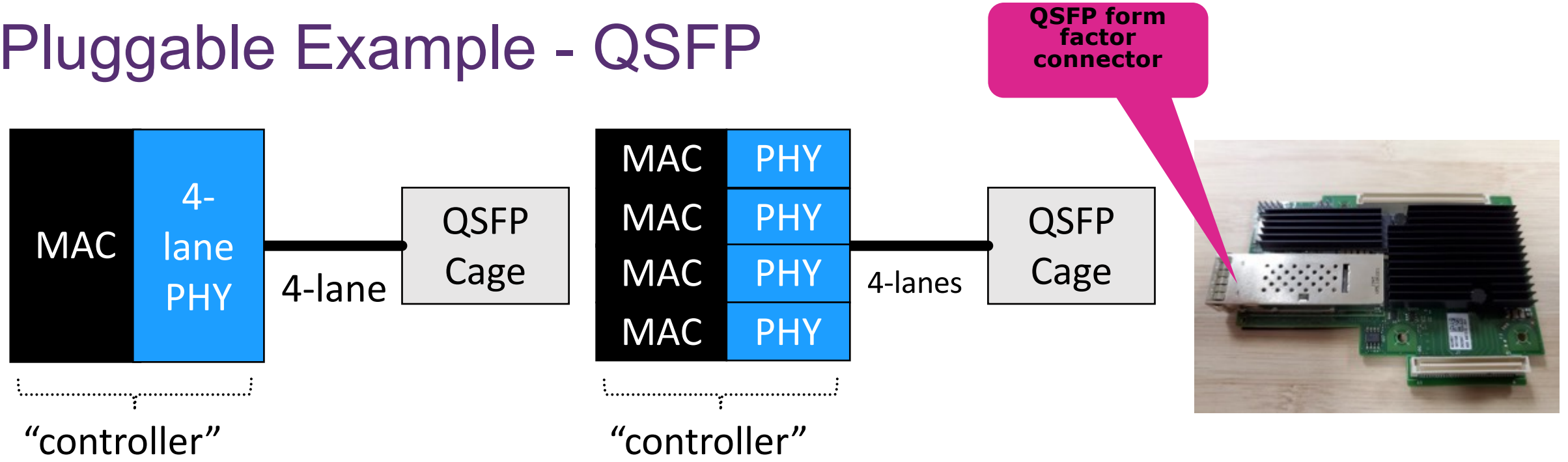


Pluggable Example - SFP



- 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.

Pluggable Example - QSFP



- 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

Photo courtesy, Dan Case (Intel)

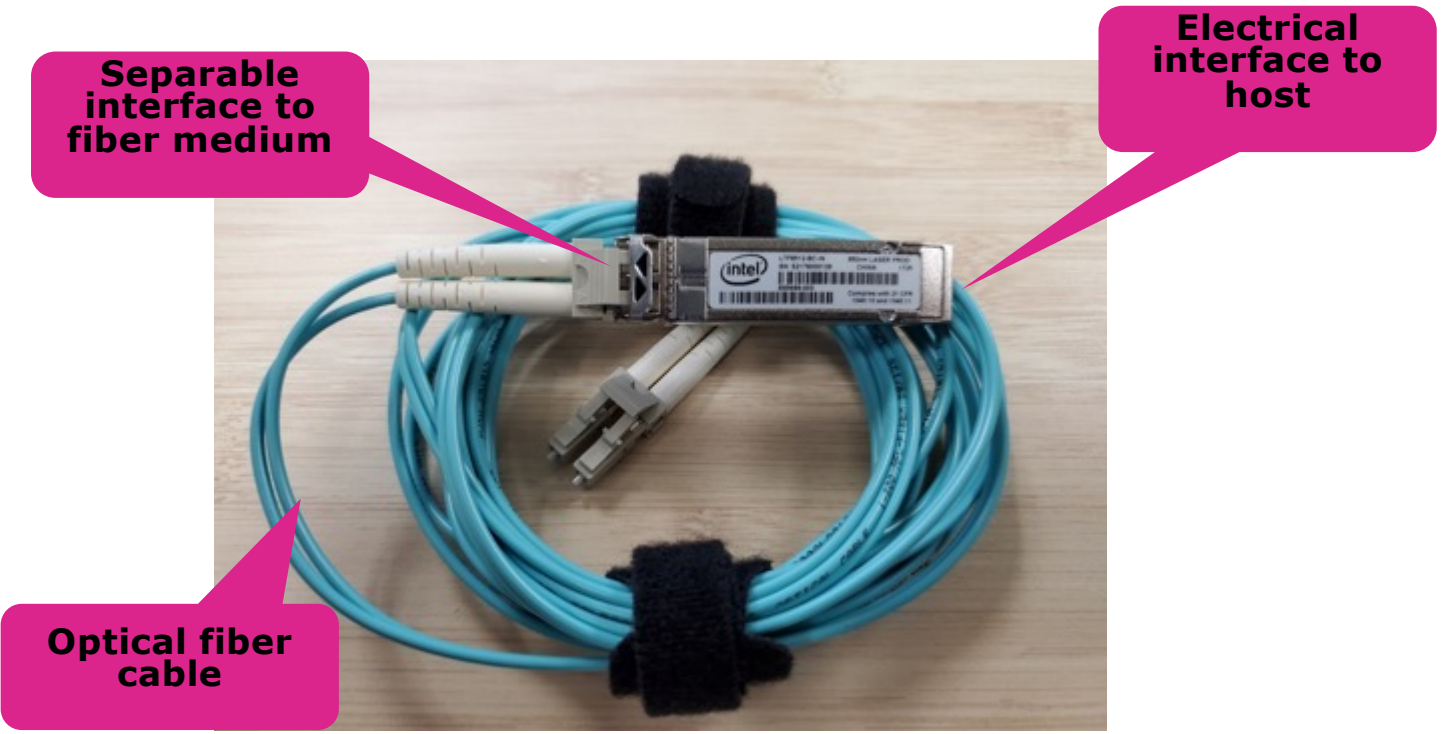
- Twin-axial Breakout Cable
  - Point to multi-point

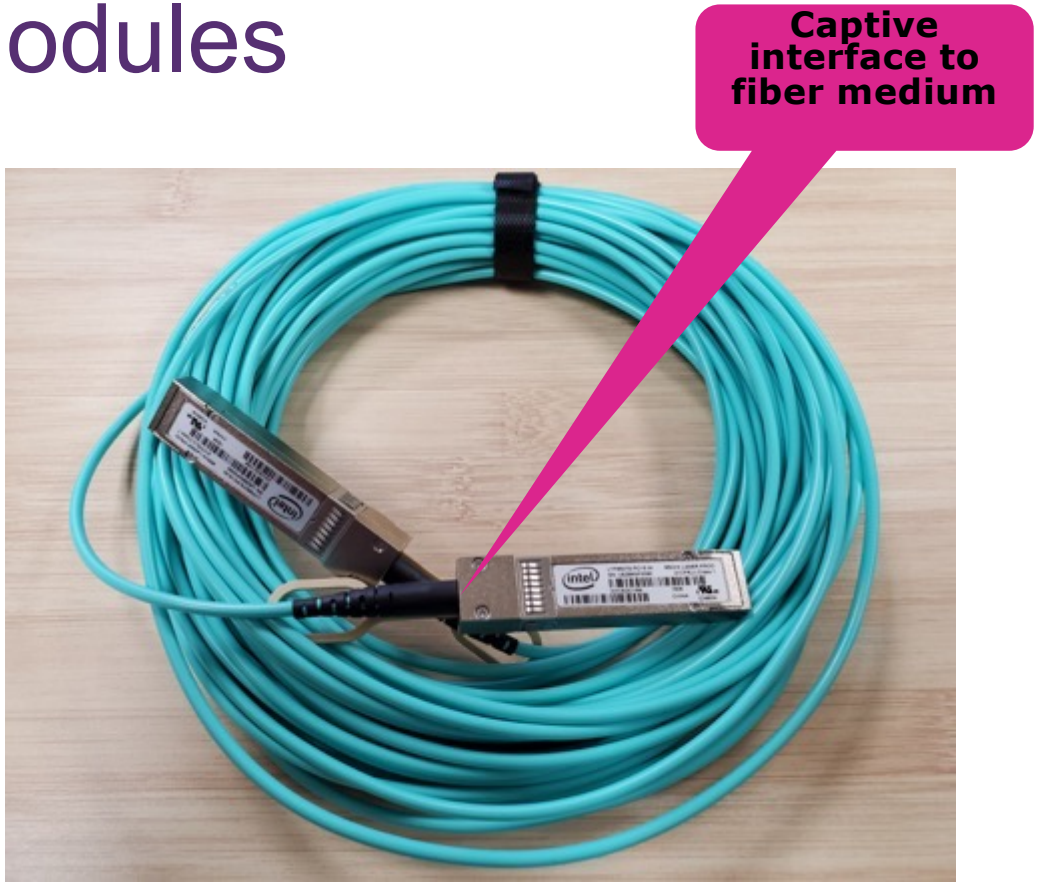
Photo courtesy, Dan Case (Intel)

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 has active devices such as redrivers or retimers inside the connector

Pluggable Optical Transceiver Modules

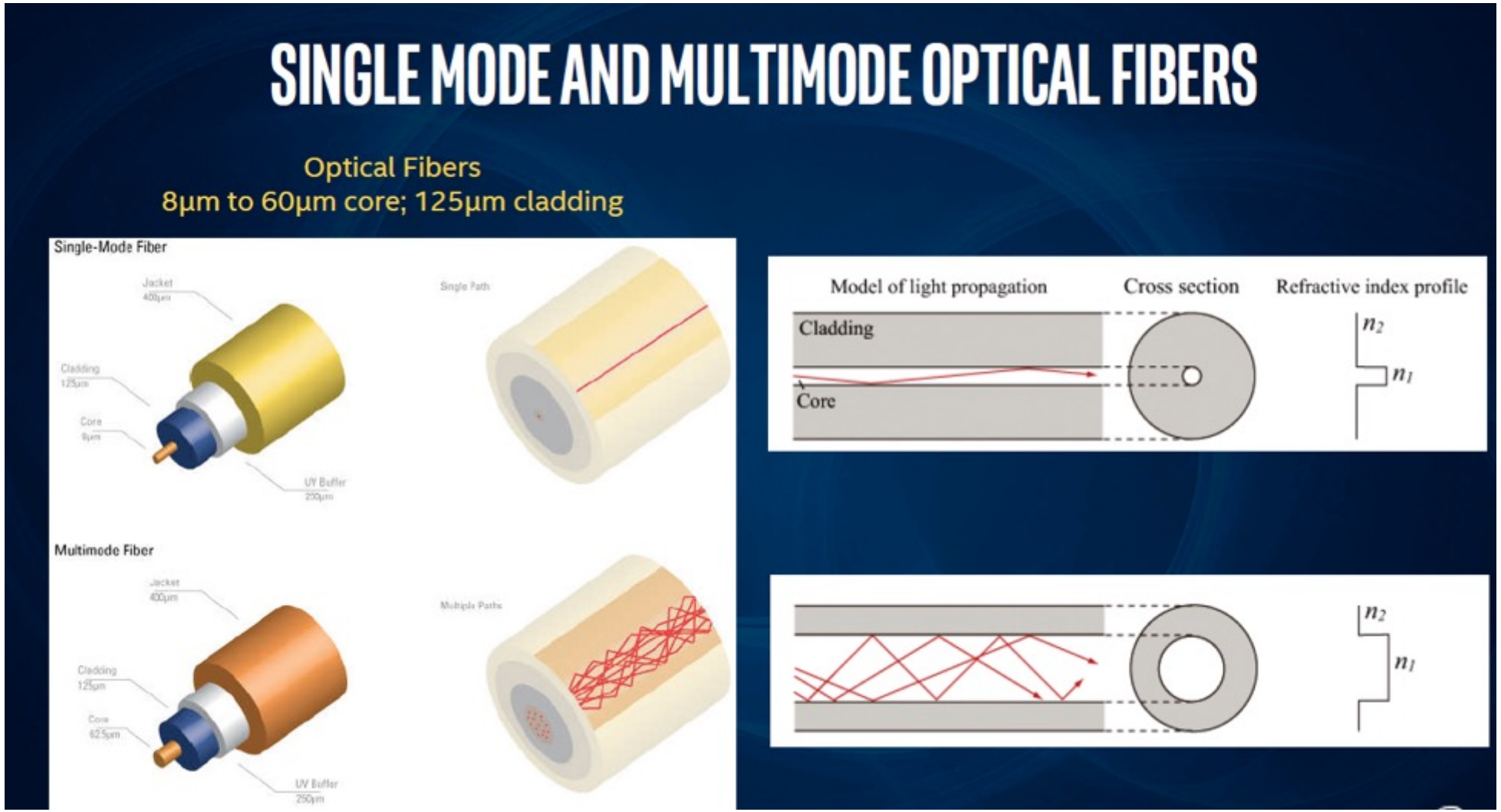




Pluggable Transceiver is a module with an electrical connector in one end (to host) and optical connector in the other end (to the medium/fiber). It can have one or more parallel lanes in each direction (transmit and receive)

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

Optical Fiber Types



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
  - First came the Ethernet Technology Consortium (ETC) version of 25GbE: 25GBASE-CR1
  - 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

			Supported PHY types		
Cable Type	**Reach**	**FEC Modes**	IEEE 25GBASE-CR	IEEE 25GBASE-CR-S	Ethernet Technology Consortium 25GBASE-CR1
CA-25G-L	5m	RS(528)	YES	NO	NO
CA-25G-S	3m	RS(528), BASE-R	YES	YES	NO
CA-25G-N	3m	RS(528), BASE-R, No FEC	YES	YES	YES

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

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

Pre-emphasis Or DSP IC

AKA:
"Active DAC"
"Active Electric Cables (AEC)"

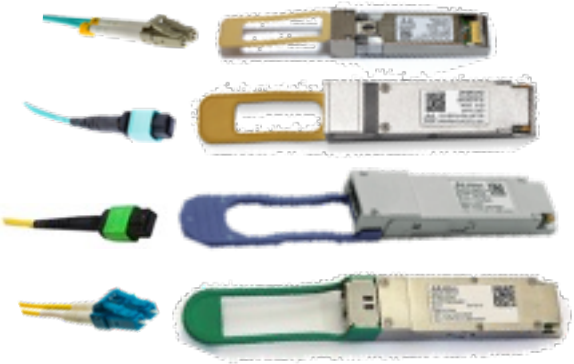
Active Optical Cables (AOC)

2 Transceivers w/optical fibers bonded inside.
Lowest-Priced Optical Link
Up to 100m Reaches

Transceivers with Integrated Fibers

Optical Transceivers

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



Transceivers with Detachable Optical Fiber Connectors

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

Transceivers with Integrated Fibers

Inside DAC Cables

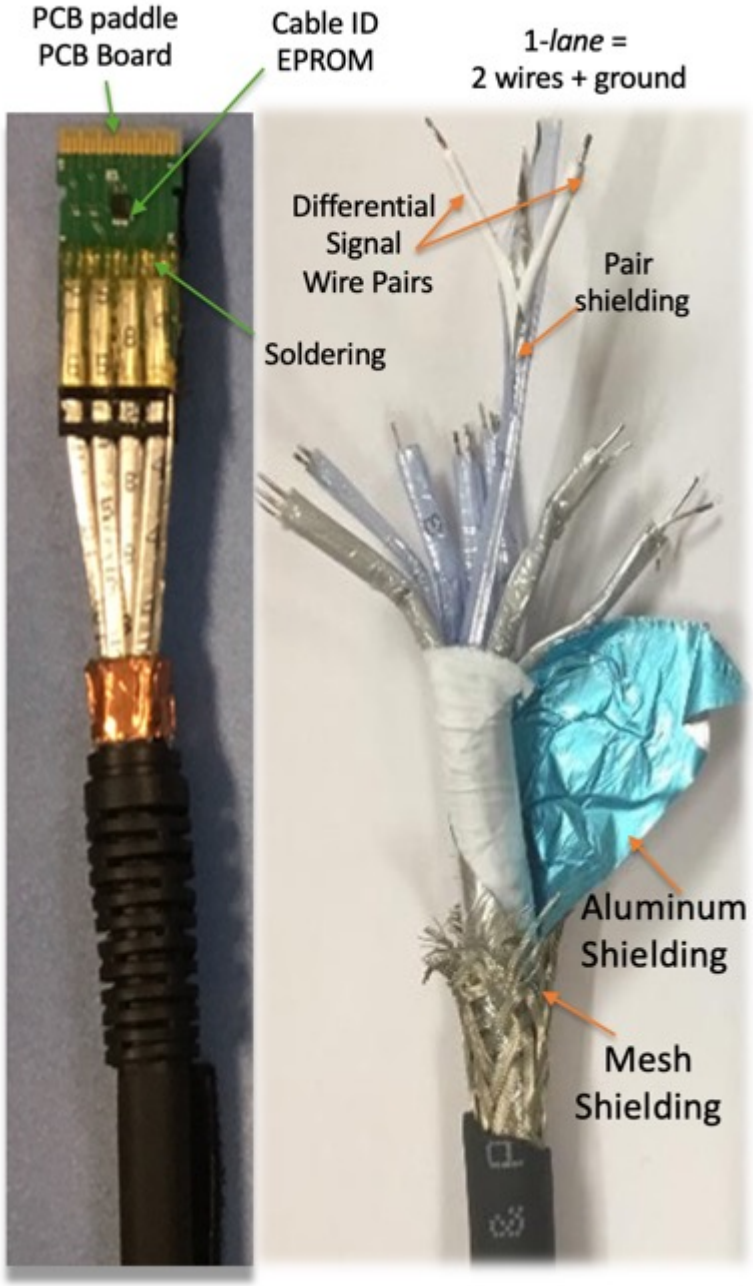
Simplest high-speed interconnect

2x400G DAC has 48 wires
3 wires/direction (2+ground)
8-channels x 2 directions x 3 wires = 48 wires

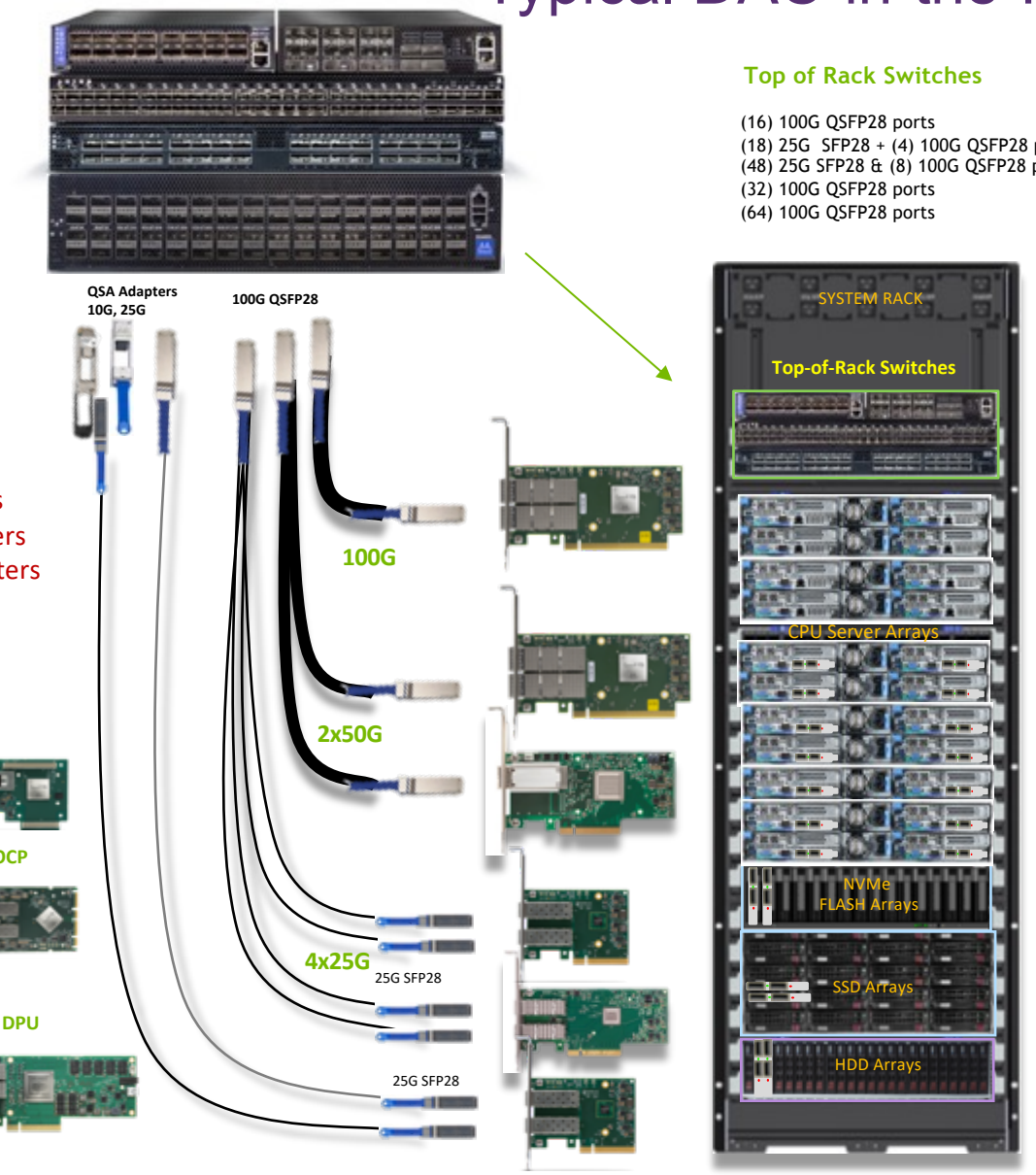
Wires, shielding, PCB, soldier, shell, label

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

4-Channel 100GbE, EDR DAC



Typical DAC-in-the-Rack Use



Passive
Copper DACs

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

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

Advantages:

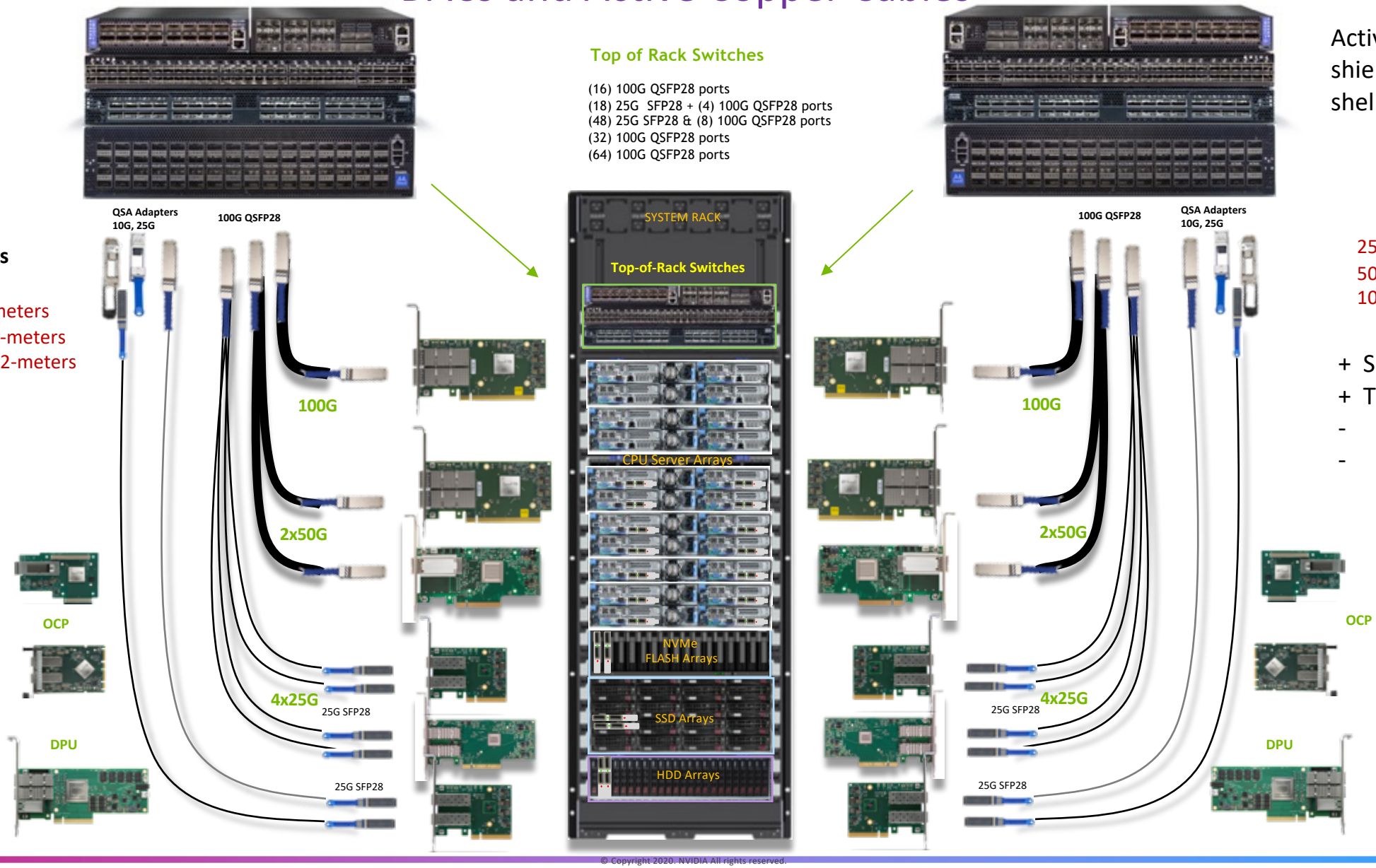
- Zero-latency
- Zero-power consumption
- Lowest-cost
- Simplicity
- High-reliability
- Long reach

Disadvantages:

- Heavy
- Bulky
- Shorter with faster line rates

25G-NRZ = up to 5-meters

DACs and Active Copper Cables



Passive Copper DACs

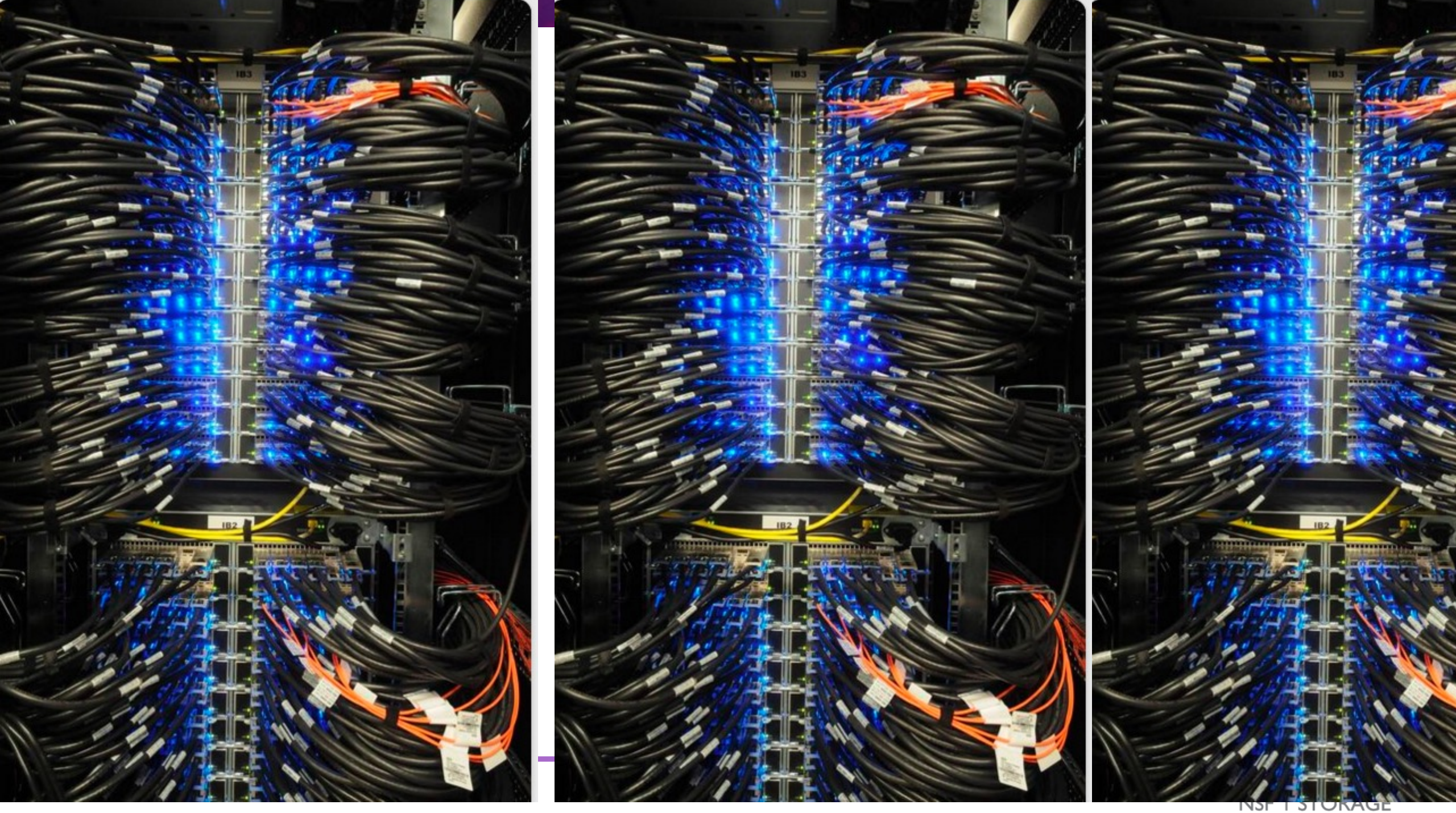
25G-NRZ up to 5-meters
50G-PAM4 up to 3-meters
100G-PAM4 up to 2-meters

Active DACs = Wires, shielding, PCB, soldier, shell, label **+ ICs**

Active Copper DACs

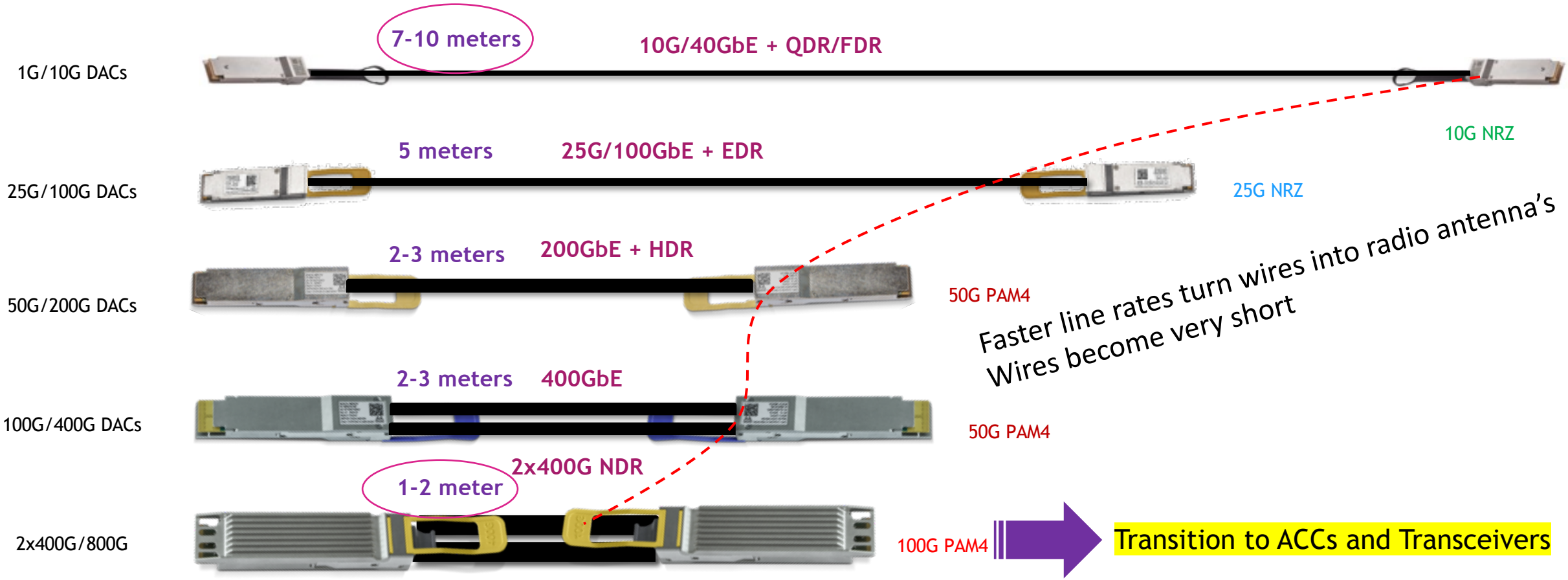
25G-NRZ up to 7-meters
50G-PAM4 up to 5-meters
100G-PAM4 up to 3-meters

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


IB3
IB2

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



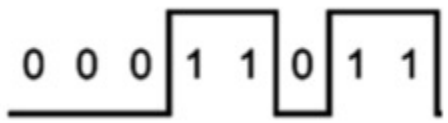
Signal Modulation: NRZ and PAM4

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

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

Bits Clock

25/100G Same Intensity Level



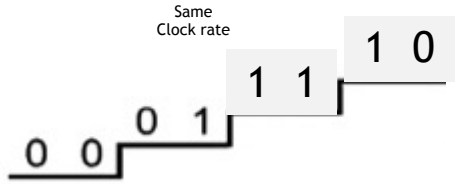
25G-NRZ
(Non-return to Zero)

1
0

25GHz

Traditional one data bit (1,0s); 1-bit x 25GHz clock = 25Gb/s

200G Same Intensity Level

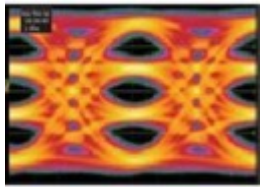


50G-PAM4

1,0
1,1
0,1
0,0

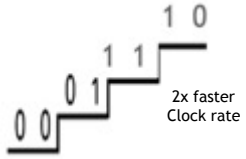
25GHz

Stacks 2 data bits on top of each other, same 25GHz clock; 2-bits x 25GHz = 50Gb/s



(Pulse Amplitude Modulation 4-levels)

400G Same Intensity Level



100G-PAM4

1,0
1,1
0,1
0,0

50GHz

Twice faster clock than 50G-PAM4, 2-bits x 50GHz = **100Gb/s**

Applies to both Electrical and Optical Modulation

Form-factor Buzzword-ology

Why all the new form-factors?

- SFP = **S**mall **F**orm-factor **P**lug (1-channel)

- **Q**uad SFP (4-channels)

- QSFP**-DD** = Double Density (8-channels)
-
- **OSFP = O**ctal SFP (8-channels)

SFP+ = 1x10G SFP28 = 1x25G SFP56 = 1x50G **SFP112** = 1x100G

QSFP+ = 4x10G QSFP28 = 4x25G QSFP56 = 4x50G QSFP112 = 4x100G

Max line rates

NEW

SFP-DD = **2**x50G, 2x100G

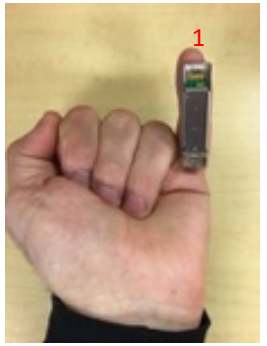
QSFP56-DD = **8**x50G,

QSFP112-DD = **8**x100G

OSFP = **8**x50G, 8x100G

OSFP-XD = 16x100G, 16x200G

SFP



QSFP

QSFP = 4-channel

OSFP

OSFP = 8-channel

MSA: Multiple Source Agreement

Industry standard. enclosures for transceiver

Holds wires and transceiver components.

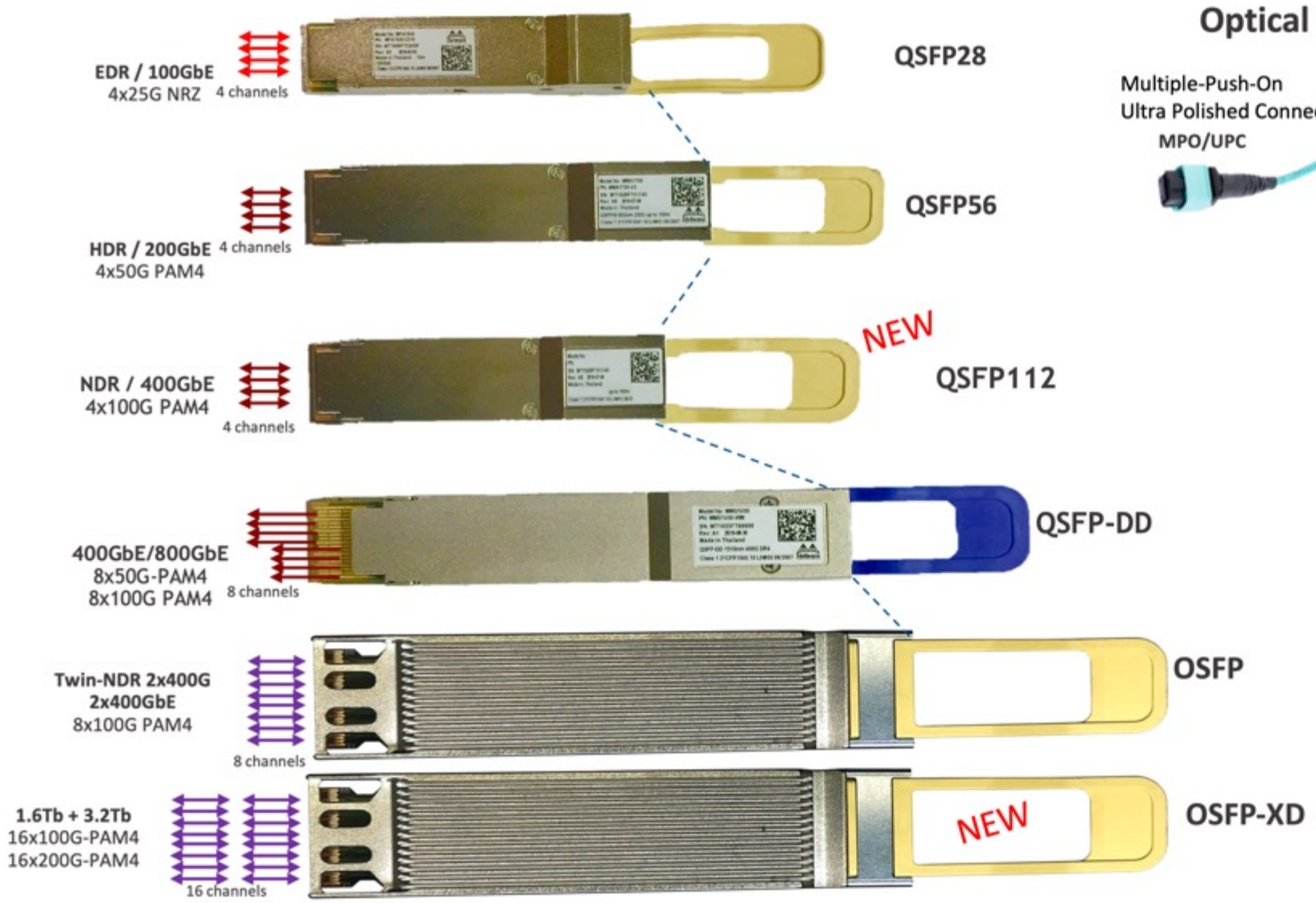
Zinc/Copper Shells

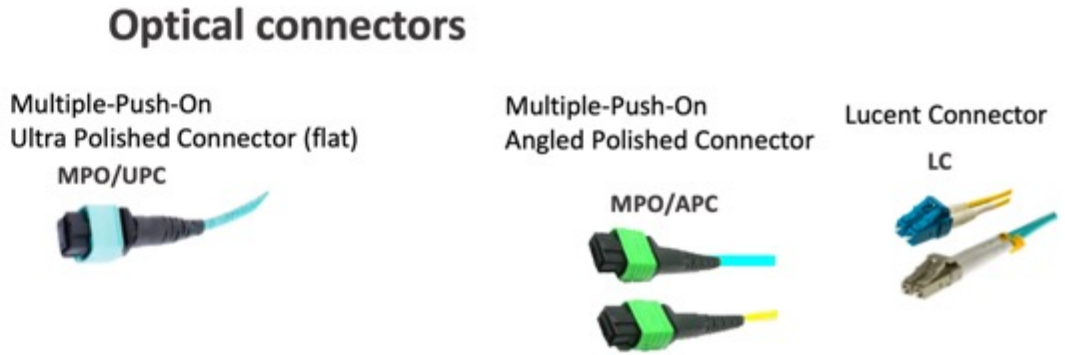
Can be used with

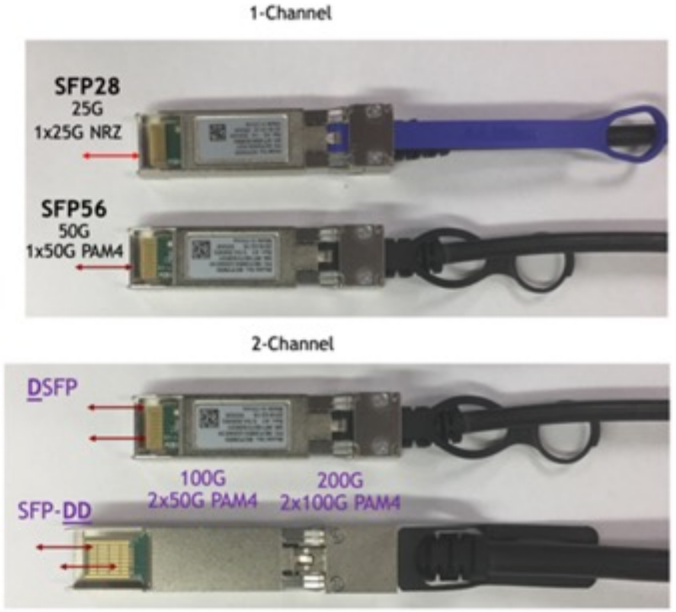
DAC, AOCS, MM or SM transceivers.

Speed rated (EFI limit)

Data center : QSFP + OSFP Transceivers and Form-factors







Single mode = yellow
Multimode = Aqua + Tan

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

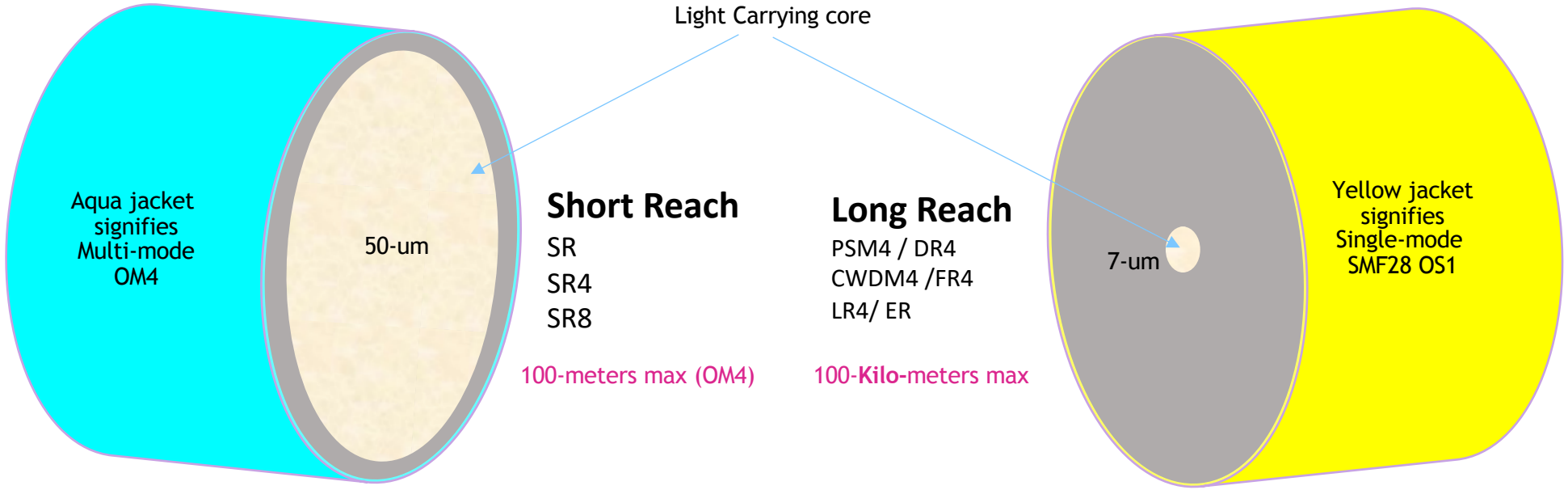
Multi-mode & Single-mode Fibers

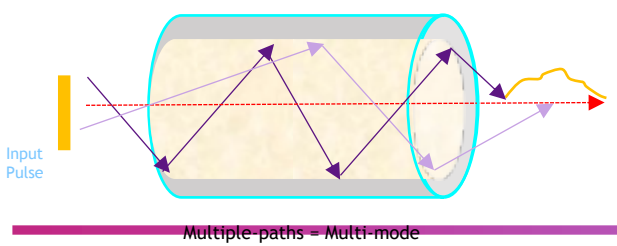
Multi-mode

- *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
- o Used with only one signal at 850nm wavelength
- o Used with SR, SR4, SR8 transceivers

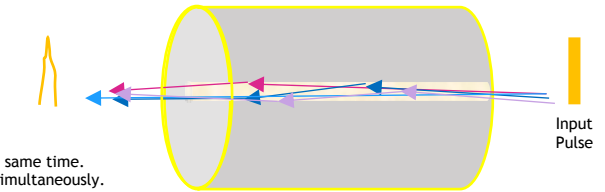
Single-mode

- *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
- o Can multiplex 4 or 8 signals simultaneously into 1 fiber at 1310nm wavelength
- o Used with DR1, DR4, FR4, LR4 transceivers





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

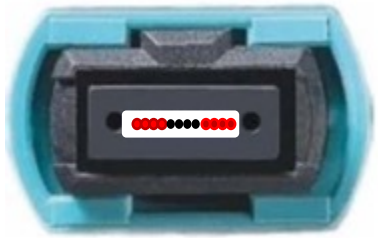


Pulse light rays stay concentrated, intense & arrive at same time.
Multiple colored wavelength signals can be transmitted simultaneously.
Max lengths up to 40km and 4-8 channels.

Optical Connectors

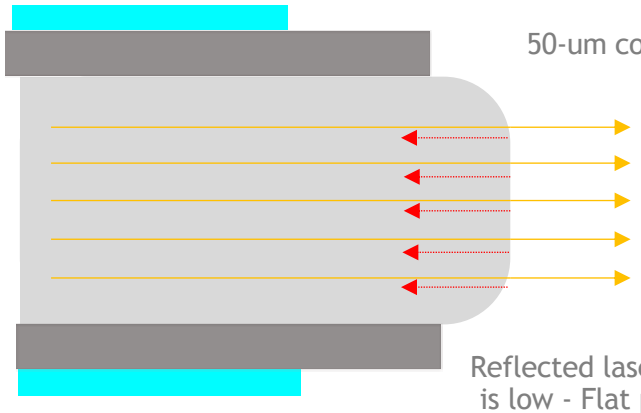
25G-NRZ, 50G-PAM4

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



Aqua fiber = Multi-mode

-55dB back reflection



Used with:
- 100GbE/**EDR** SR4 transceivers
- 200GbE/**HDR** SR4 transceivers

Multiple-Push-On
Angled Polished Connector
MPO/APC

MPO-12
Using only 8-fibers

Transmit Fibers | Unused Fibers | Receive Fibers

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

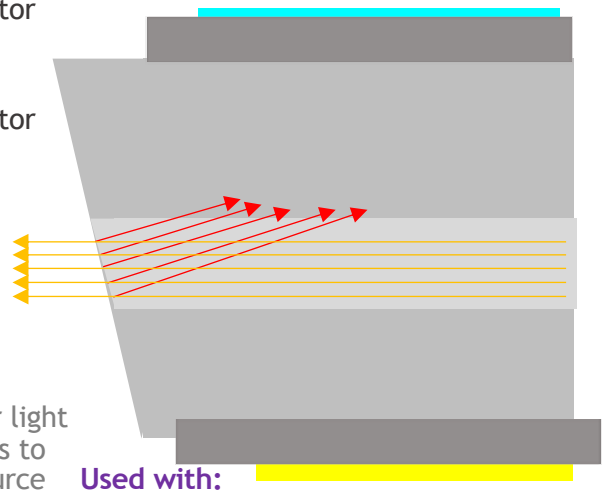
100G-PAM4

Green shell = APC

-65dB back reflection

Single-mode fiber
Angled Polished Connector
7-um core

Multi-mode fiber
Angled Polished Connector
50-um core



Concentrated laser light
reflected off axis to
protect laser source

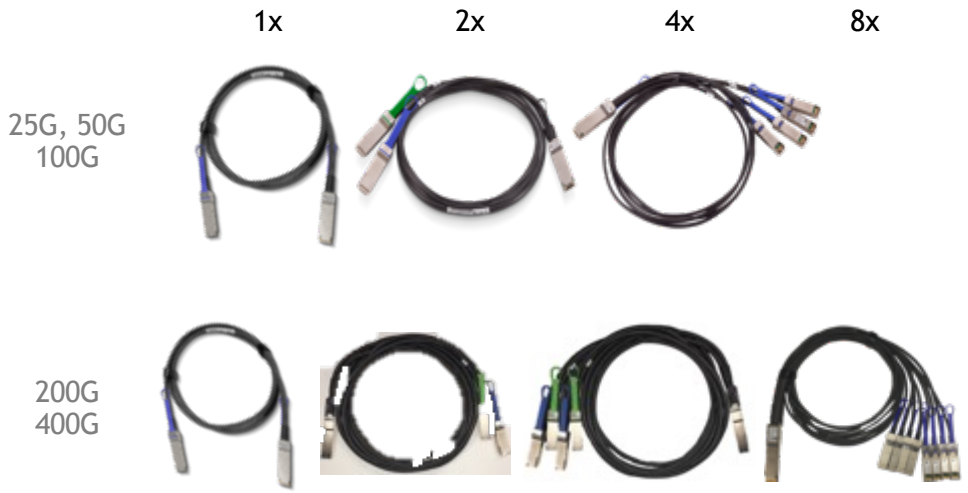
Used with:
- 400G/800GbE **SR8, DR4** transceivers
- **NDR** transceivers
  both single-mode & multi-mode

> 500 Different Cables & Transceivers Products

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

Direct Attach Copper (DAC)

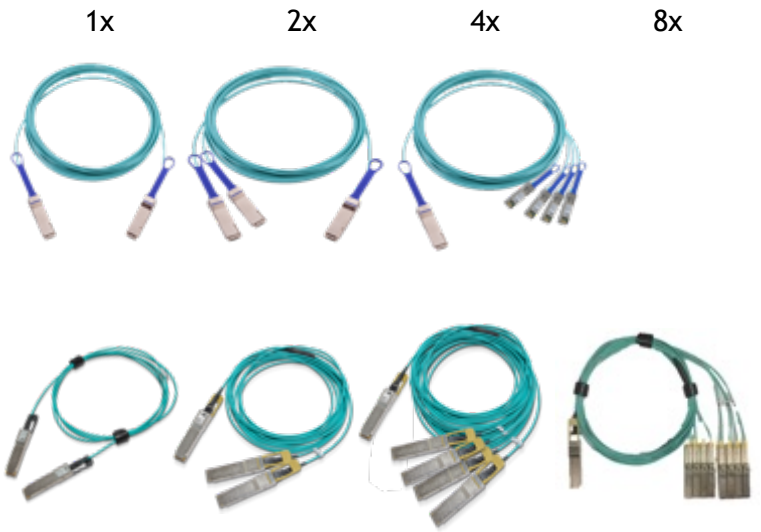
Copper wires "Directly Attaches" system together
Lowest Priced, Zero-Latency, Zero-Power
Up to 2,3 and 5m Reaches



Copper Cables
Active Copper Cables

Active Optical Cables (AOC)

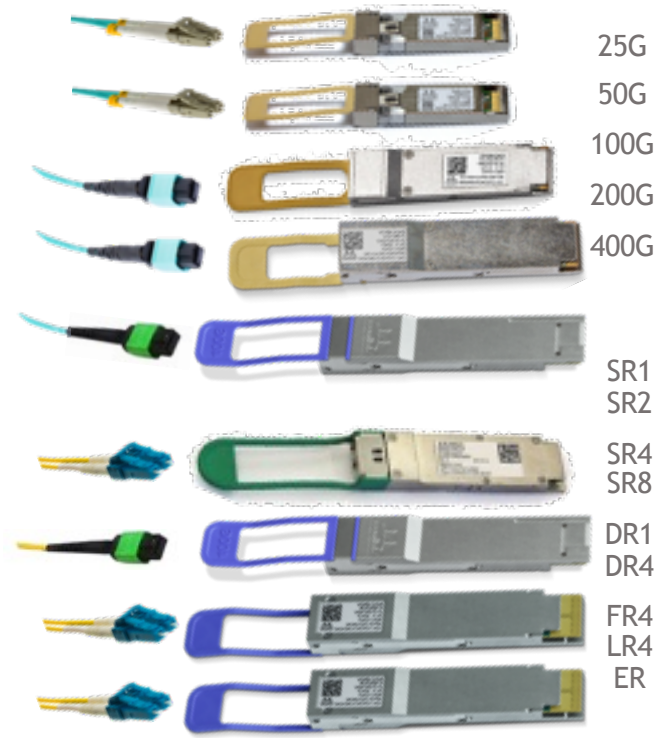
2 Transceivers w/optical fibers bonded inside.
Lowest-Priced Optical Link
Up to 30m and150m Reaches



Transceivers with
Integrated Fibers

Optical Transceivers

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



Transceivers with Detachable
Optical Connectors

Interconnects are Cost Optimized for Each Reach

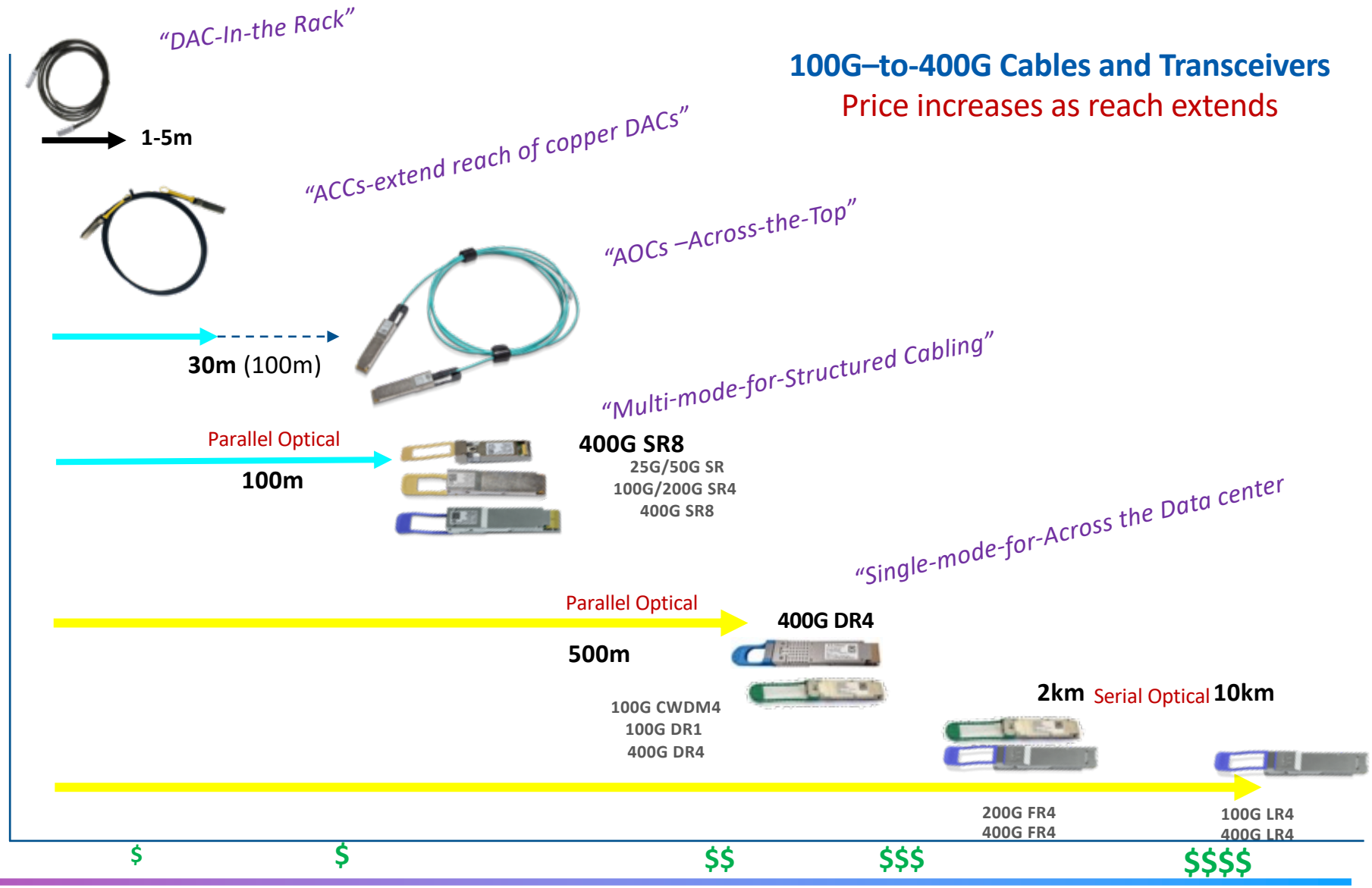
DAC Copper Cables (DACs)

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

Optical Descriptors and Codes

Multi-mode transceivers up to100-meter reach

- **SR:** Short-Reach 1-channel *NDR current 30-meters*
- **SR4:** Short-Reach 4-channel
- **SR8:** Short-Reach 8-channel

Single-mode transceivers 500m, 2km, 10km reach

- **DR1:** Datacenter-Reach 1-channel (500m)
- **DR4:** Datacenter-Reach 4-channel (500m) New PSM4
- **DR8:** Datacenter-Reach 8-channel (500m)
- **FR4:** Far-Reach 4-channel (2km)(multiplexed) New CWDM4
- **LR4:** Long-Reach 4-channel (10km)(multiplexed) 100G FR4 (CWDM4), 200G FR4, 400G FR4, 2x400G FR4

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

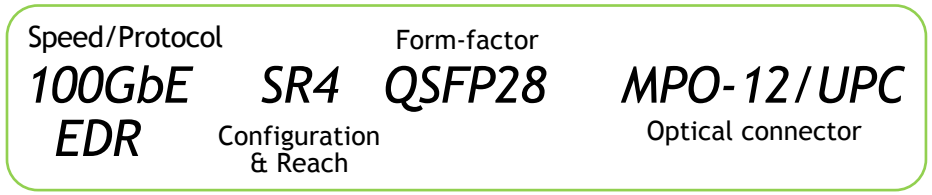
Optical fibers:

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)

How it all goes together



Multimode Optics: Transceivers or AOCs

25G SR SFP28

100G SR4 QSFP28

QSA Port Adapter

LC

MPO

Linking within System Racks

100G

100G QSFP28

2x50G

50G QSFP28

4x25G

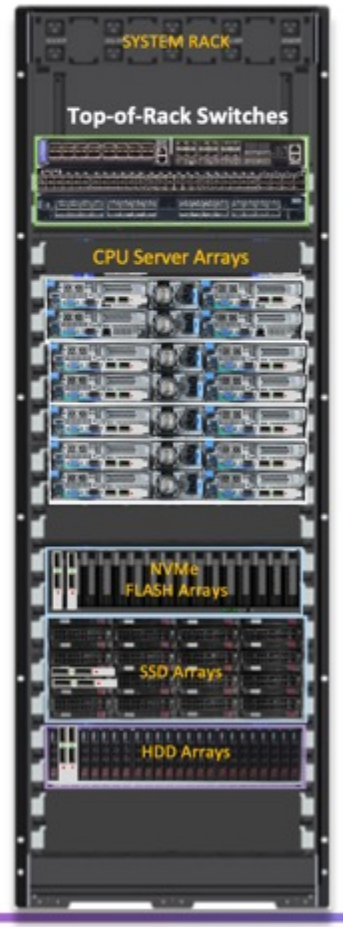
25G SFP28

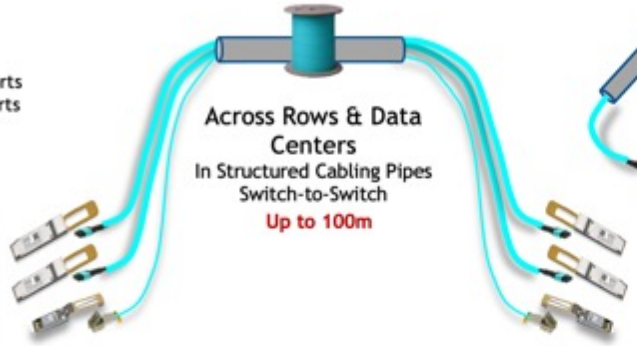
25G SFP28

1x25G

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





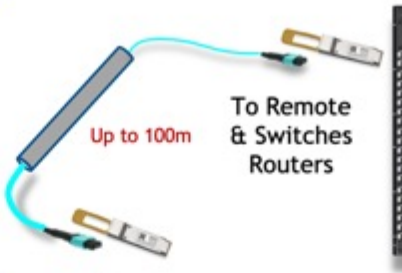
Across Rows & Data Centers
In Structured Cabling Pipes
Switch-to-Switch
Up to 100m

Linking Between Systems

Short Reach Transceivers

25G SR 100m
100G SR4 100m
100G SWDM4 100m
200G SR4 100m
400G SR8 100m
2x400G SR8 50m

Up to 100m

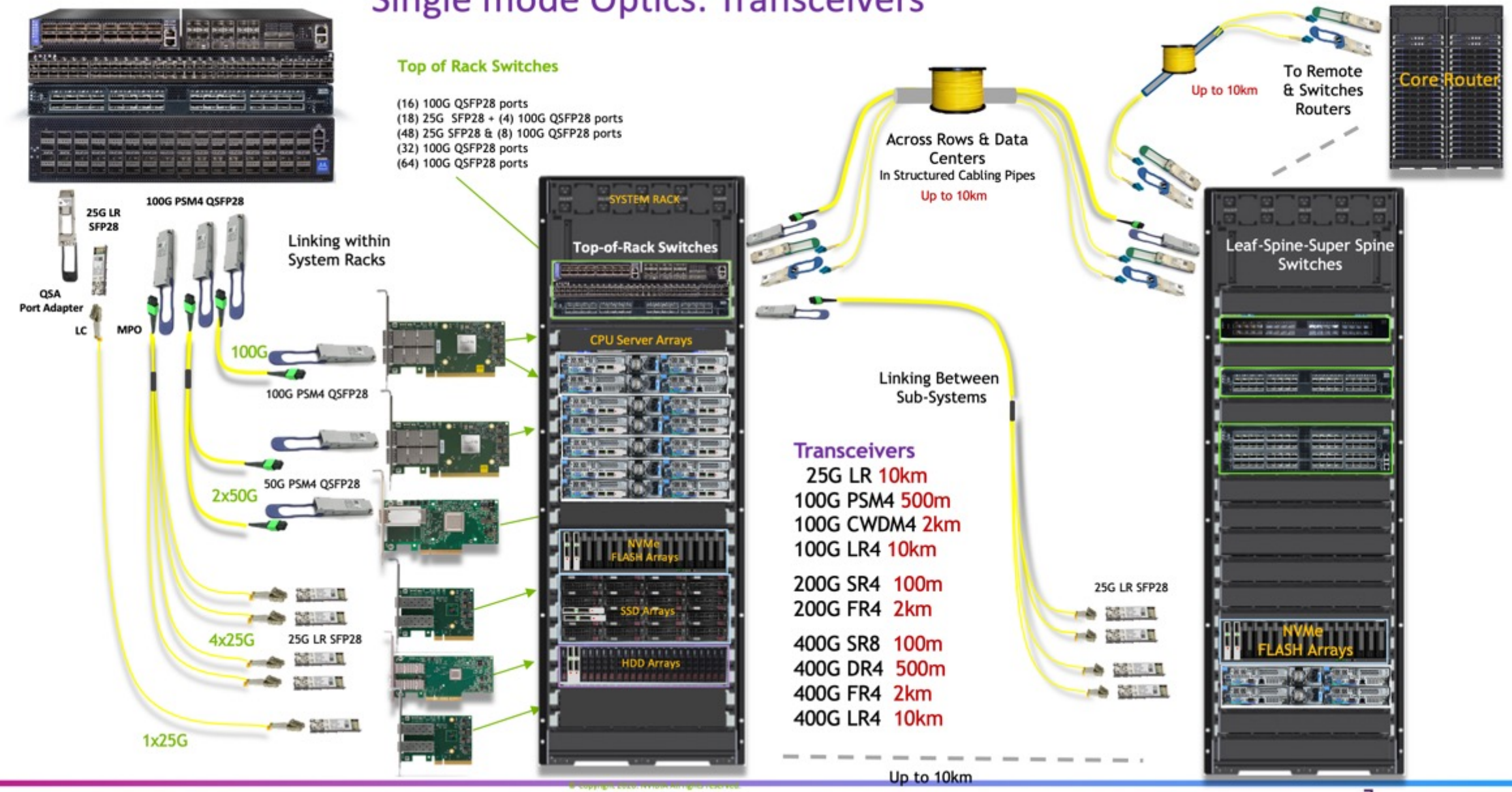


Up to 100m

To Remote & Switches Routers

Router

Single mode Optics: Transceivers



Cables & Transceivers for 100g/400GbE Networks

FOR SHORT REACH, LONG REACH, STRUCTURED & BREAKOUT CABLING

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

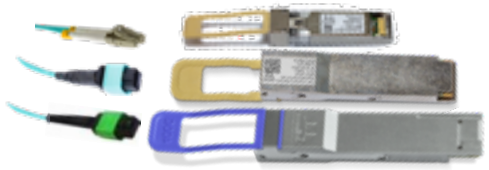
AOC
Between Switches
Within Racks
3-30m…100m

SR/SR4/SR8
For structured cabling
Between Racks
Up to 100m

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



2-fiber Serial

8, 16- fiber Parallel

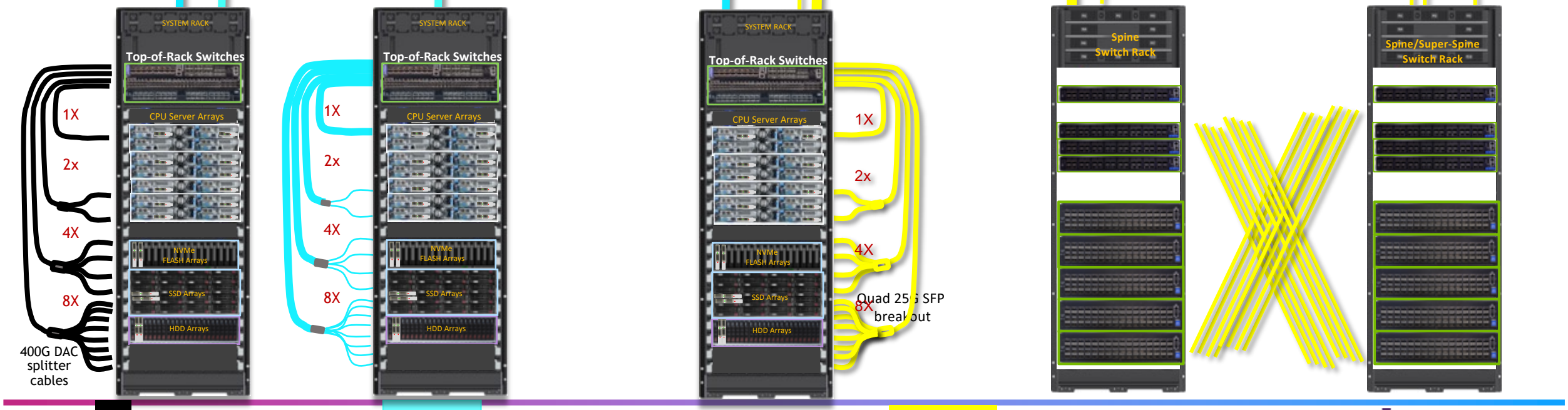
DAC

AOC: 3-100m

Splitter Cables & Transceivers
For Inside & Between Racks

8-fiber Parallel

Single-mode For Structured Cabling
Long Reaches
Leaf-Spine Links



DAC *"In-the-Rack"*
Up to 2.5m

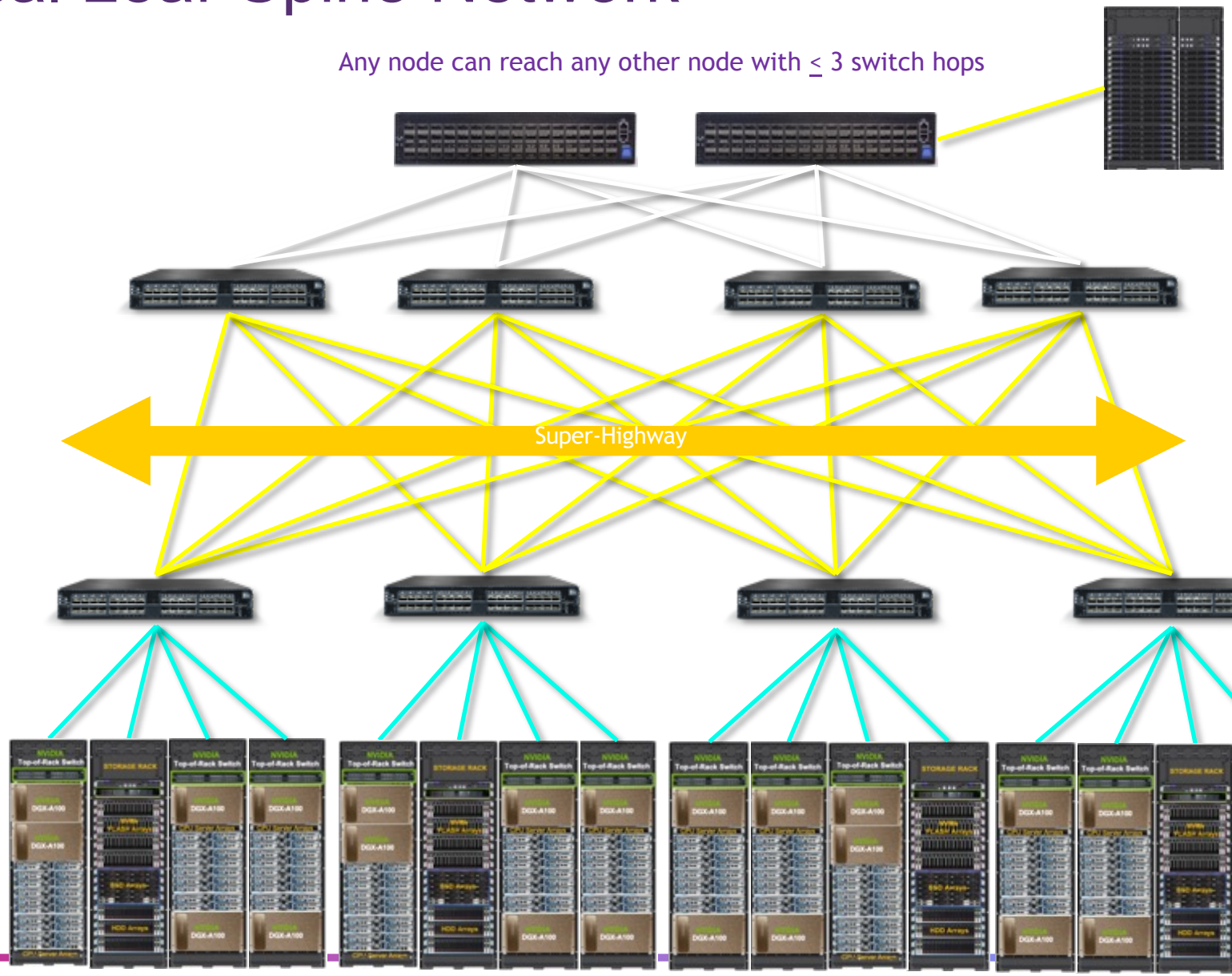
Multi-Mode Optics
3m-100m

Single-Mode Optics
Up to 10Km

SNIA NSF | NETWORKING STORAGE

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

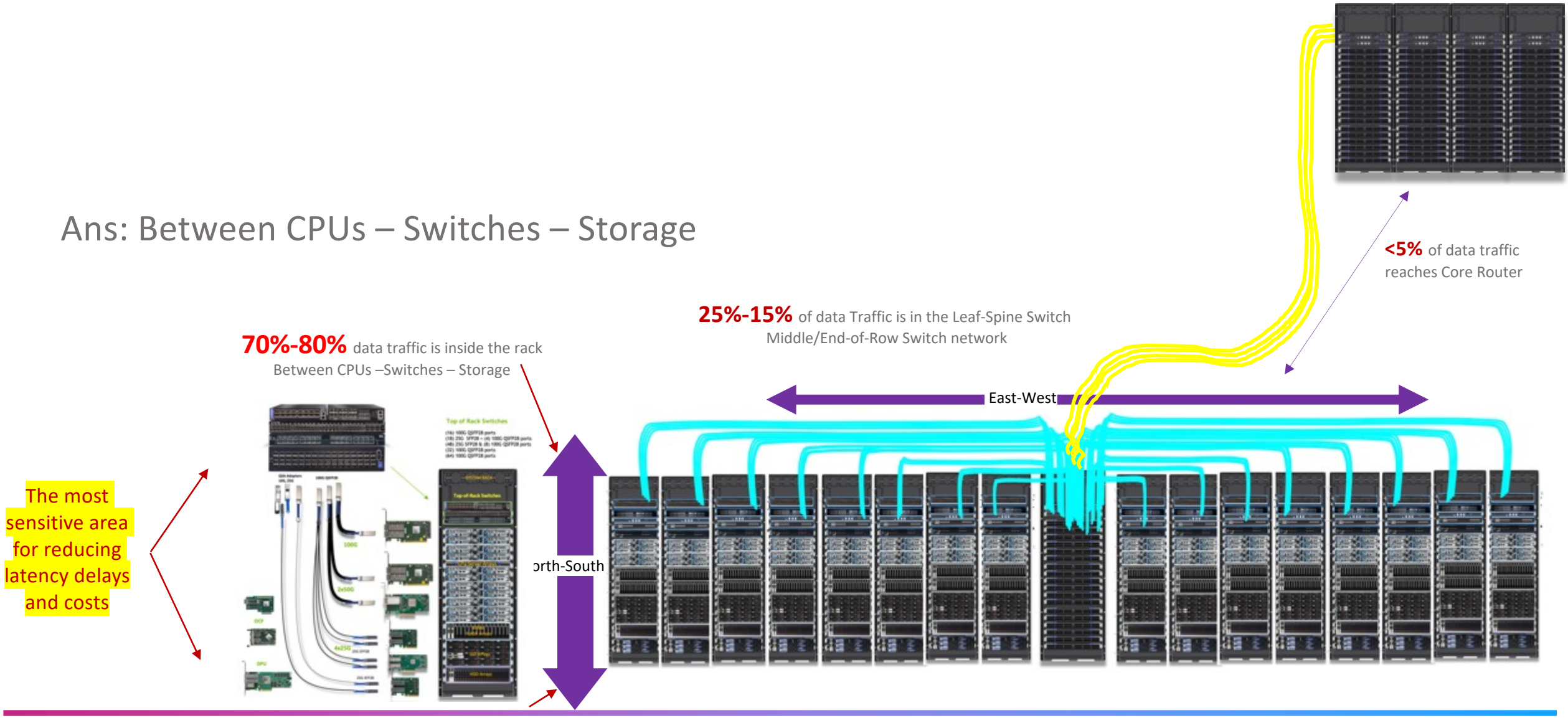
Leaf Network
Layer 2
Single-mode Transceivers:
CWDM4/FR4
PSM4/DR4
500m-2km

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

Compute/Storage
DACs, ACCs

Where is Most of the Ethernet Data Traffic?

Ans: Between CPUs – Switches – Storage



Typical Data Center Cluster

DACs-in-the-Rack

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

Where Connectorized Transceivers Used

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

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

Fujitsu Riken FUGAKU HPC - #1 in Top500

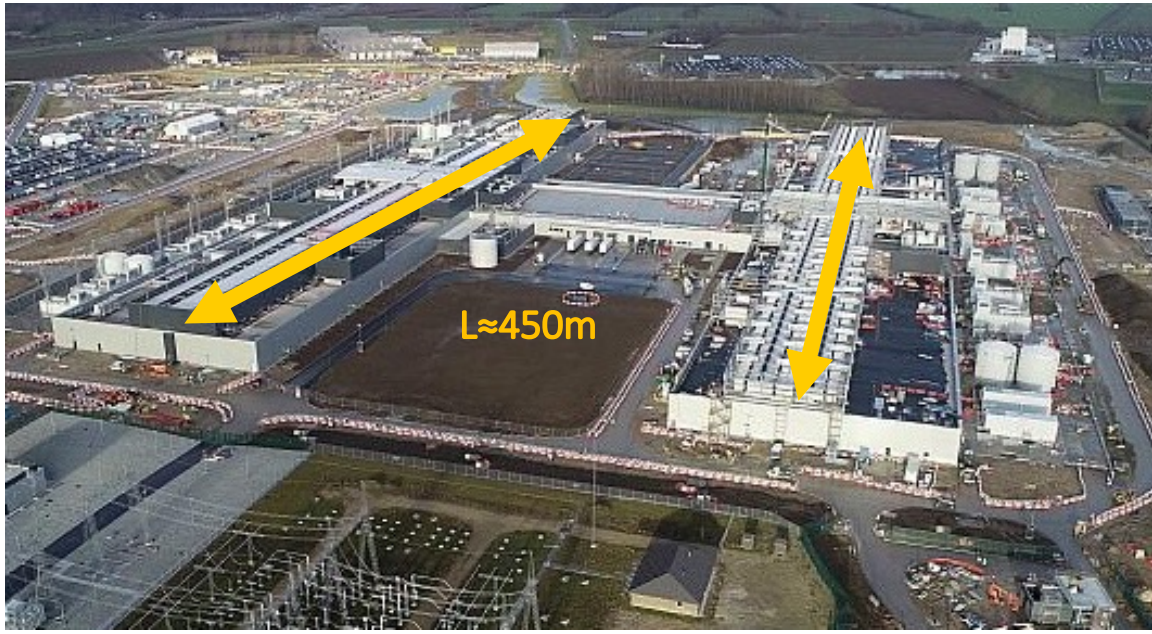
>100 thousand AOCs

Photo: Riken

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!



Facebook Datacenter, Fyn Denmark, 56.000m², 40MW

- 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

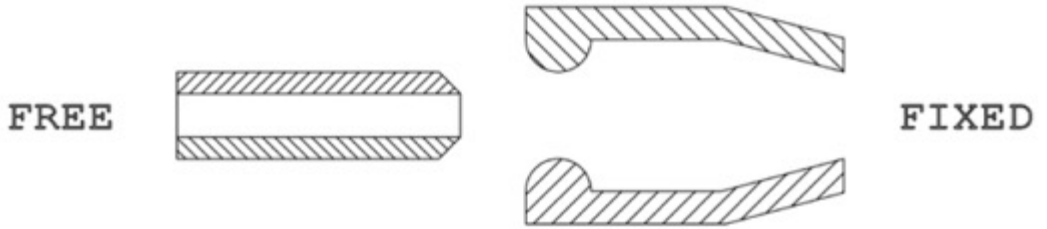
- About: DACs, ACCs, AOCs, multimode + single mode transceivers
- Why so many different cables and transceivers
- Quick look at what's coming 800GbE..3.2T
- How they are all used in various systems

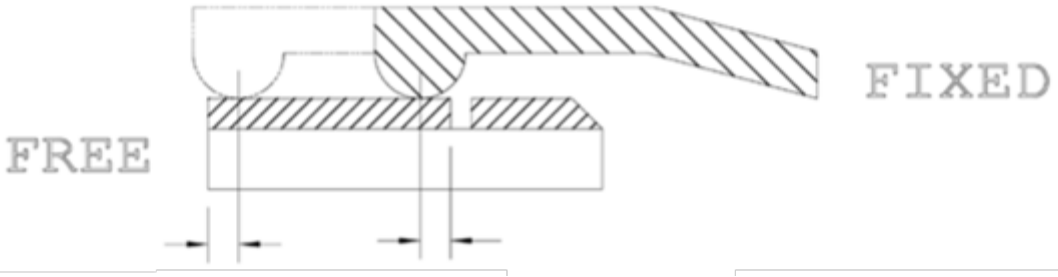
Real World Use Cases

Sam Kocsis

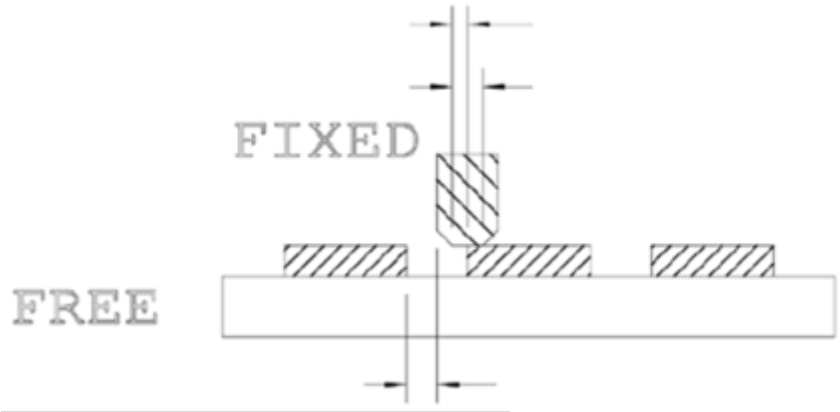
Critical Importance by way of Functional Design

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

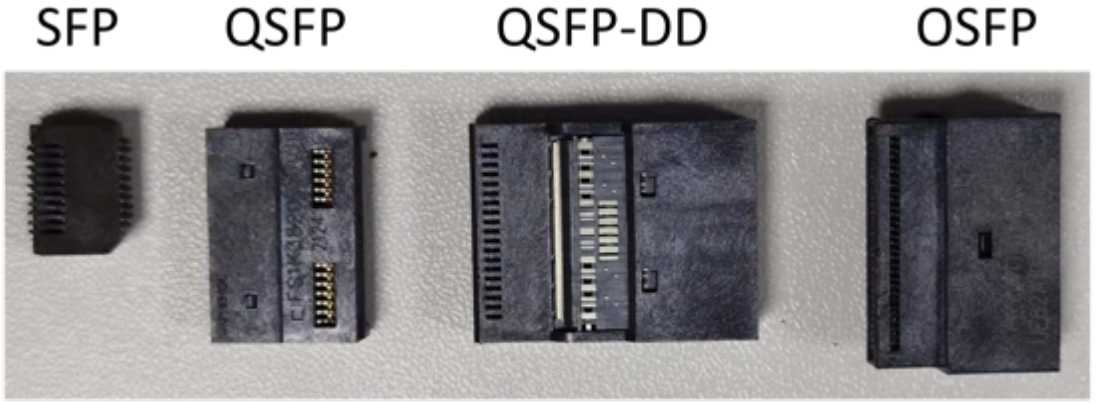
Pluggable IO Examples

SFP

QSFP

OSFP

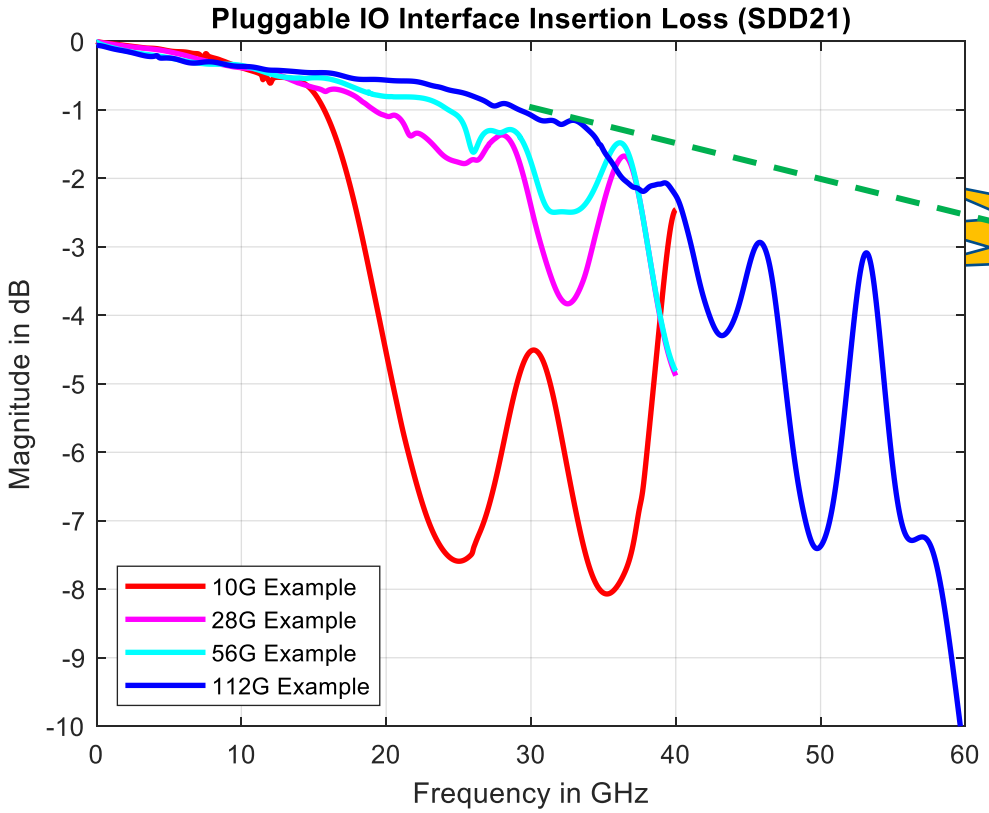
QSFP-DD



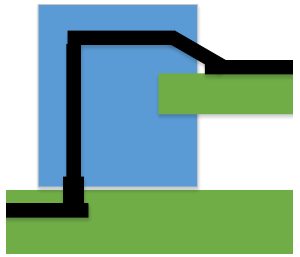
1-lane 4-lane 8-lane 8-lane

Denser form factors attractive from an application perspective, but present new design challenges

Evolution of the Performance of Pluggable IO



224G Example?!



Pluggable IO Interface =
Mating interface +
Connector Body +
Attachment to PCB

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

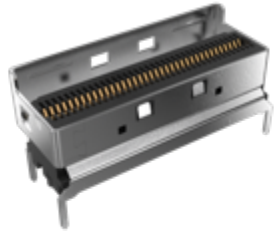
Cables and IO Solutions “Inside the Box”

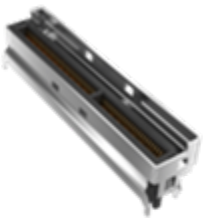
“Inside the Box” solutions are focused on chip-to-chip links

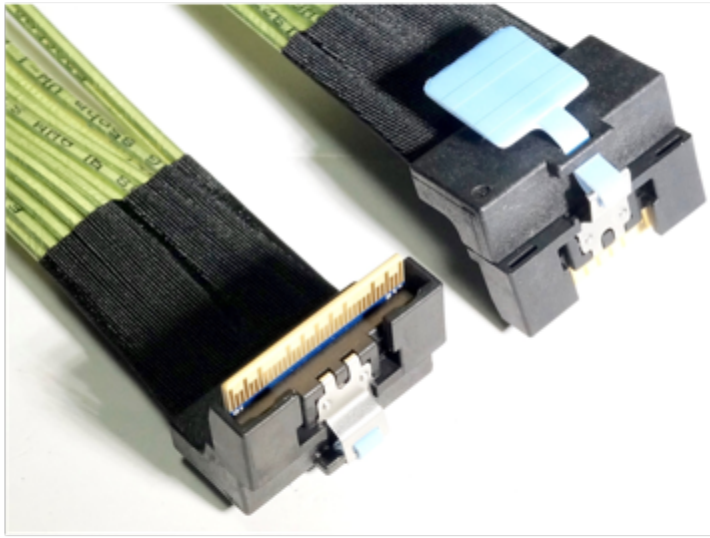
Form factors tend to be a little bit different compared to other pluggable IO solutions discussed today

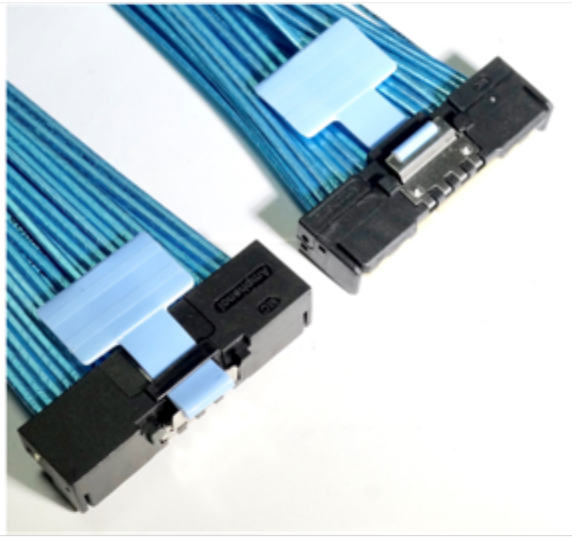
The same tolerance stack-up applies to these IO solutions as well

Interaction of the fixed and free sides of the interface significantly impact the capabilities of the interface





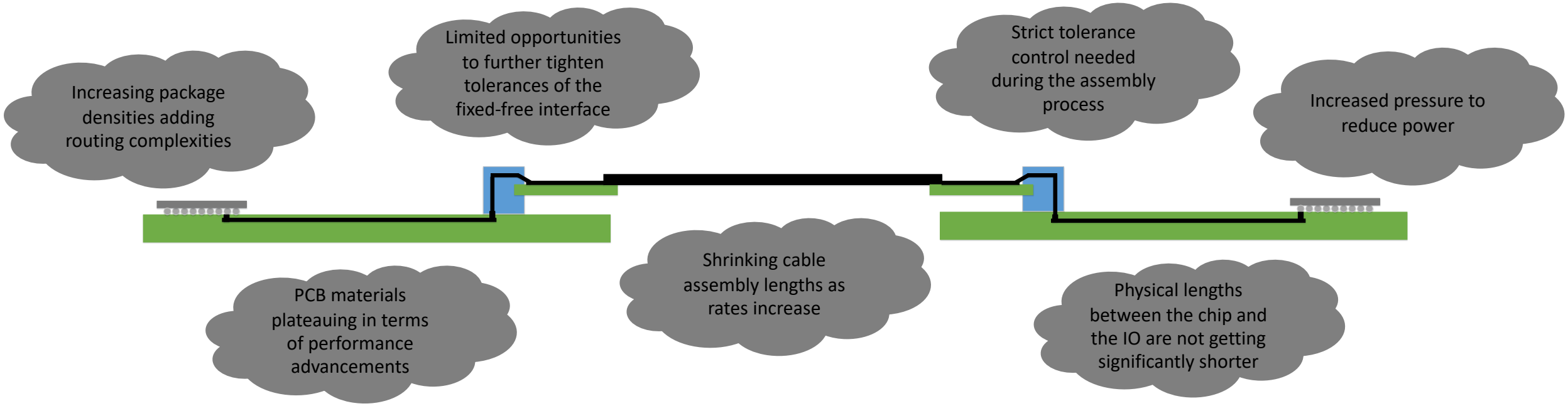




SFF-8654

SFF-TA-1016

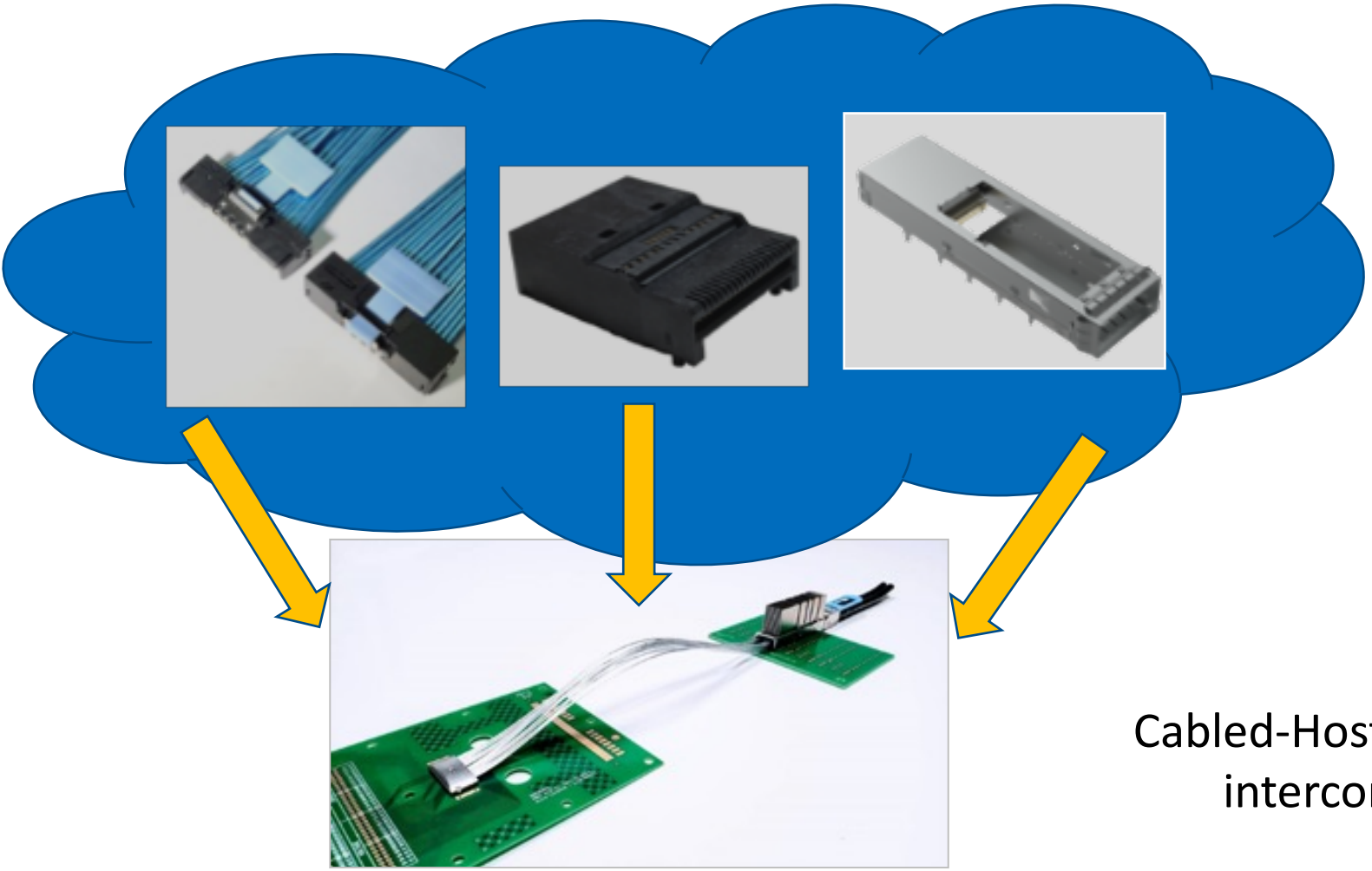
Challenges Ahead for Pluggable IO



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
- 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/
- Follow us on Twitter @SNIANSF

Thank You

