ESNOG-32
400G And Beyond

Manuel Mendez
Systems Engineering
manuel@

About Arista Networks

Founded 2004
20th Year Anniversary

IPO 2014
June 6th

S&P 500
Added in 2018

#1 High Speed Ethernet
10G-100G-400G

10,000
Customers

$124B Market Cap
As of Oct'24

CATNIX Catalonia Neutral Internet Exchange

Network Provider
Since 2016

Agenda

Transceivers 101

Switch Silicon & Optics Evolution

400G

800G and 1.6Tbps

LPO Optics

Disclaimer

The examples provided in the presentation are for illustration purposes ONLY, and are NOT indicative of ALL Arista's shipping products.

Any information contained in this presentation regarding third parties has been obtained from publicly available sources.

Transceivers 101

Refresher, Breakouts and Gearboxes

Refresher

*Serializer Deserializer
**Physical Layer device
***Multi-Source Agreement

- Transceivers convert a host-side (internal) signal to the line-side (external) signal

- **Host-side** interfaces are known as **SerDes***
  - Defined by IEEE standards
  - Designed for short range high speed signalling (i.e. chip to chip)
  - Connect the PHY** to the transceiver.
  - PHY may be inside the switching silicon or an external component (e.g. Base-T PHY)

- **Line-side** interfaces may be electrical or optical
  - Defined by IEEE standards or by industry groups (e.g. 10G-LR)

- Transceiver format and electrical interface design
  - **Physical** form factor, electrical interface defined by MSA***
  - **Internal** implementation is the vendor's secret sauce

To interoperate with hosts and link partners, transceivers must meet many common standards

Host Side
(Electrical)

Line Side
(Optical or Electrical)

Transceiver Format	Host Side Interface (aka SerDes)	Example Type	Line Side Interface
SFP+ (Small Form factor Pluggable +)	1 x 10Gbps	10G-SR	1 x 10G wavelength (on single MMF pair)
SFP+ (Small Form factor Pluggable +)	1 x 10Gbps	10G-T	1 x Cat5e or better
SFP28 (Small Form factor Pluggable 28Ghz)	1 x 25Gbps	25G-SR	1 x 25G wavelength (on single MMF pair)
SFP28 (Small Form factor Pluggable 28Ghz)	1 x 25Gbps	25G-LR	1 x 25G wavelength (on single SMF pair)
SFP-DD or DSFP (Small Form factor Pluggable + Double Density) (Double SFP is an alternative standard)	2 x 50Gbps	100G-DR	1 x 100G wavelength (on a single SMF pair)

Transceiver format defines the host-side electrical characteristics, line side varies by use case

Host Side
(Electrical)

Line Side
(Optical or Electrical)

Transceiver Format	Host Side Interface (aka SerDes)	Example Type	Line Side Interface
QSFP+ (Quad Small Form factor Pluggable +)	4 x 10Gbps	40G-SR4	4 x 10G wavelengths (on 4 parallel MMF pairs)
QSFP+ (Quad Small Form factor Pluggable +)	4 x 10Gbps	40G-LR4	4 x 10G wavelengths (on a single SMF pair)
QSFP28 (Quad Small Form factor Pluggable 28Ghz)	4 x 25Gbps	100G-SR4	4 x 25G wavelengths (on 4 parallel MMF pairs)
QSFP28 (Quad Small Form factor Pluggable 28Ghz)	4 x 25Gbps	100G-LR4	4 x 25G wavelengths (on a single SMF pair)
QSFP28 (Quad Small Form factor Pluggable 28Ghz)	4 x 25Gbps	100G-DR	1 x 100G wavelength (on a single SMF pair)

Transceiver format defines the host-side electrical characteristics, line side varies by use case

Host Side
(Electrical)

Line Side
(Optical or Electrical)

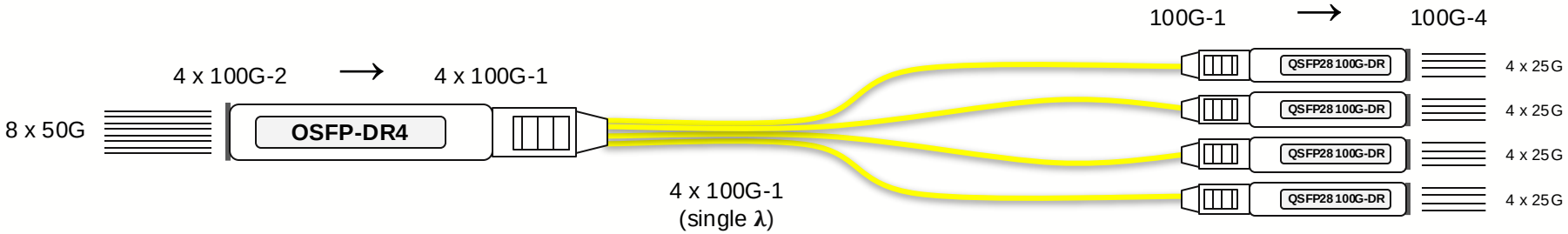
Transceiver Format	Host Side Interface (aka SerDes)	Example Type	Line Side Interface
QSFP56 (Quad Small Form factor Pluggable 56Ghz)	4 x 50Gbps	200G-LR4	4 x 50G wavelengths (on a single SMF pair)
QSFP-DD/QSFP56-DD (Quad Small Form factor Pluggable - Double Density)	8 x 25Gbps or 8 x 50Gbps	2 x 100G or 2 x 200G DAC	2 x (4 x 25G) electrical 2 x (4 x 50G) electrical
QSFP-DD/QSFP56-DD (Quad Small Form factor Pluggable - Double Density)	8 x 25Gbps or 8 x 50Gbps	400G-SR8	8 x 50G wavelengths (on 8 parallel MMF pairs)
QSFP-DD/QSFP56-DD (Quad Small Form factor Pluggable - Double Density)	8 x 50Gbps	400G-DR4	4 x 100G wavelengths (on 4 parallel SMF pairs)
OSFP (Octal Small Form factor Pluggable)	8 x 50Gbps	400G-LR4	4 x 100G wavelengths (on a single SMF pair)
OSFP (Octal Small Form factor Pluggable)	8 x 50Gbps	400G-ZR	1 x 400G wavelength (on a single SMF pair)

Physical Link Naming Convention

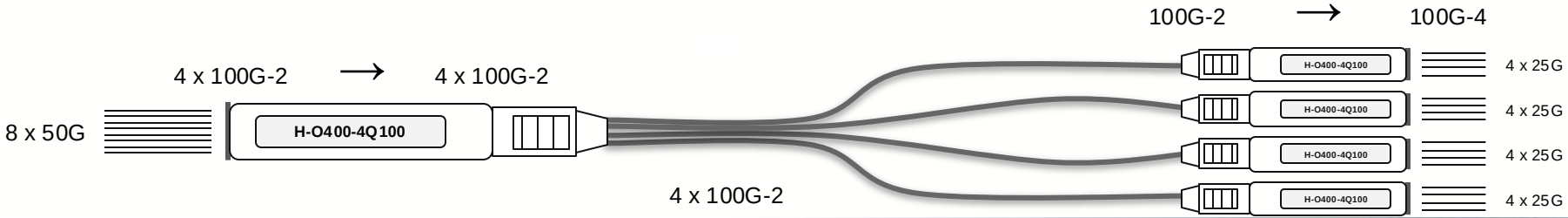
Nomenclature	Link Speed	Underlying Carrier(s)	Example Deployment
50G-2	50G	2 x 25G NRZ	50G on 2 lanes of a QSFP28
50G-1	50G	1 x 50G PAM4	50G on 1 lane of an SFP-DD / QSFP56 / Q-DD / OSFP
100G-4	100G	4 x 25G NRZ	100G on 4 lanes of a QSFP
100G-2	100G	2 x 50G PAM4	100G on 2 lanes of an SFP-DD / QSFP56 / DD / OSFP
100G-1	100G	1 x 100G PAM4	100G on 1 lane of a QSFP112 / DD112 / OSFP800
200G-4	200G	4 x 50G PAM4	200G on 4 lanes of a QSFP56 / DD / OSFP
200G-2	200G	2 x 100G PAM4	200G on 2 lanes of a QSFP112 / DD112 / OSFP800
400G-8	400G	8 x 50G PAM4	400G on 8 lanes of a QSFP-DD / OSFP
400G-4	400G	4 x 100G PAM4	400G on 4 lanes of a QSFP112 / DD112 / OSFP800

Breakouts

Optical Breakout



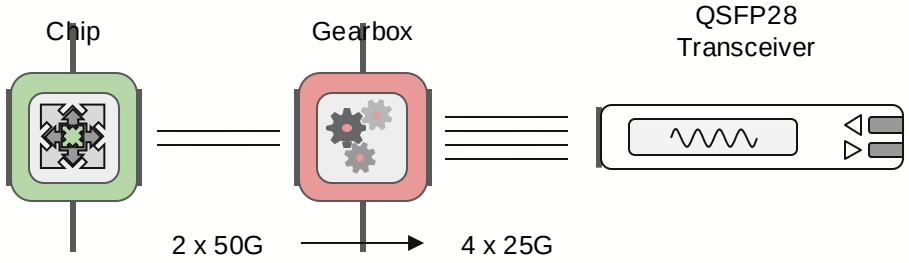
Active Copper Breakout



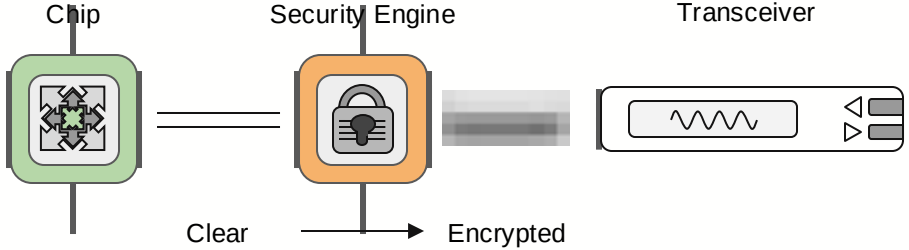
Gearboxes

Provides rate conversion and may add other features:

Rate Conversion:



Encryption:

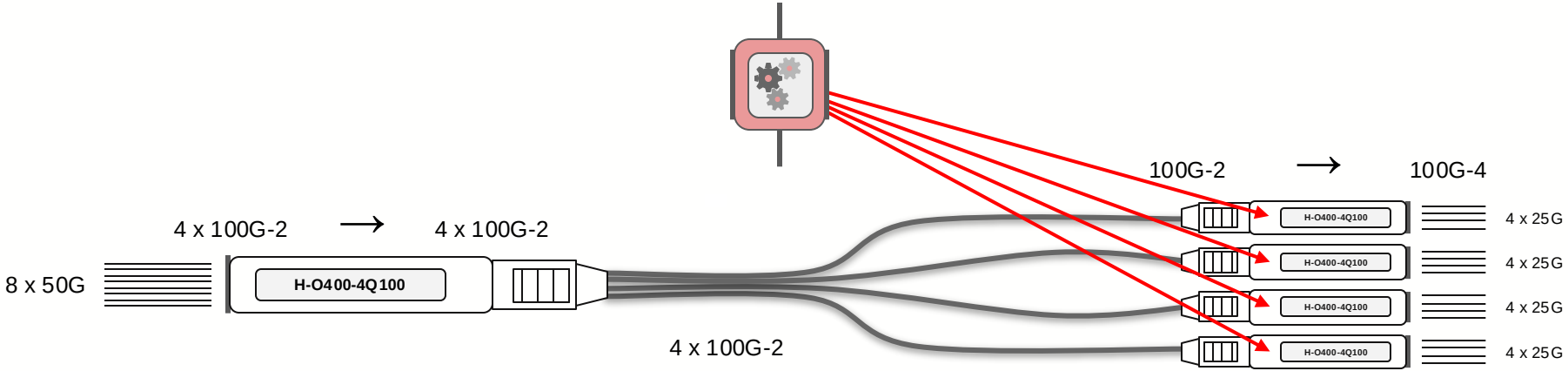


Gearbox converts 2 x 50G to 4 x 25G to support QSFP28

ARISTA

Gearboxes

Active Copper Breakout

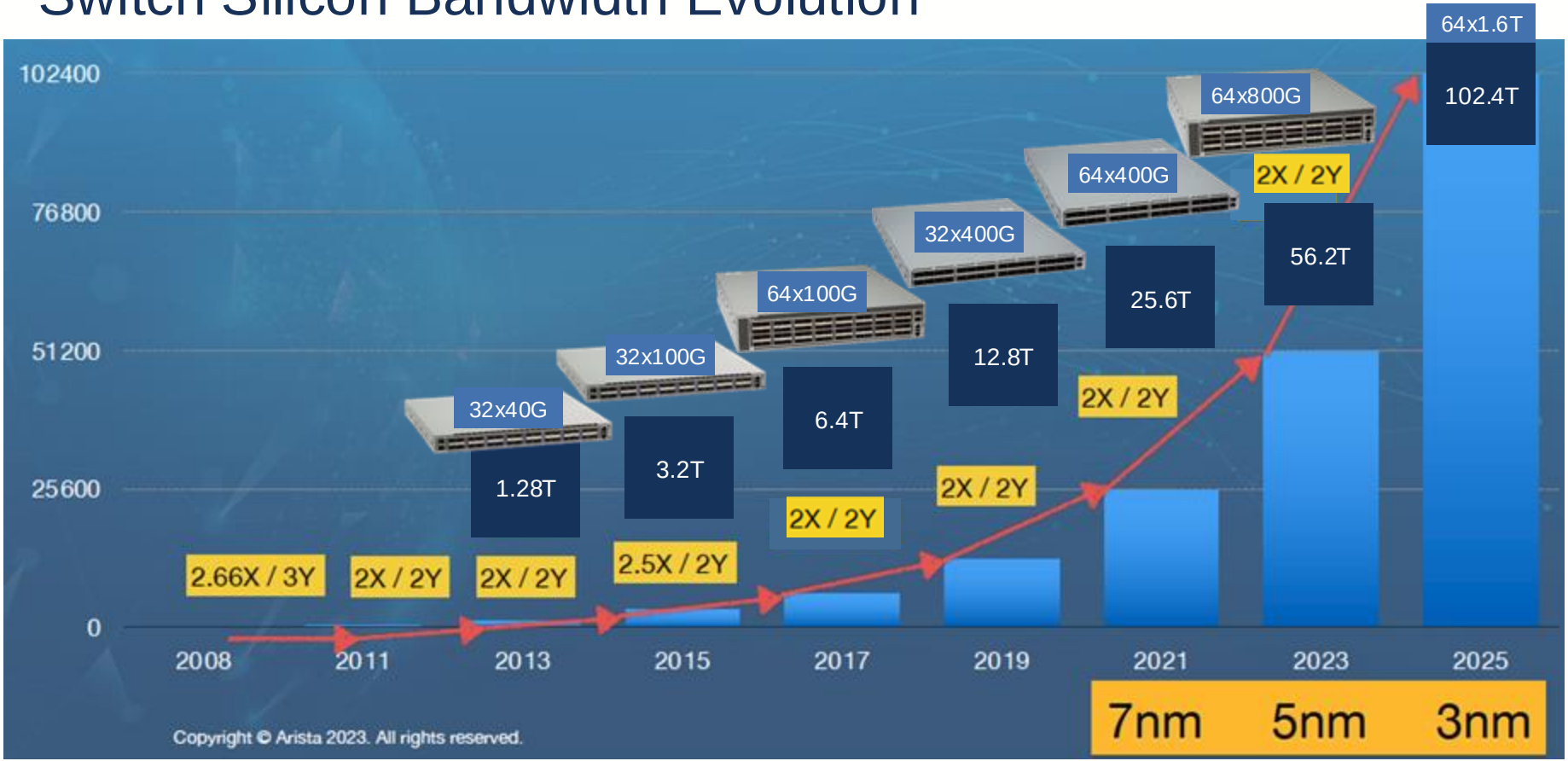


ARISTA

Founding Member

Switch Silicon & Optics Evolution

Switch Silicon Bandwidth Evolution

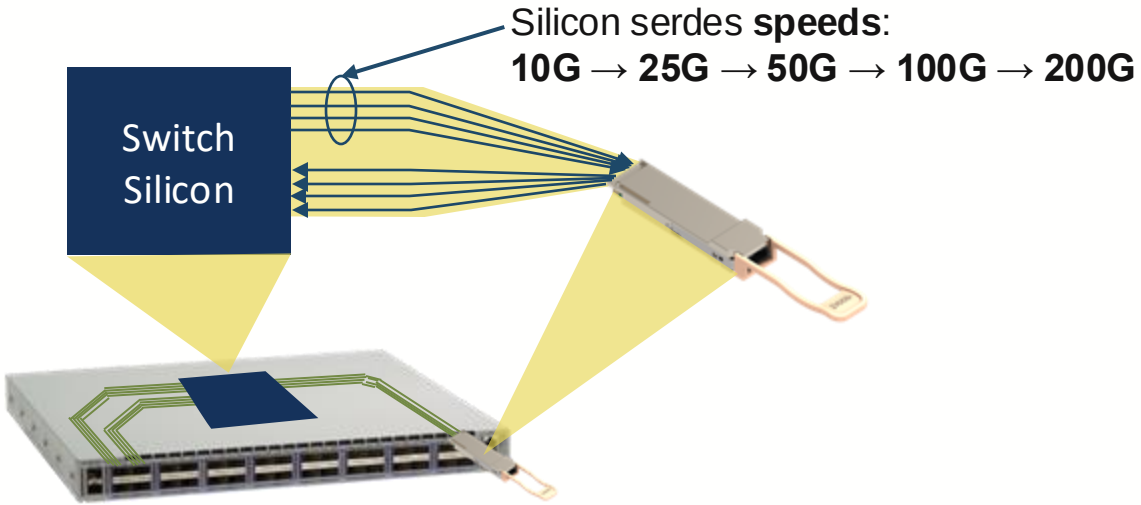


ARISTA

What Limits Switch Silicon Bandwidth?

Switch silicon bandwidth is constrained by two fundamental factors:

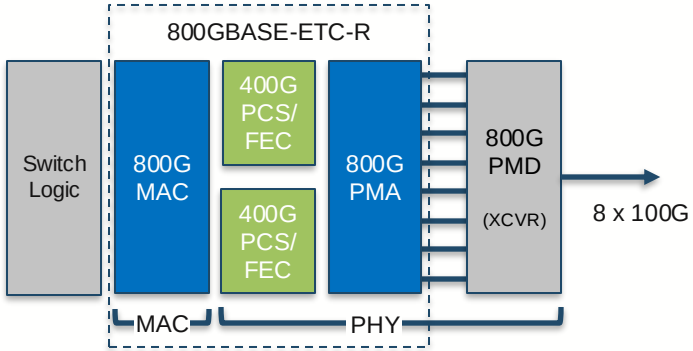
1. Number of I/O on a silicon die
2. How fast the I/O (SerDes) can be driven



The easiest way to go faster is (for SerDes speeds) to go faster

State of 800/1.6 TGbE

- 800 GbE Requirements:
  - 100G SerDes
  - 800G Transceivers
  - 800G MAC Layer



Ethernet Technology Consortium

Founding Member

Standard	Status	Description	Electrical Interfaces	Silicon
IEEE 802.3ck	Sept 2022	100, 200 & 400 GbE using 100G lanes	100G-1, 200G-2, 400G-4	Shipping
Ethernet Technology Consortium 800GBASE-ETC-R	October 2020	800G using 100G lanes	800G-8	Shipping
IEEE 802.3df	March 2024 Approved	800G using 100G lanes	800G-8	Future
IEEE P802.3dj	2H 2026	200, 400 Gb/s, 800 Gb/s, and 1.6 Tb/s using 200 Gbit/s lanes 1.6 Tbps MAC	200G-1, 400G-2, 800G-4, 1600G-8	

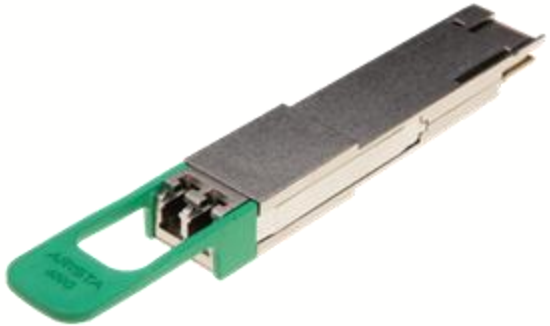
ARISTA

Optics Form factors

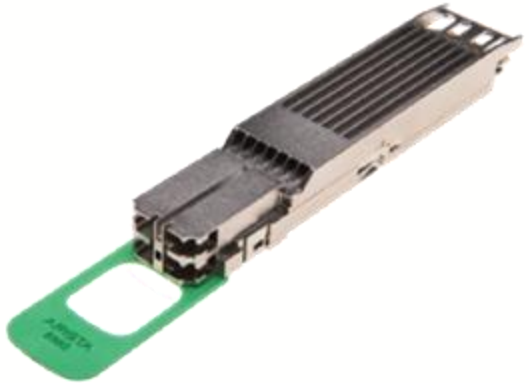
Founding Member

Transition from 400G QSFP-DD to 800G OSFP. Why?

- Better thermal performance (integrated heatsink, larger surface area)
- Robust signal integrity (single row connector)
- Better fit for dual-LC connectors for 800G
- Industry aligning with largest volume consumers



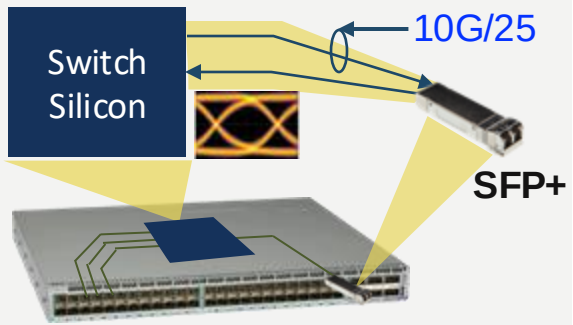
400G QSFP-DD



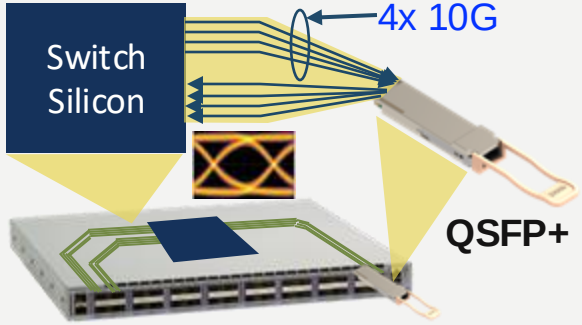
800G OSFP

SerDes Speed Transition

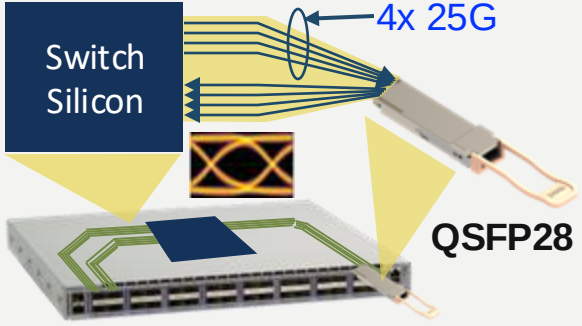
10G / 25G Port: 10G/25G NRZ



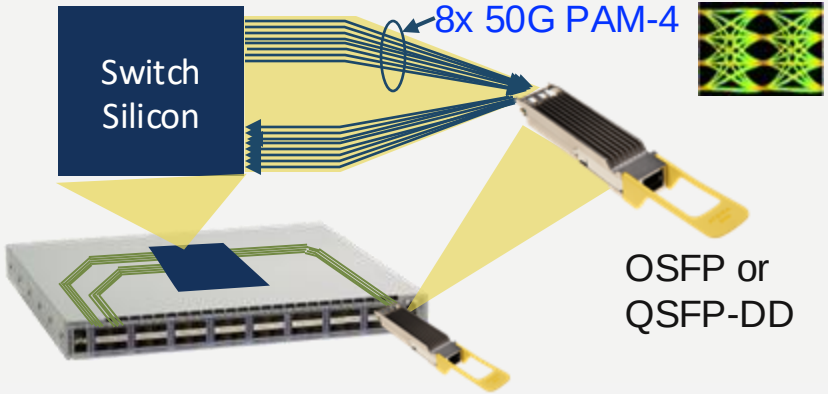
40G / Port: 4x 10G



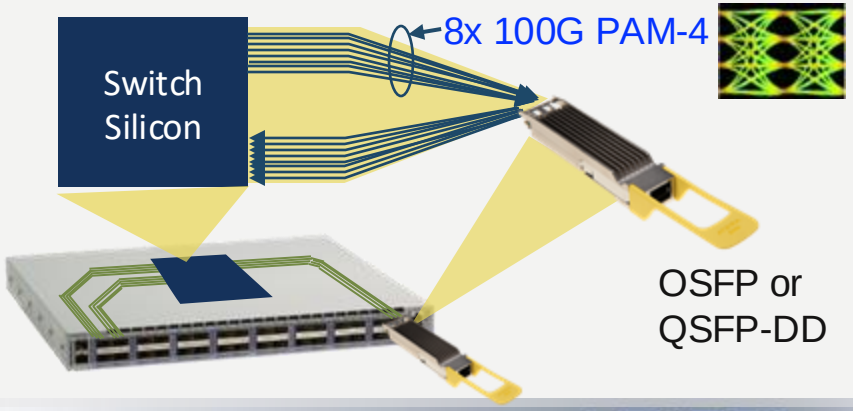
100G / Port: 4x 25G



400G / Port: 8x 50G PAM-4



800G / Port: 8x 100G PAM-4



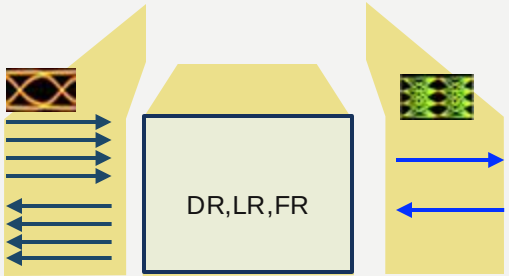
ARISTA

Serdes Speeds & Optical Speeds for 100G, 400G & 800G

100G QSFP Optics

Elec interface:
4 x 25G NRZ

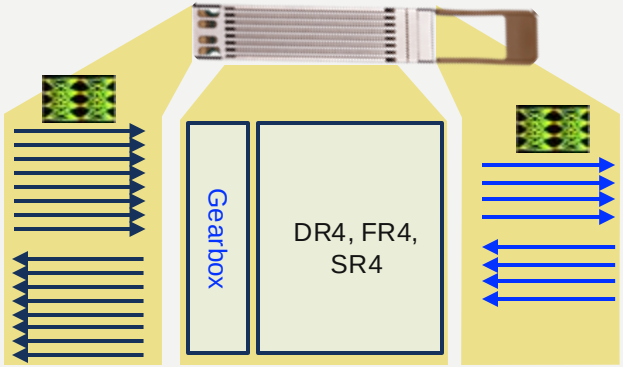
Optical waves:
4 x 25G NRZ



Elec interface:
4 x 25G NRZ

Optical wave:
1 x 100G

400G Optics: 4x 100G-λ or 8x 50G-λ



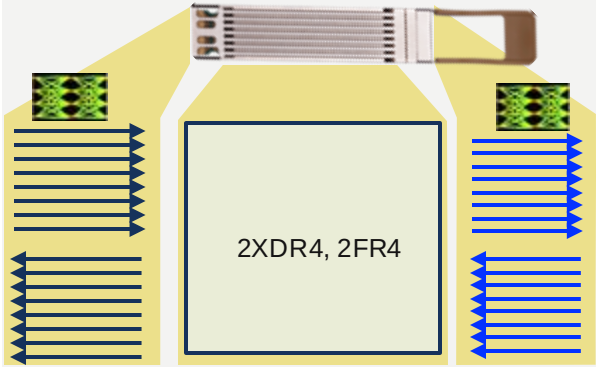
Elec Interface:
8 x 50G

Optical waves:
4x 100G PAM-4

Elec Interface:
8 x 50G

Optical waves:
8x 50G PAM-4

800G Optics: 8x 100G-λ



Elec Interface:
8 x 100G

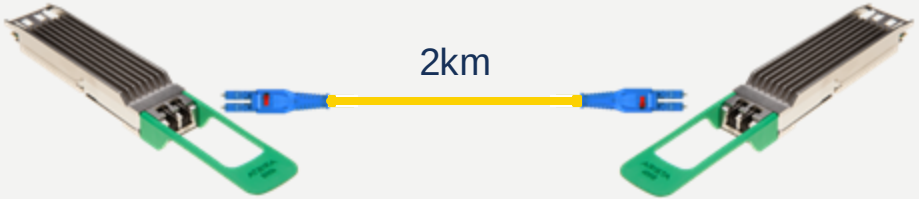
Optical waves:
8x 100G PAM-4

ARISTA

400G

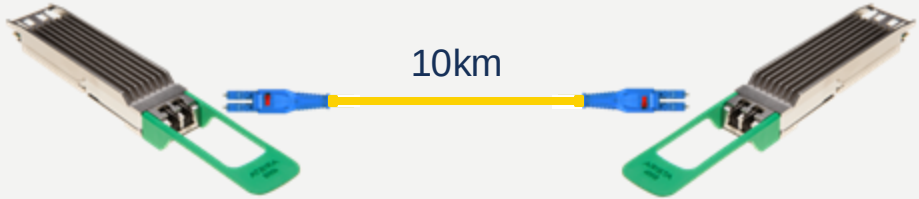
SMF 400G Optics

400G-FR4



400GE over duplex SMF

400G-LR4



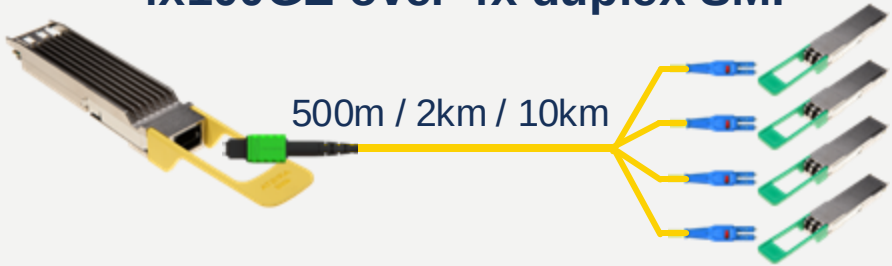
400GE over duplex SMF

400G-DR4 / XDR4 / PLR4

400GE over parallel SMF

OR

4x100GE over 4x duplex SMF

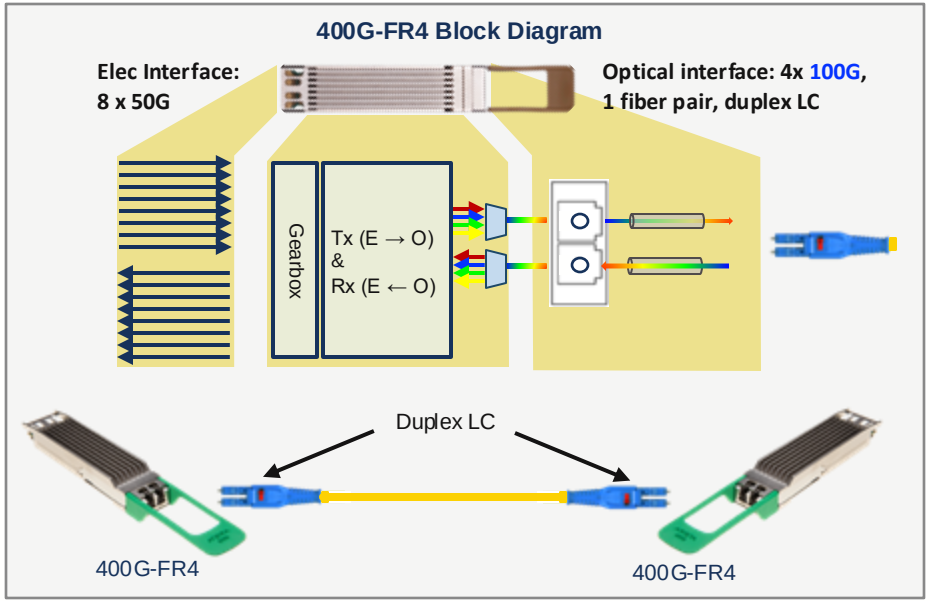


4x 100G-DR / FR / LR

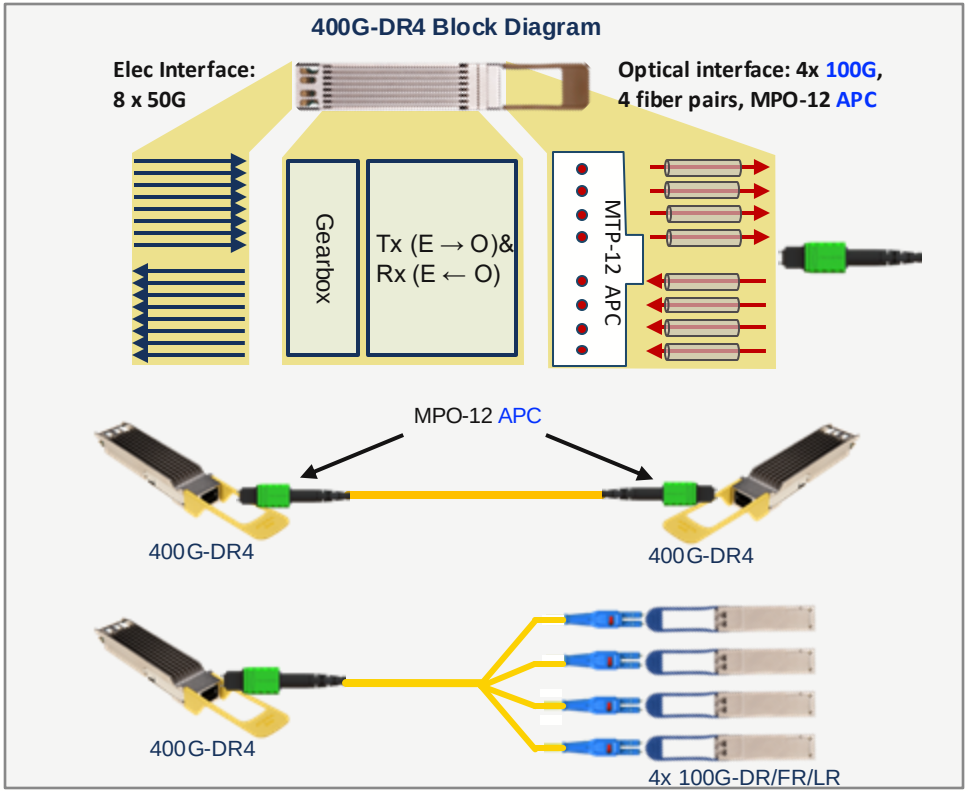
ARISTA

SMF 400G Optics Building Blocks

400G-FR4 / LR4: 2km / 10km over duplex SMF

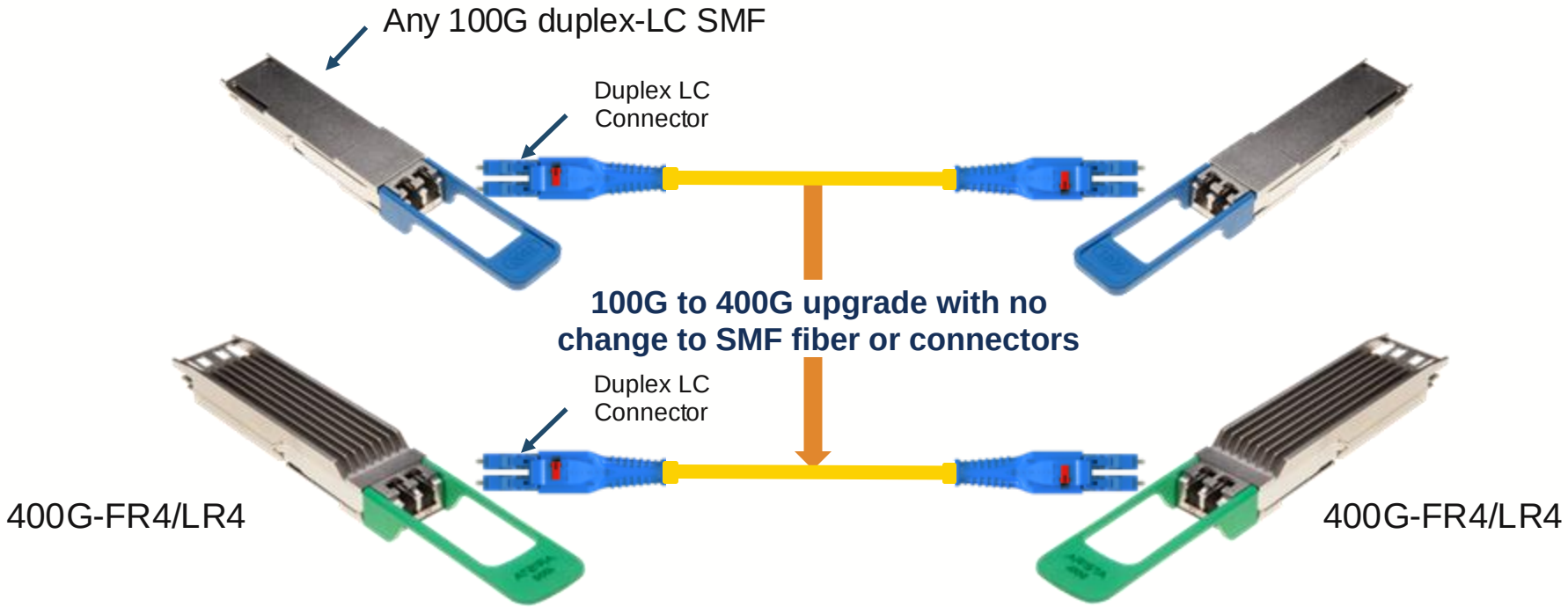


400G-DR4 / XDR4 / PLR4: 500m / 2km / 10km, over Parallel SMF



ARISTA

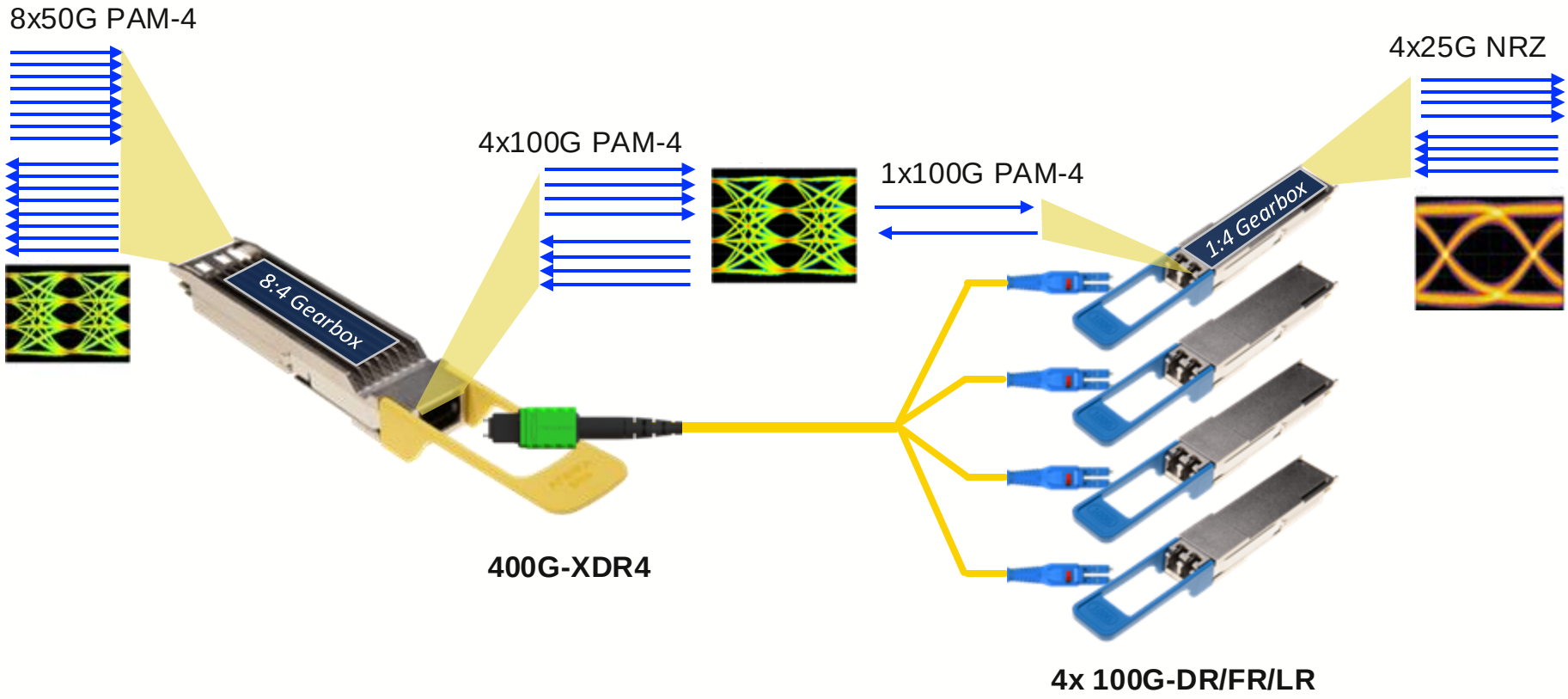
Upgrade duplex SMF from 100G to 400G



400G-FR4/LR4 uses the same fiber & connectors as 100G duplex-LC SMF Optics

ARISTA

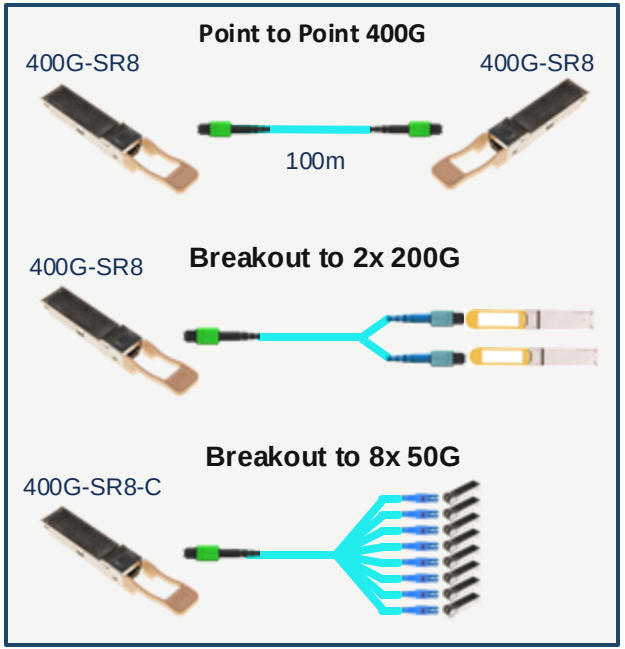
400G-DR4/XDR4/PLR4 Optical Breakout to 4x100G



MMF 400G Optics

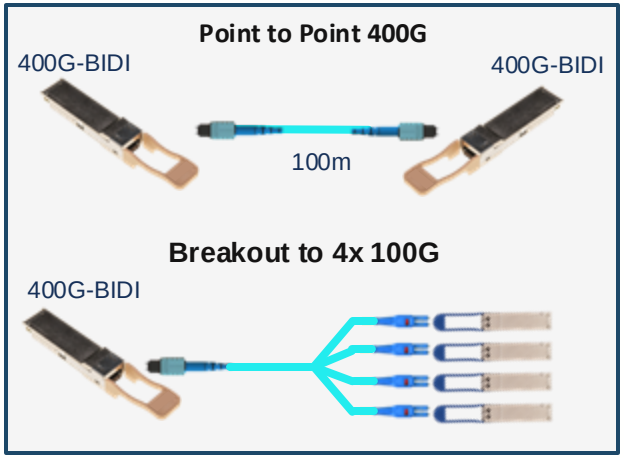
400G-SR8

- Reach: 100m
- Optical Connector: MPO-16 APC
- Breakouts: 2x 200G, 8x 50G



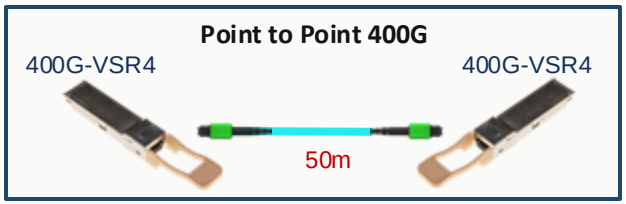
400G-BIDI

- Reach: 100m
- Optical Connector: MPO-12 UPC
- Breakouts: 4x 100G-BIDI

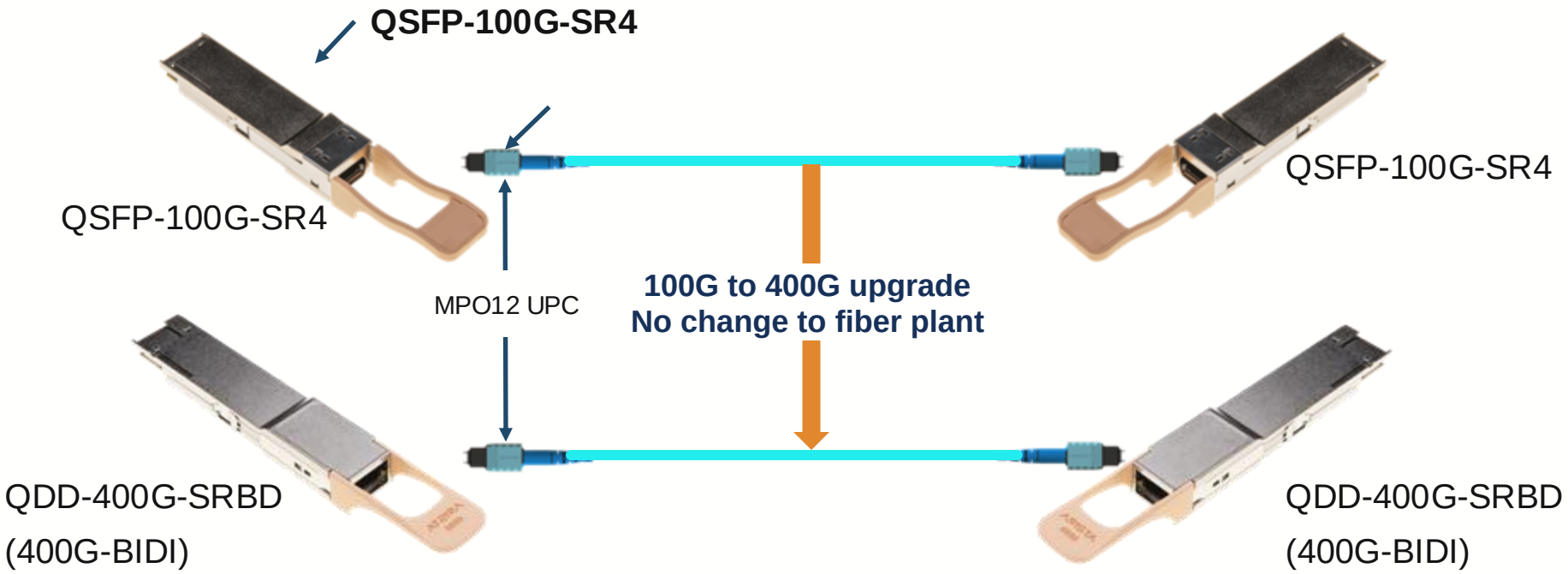


400G-VSR4

- Reach: 50m
- Optical Connector: MPO-12 APC
- Breakouts: None at present



Upgrade MMF Networks from 100G to 400G with **400G-BIDI**

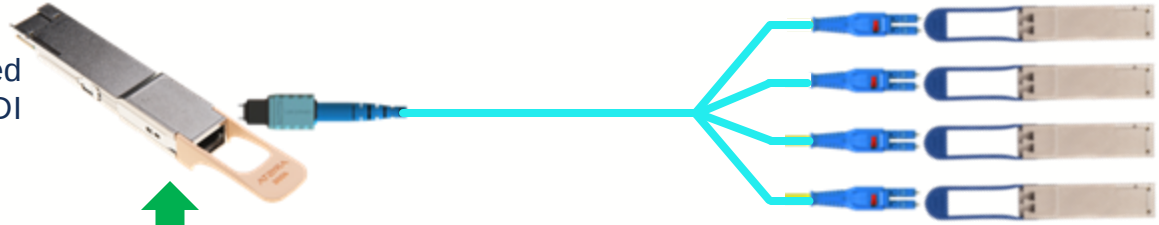


400G-BIDI is the **ONLY** 400G Transceiver that uses EXACTLY the same fiber & connectors as 100G-SR4

ARISTA

Flexible Breakout to 4x 100G with 400G-BIDI

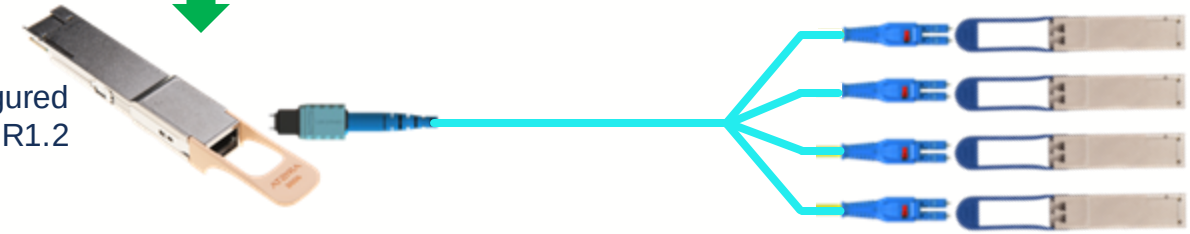
400G-BIDI, configured as 4x 100G-BIDI



4x 100G-BIDI (Installed base)

SW Configurable for 2 breakout modes
4x 100G-BIDI (Installed base) OR
4x 100G-SR1.2

400G-BIDI, configured as 4x 100G-SR1.2



4x 100G-SR1.2 (New “100G-BIDI”)

400G-BIDI is the **ONLY** 400G MMF Transceiver that enables simple 4x 100G Breakout

ARISTA

400G-VSR4 for 400G Switch-to-NIC Connectivity

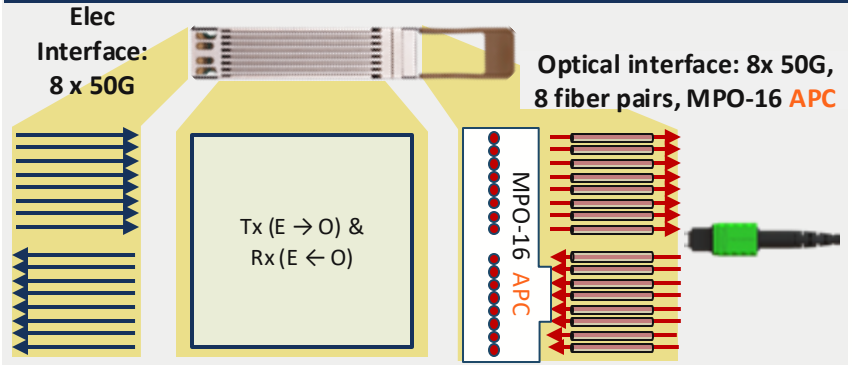
Key Application: 400G-8 switch ports to 400G-4 QSFP112 NIC

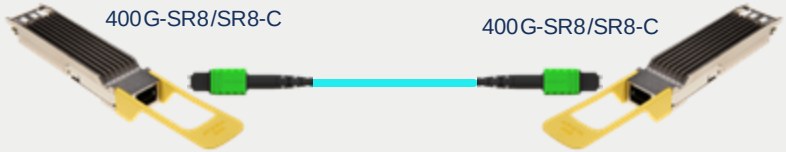
400G-VSR4 is the **ONLY** 400G MMF Transceiver with
4x 100G Optical Channels for interop with 400G-SR4/VR4 QSFP112

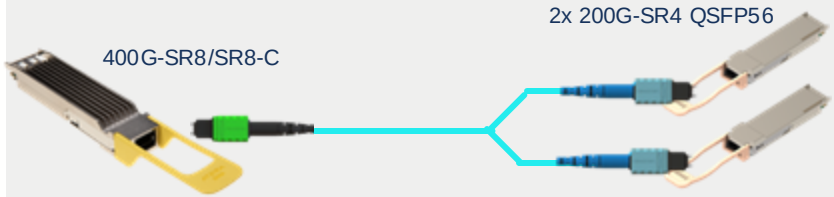
ARISTA

MMF 400G Optics Block Diagram

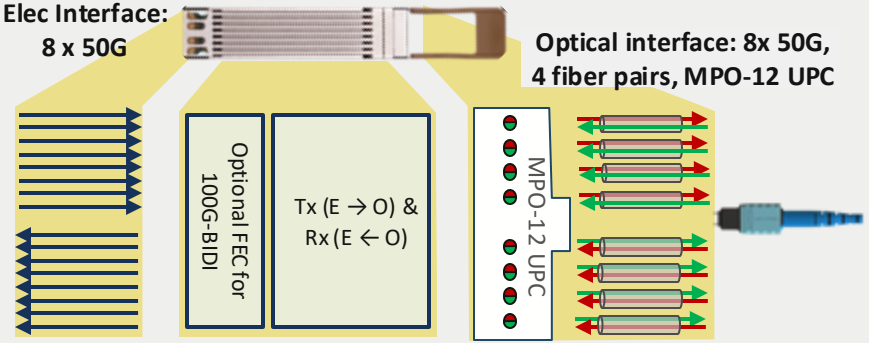
400G-SR8, 100m over parallel MMF

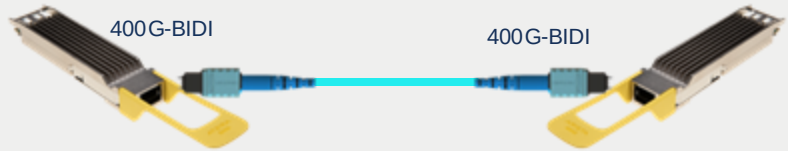


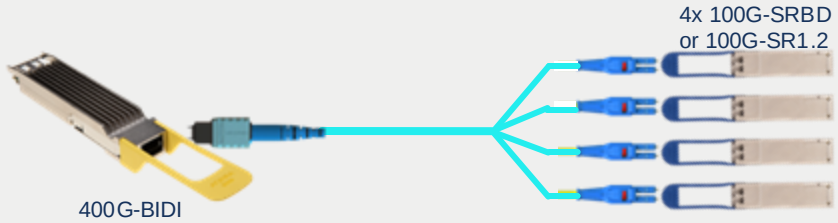




400G-BIDI, 100m over parallel MMF (AKA SR4.2)

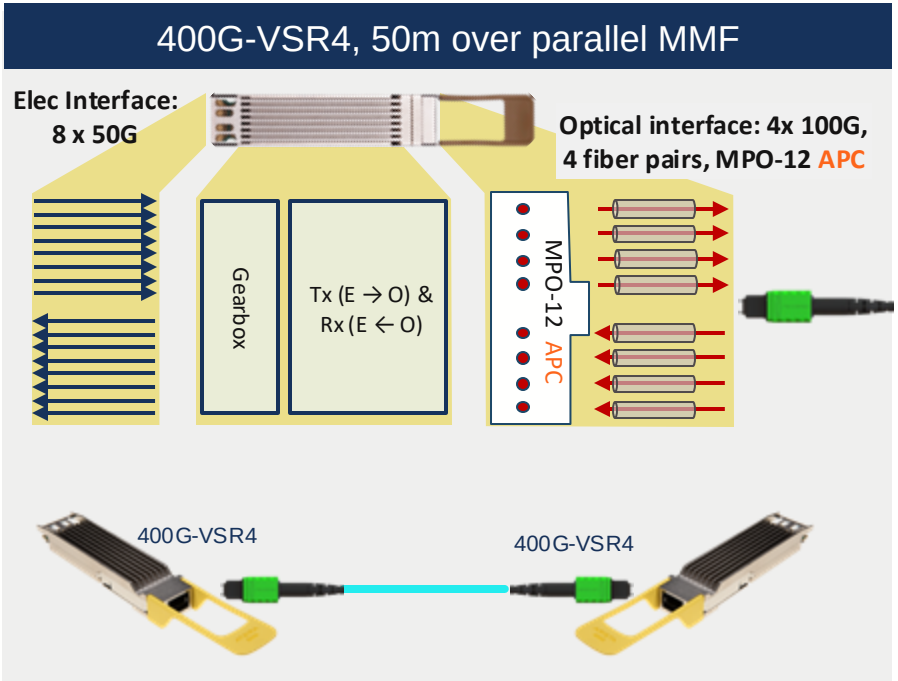






ARISTA

MMF 400G Optics Block Diagram



Note:

IEEE802.3 has standardized “400G-SR4” for 100m reach, and “400G-VR4” for 50m reach.

The Arista 400G-VSR4 will interop with both 400G-SR4 **and** 400G-VR4 optics over 50m

There are 50m reach optics on the market labelled as “400G-SR4”

400G AOCs and Passive Copper Cables (DACs)

400G to 400G Active Optical Cables (AOCs)

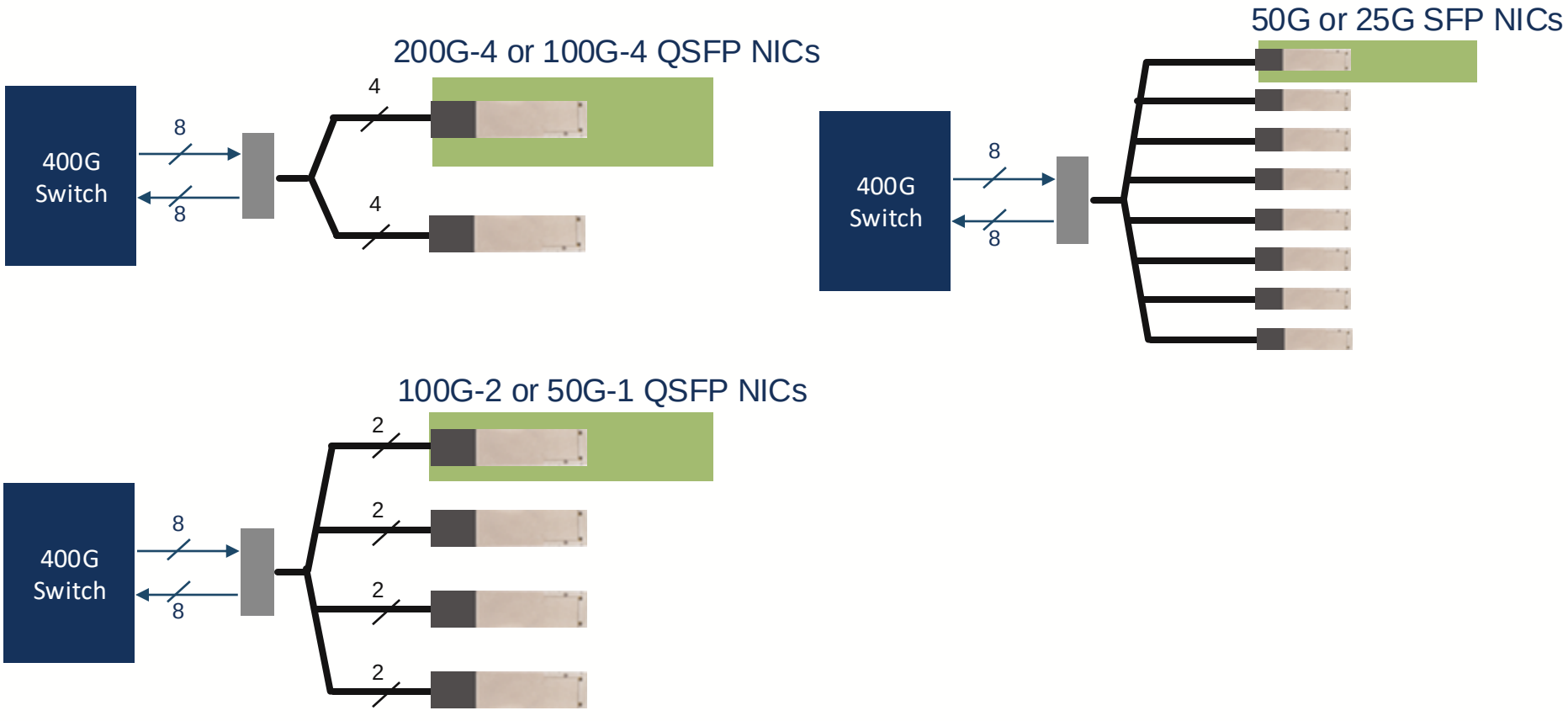
400G to 400G DACs

Breakout to 2x200G QSFPs

Breakout to 4x100G-2 QSFPs

Breakout to 8x50G SFPs

400G Copper Breakout Cables for Server Connections



400G AOCs for NIC and GPU Connectivity

400G QSFP-DD to 400G QSFP112 AOC

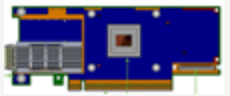
Elec Interface: 8 x 50G

8x 50G to 4x
100G gearbox

400G OSFP
or QSFP-DD

Elec Interface: 4 x 100G

400G QSFP112



400G QSFP112 NICS

400G QSFP-DD to 2x 200G QSFP56 Breakout AOC

Elec Interface: 8 x 50G

400G OSFP
or QSFP-DD

Elec Interface: 4 x 50G

2x 200G QSFP56

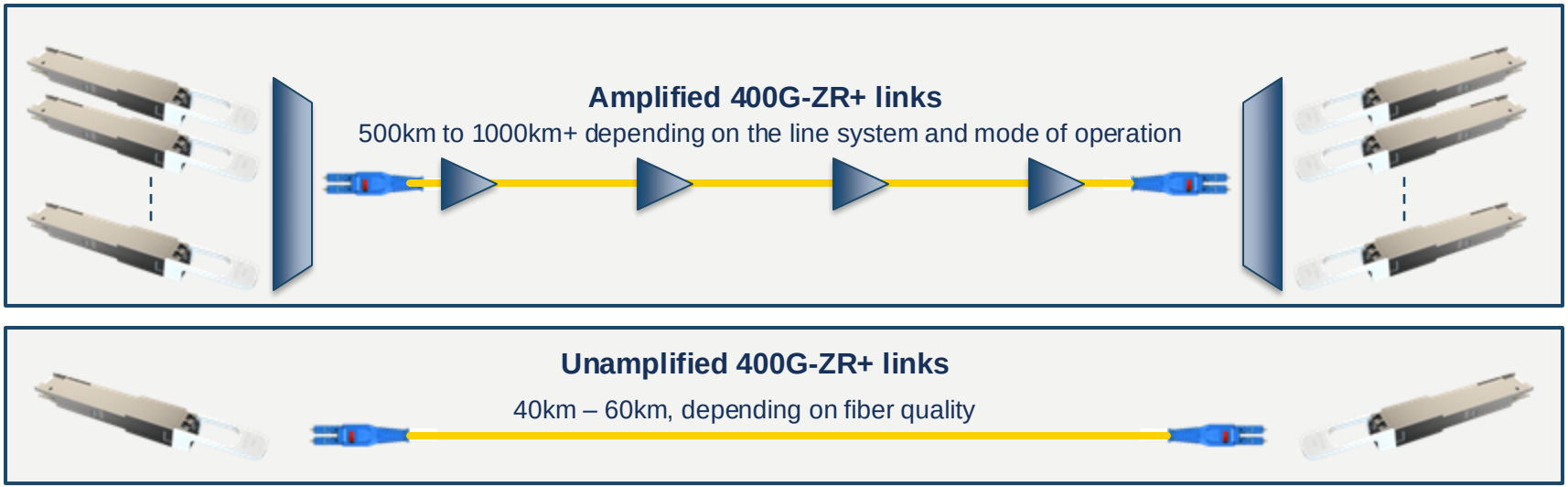
200G QSFP56 NICS

400G-ZR+ Applications

OpenZR+ MULTI-SOURCE AGREEMENT

Promoter

1. Amplified, using a **3rd party line system**: 500km – 1000km+ depending on the mode of operation and line system design
2. Unamplified: 40km – 60km reach, depending on fiber quality



ARISTA

800G

400G to 800G Optics Transition

400G Optics

8 x 50G PAM-4

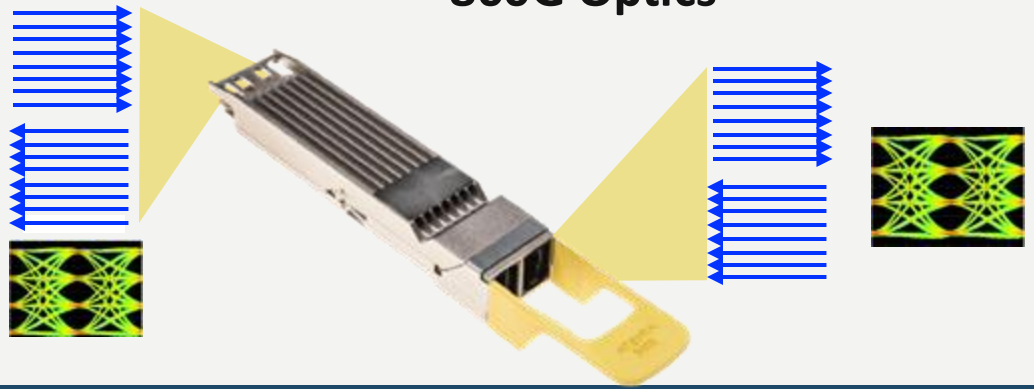
4x 100G PAM-4

Elec serdes
50G to 100G

Number of optical channels: 4x to 8x
Optical modulation is identical (100G)

800G Optics

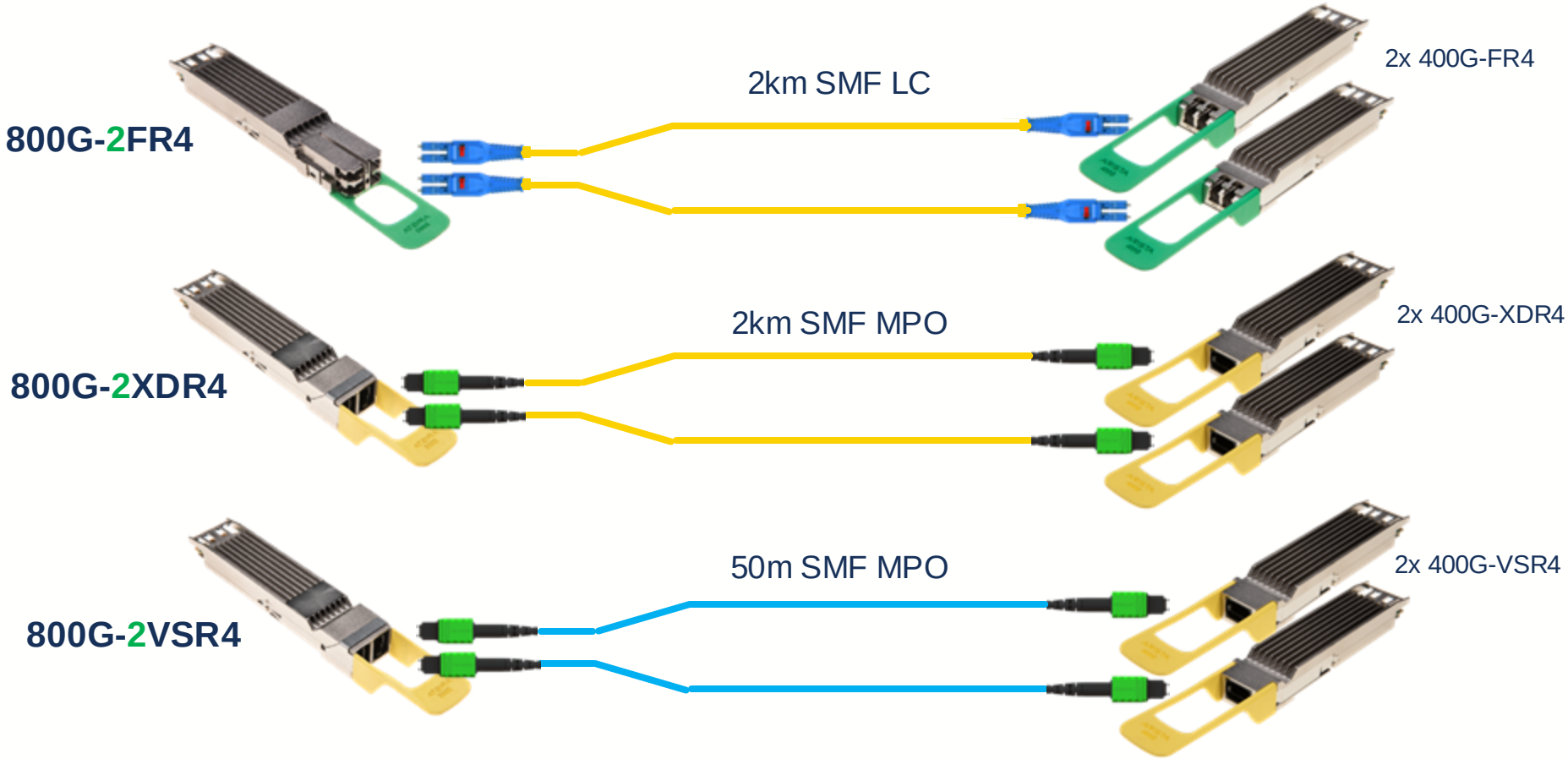
8 x 100G PAM-4



8x 100G PAM-4

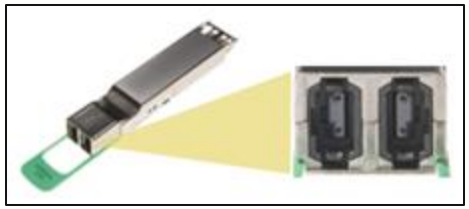
ARISTA

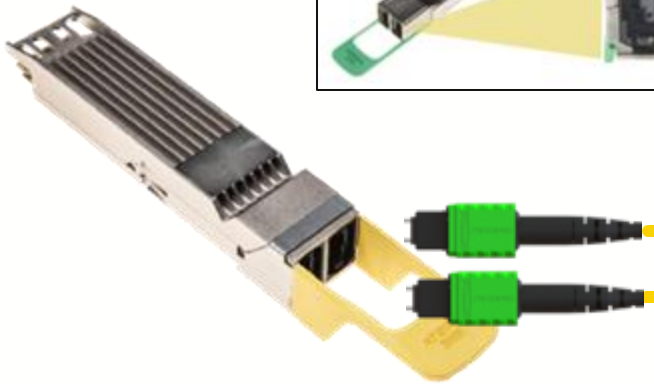
800G Optics = Double Density 400G Optics



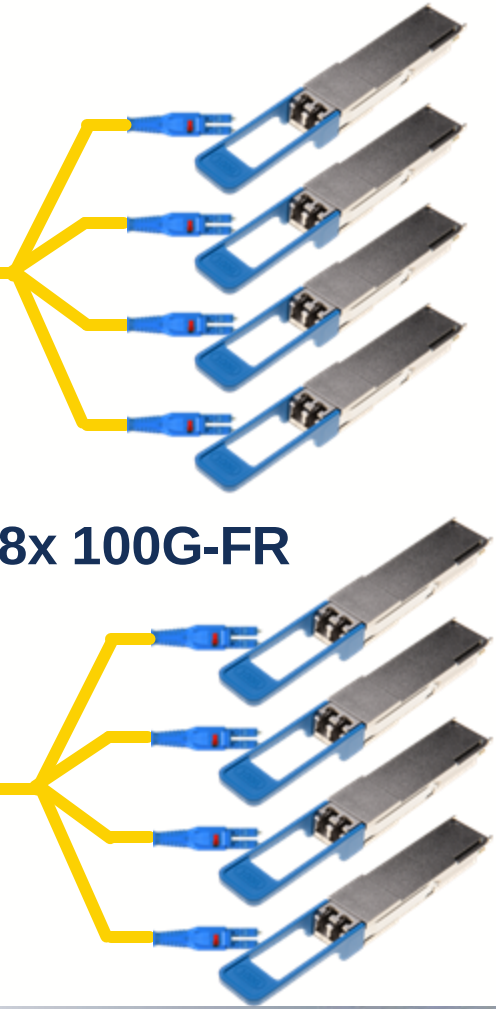
ARISTA

800G Optics: Breakout to 8x100GE





800G-2XDR4
(configured as 8x100GE)

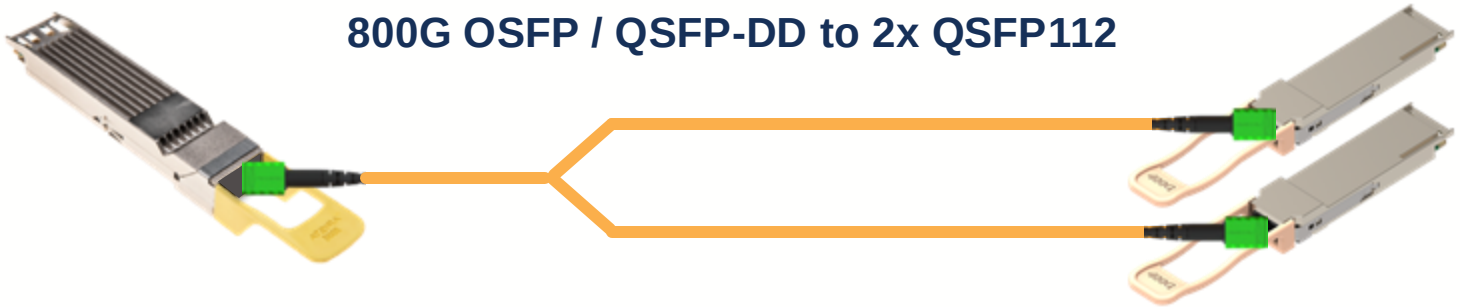


8x 100G-FR

800G AOCs

Straight Through AOC:
800G OSFP / QSFP-DD to 800G OSFP / QSFP-DD

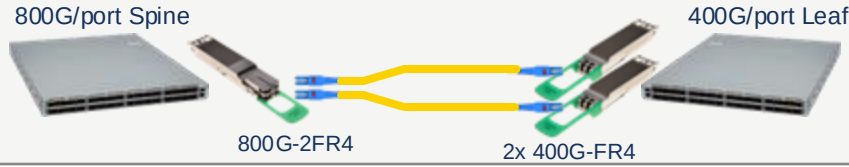
Breakout AOC:
800G OSFP / QSFP-DD to 2x QSFP112

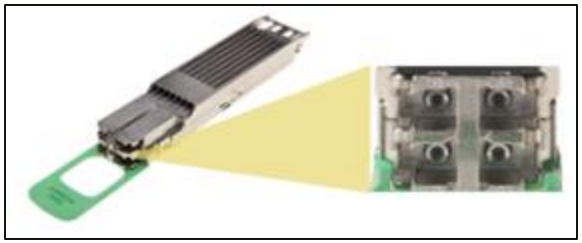


SMF 800G Optics Applications Examples

800G-2FR4

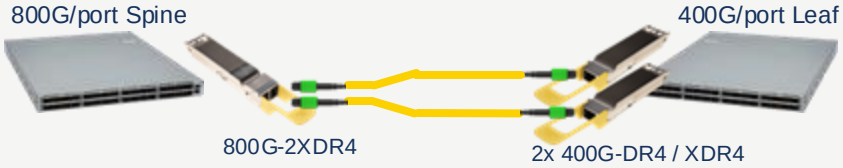
800G Spine (8x 100G/lane) to 400G Leaf (8x 50G/lane)



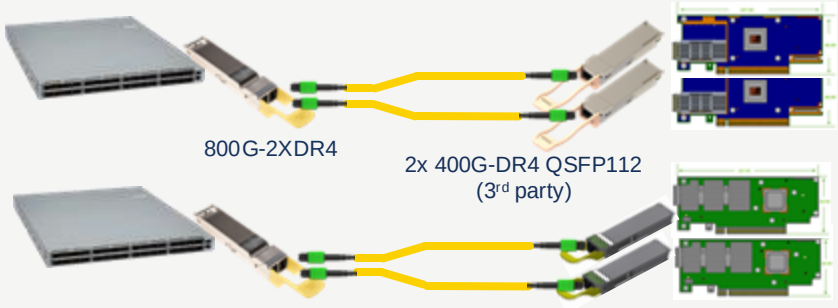


800G-2XDR4 (for SMF)

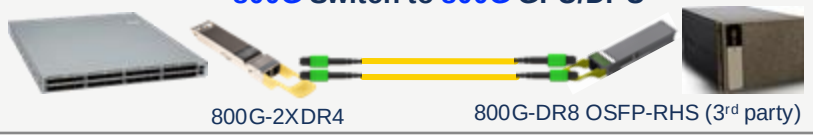
800G Spine (8x 100G/lane) to 400G (8x 50G/lane) Leaf



800G Switch (8x 100G/lane) to 2x 400G NIC (4x 100G/lane)



800G Switch to 800G GPU/DPU

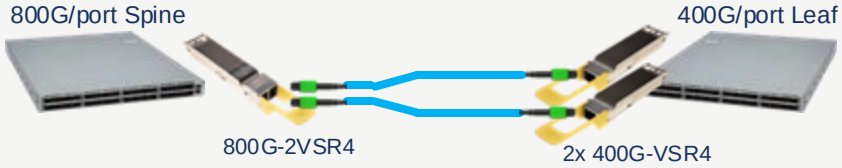


ARISTA

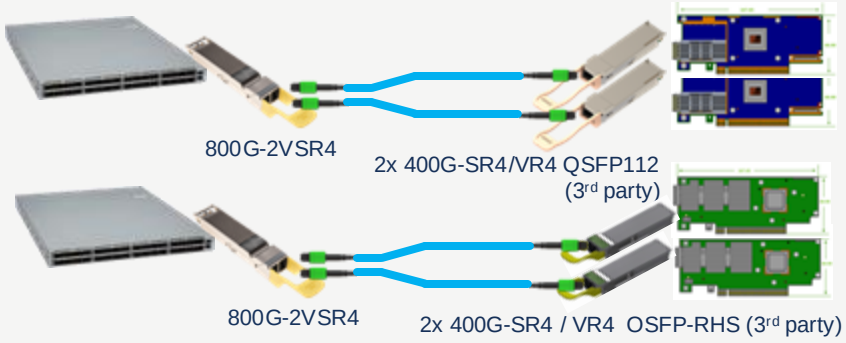
MMF 800G Optics Applications Examples

800G-2VSR4 (for MMF)

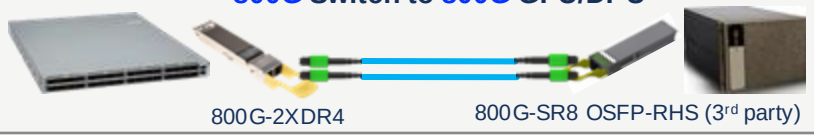
800G Spine (8x 100G/lane) to 400G (8x 50G/lane) Leaf



800G Switch (8x 100G/lane) to 2x 400G NIC (4x 100G/lane)



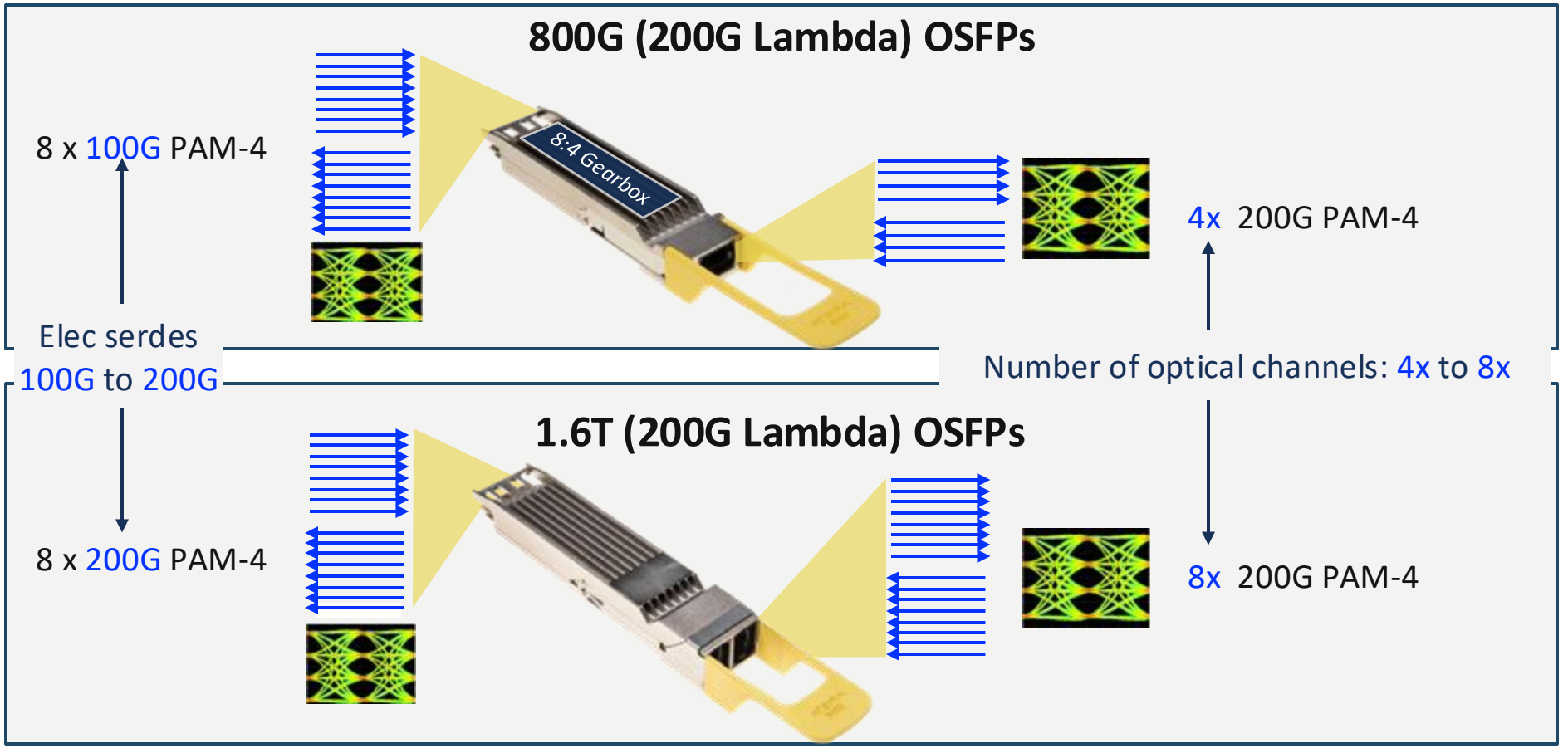
800G Switch to 800G GPU/DPU



ARISTA

Beyond 800G

800G to 1.6T Optics Transition

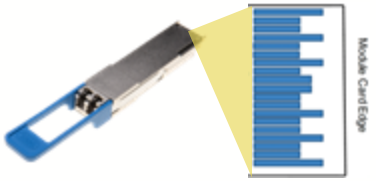


ARISTA

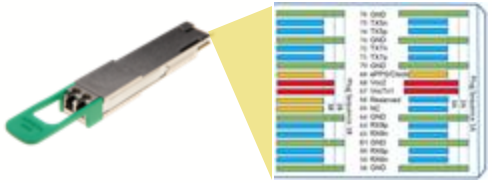
Optical Form-Factor Evolution

Founding Member

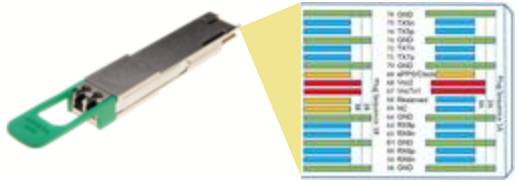
QSFP
Single row PCB connector
4x 25G Elec lanes



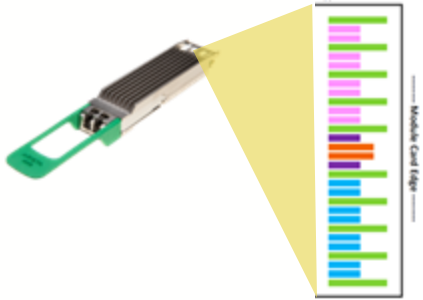
QSFP-DD: 400G
Dual row PCB connector
8x 50G Elec lanes



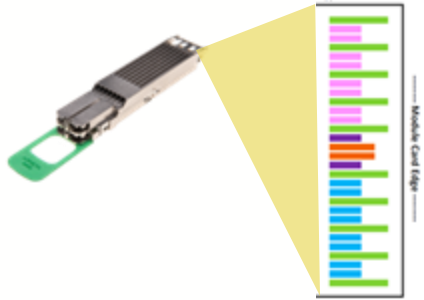
QSFP-DD: 800G
Dual row PCB connector
8x 100G Elec lanes



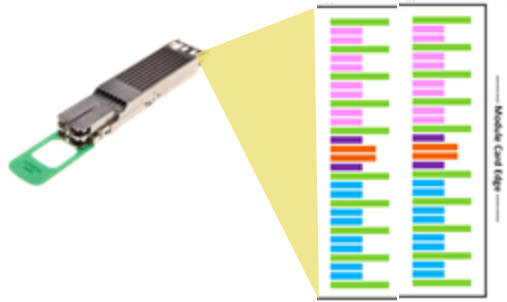
OSFP: 400G
Single row PCB connector
8x 50G Elec lanes



OSFP: 800G & 1.6T
Single row PCB connector
8x 100G/200G Elec lanes



OSFP-XD – 3.2T
Dual row PCB connector
16x 200G Elec lanes



ARISTA

ARISTA

LPO-MSA

Founding Member

LPO Optics

Linear-Drive Pluggable Optics (LPO) MSA

LPO-MSA

Founding Member

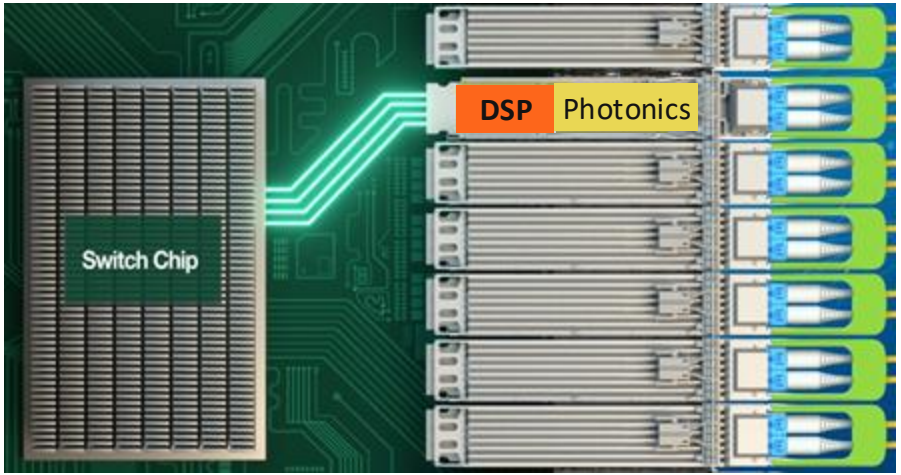
- Develops electrical and optical interoperability specifications for a diversity of high-density networking equipment and pluggable optical modules based on LPO technology
- Includes switches: NICs, and endpoints that include native Ethernet connectivity (such as GPUs)
- Ideal for high-speed, high-volume applications such as AI and high-performance computing

Linear-Drive Pluggable Optics (LPO)

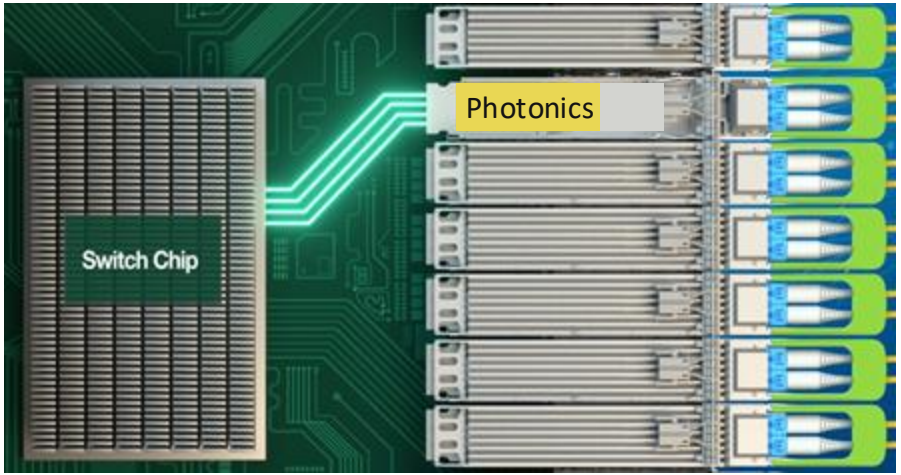
LPO-MSA

Founding Member

Traditional pluggable optical modules



LPO Modules



- LPO means no DSP in the module
- Switch Silicon SerDes use advanced DSP technology
  - Requires excellent system design and careful serdes tuning
  - Switch Silicon already shipping

Why LPO?

LPO-MSA

Founding Member

- Significantly lower power consumption
- Meets and exceeds link reliability requirements
- Lower latency (vs. DSP-based solutions)
- Lower module cost (vs. DSP-based solutions)
- 100 Gb/s per lane operation extending to 200 Gb/s per lane
- Enables a broad ecosystem of LPO module manufacturers and LPO-enabled networking equipment

Thank You

arista.com

References

https://www.arista.com/assets/data/pdf/Arista100G_TC_QA.pdf

https://www.arista.com/assets/data/pdf/Datasheets/Arista-400G_Optics_FAQ.pdf

https://www.arista.com/assets/data/pdf/Arista-800G-Transceivers-and-Cables-FAQ.pdf

https://osfpmsa.org

https://ethernettechnologyconsortium.org

https://www.lpo-msa.org

https://openzrplus.org/