The bridge to possible

400G, 800G, and Terabit Pluggable Optics:

What You Need to Know

Mark Nowell, Cisco Fellow
Errol Roberts, Distinguished Architect

BRKOPT-2699

cisco Live!

#CiscoLive

Cisco Webex App

https://ciscolive.ciscoevents.com/ciscolivebot/#BRKOPT-2699

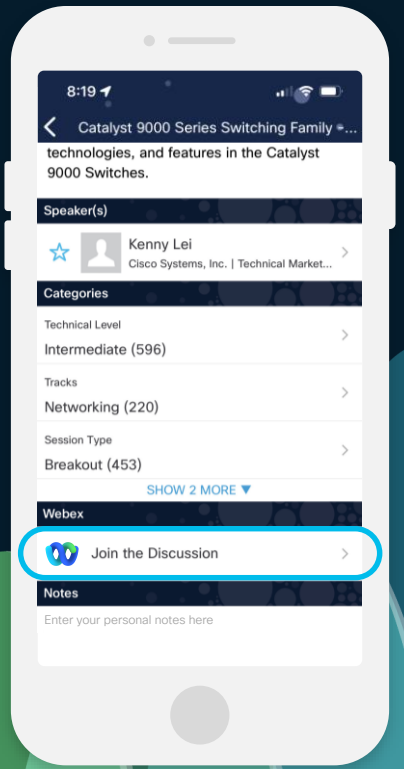
Questions?

Use Cisco Webex App to chat with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until June 7, 2024.

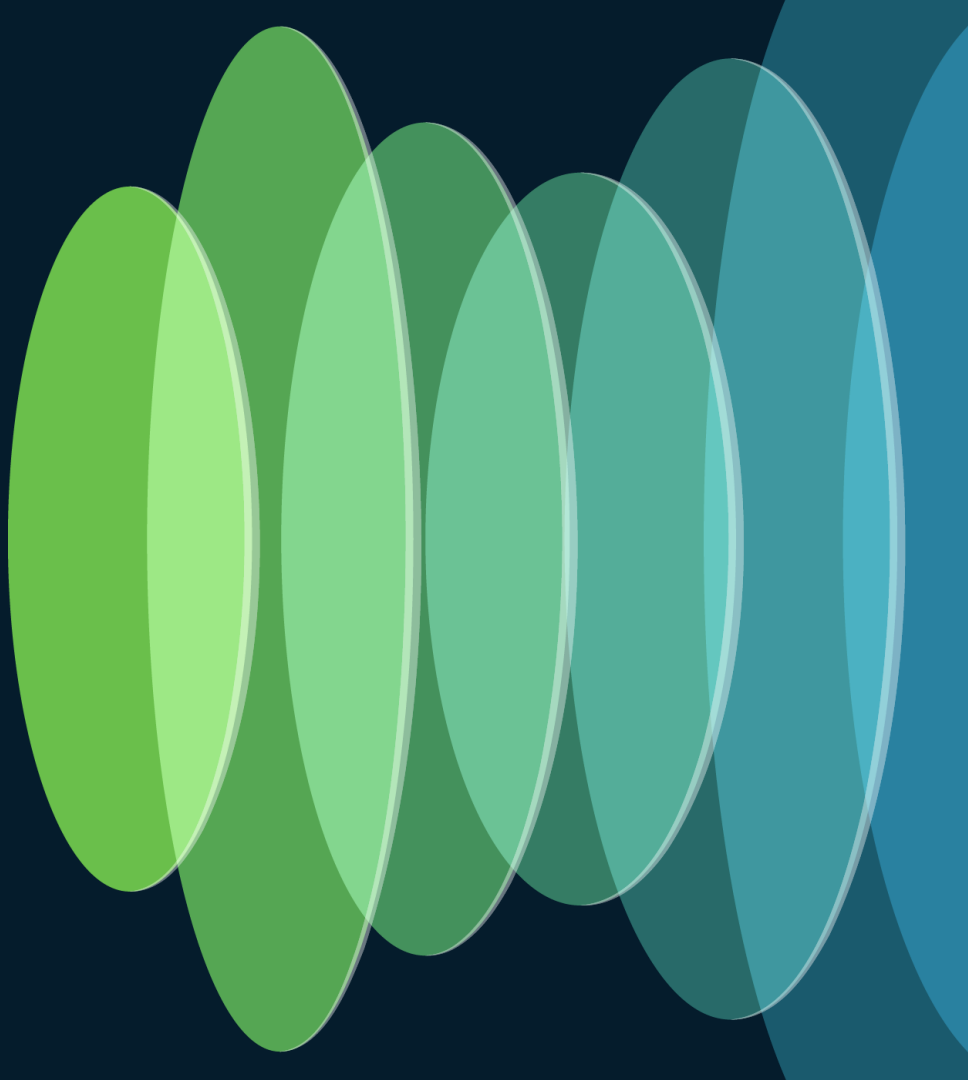


Agenda

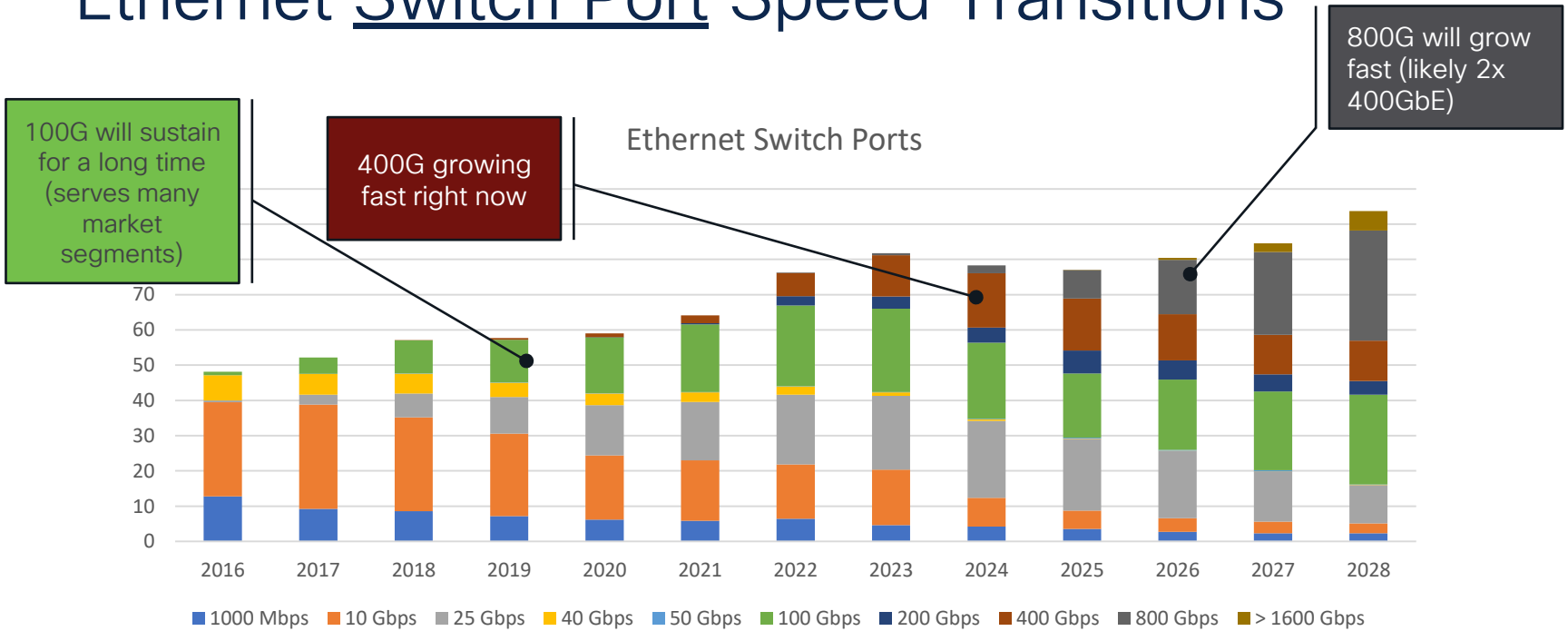
- What is driving the optics interconnect market right now?
- What does it mean for optics?
  - New speeds & altering priorities
  - New implementations
- Scaling out 400G, 800G
  - Solutions & options
  - Deployment considerations
- Wrap Up

Acknowledgements: This presentation would not exist without the inputs, expertise, and patience of many of our Cisco colleagues!

Market dynamics



Ethernet Switch Port Speed Transitions



- Majority of the highest speed transitions are webscale (top 8) customers
- Webscale will drive the speed transitions quickly to scale
- Rest of the market will leverage that scale in their own time frame

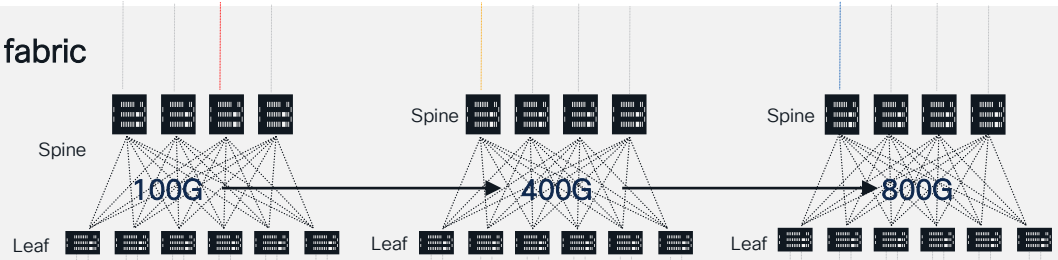
Source: Dell'Oro's Ethernet Switch – Data Center 5 Year Forecast Report 2023-2027, Jan 2024

Speed evolution in the data center

Inter-Data Center

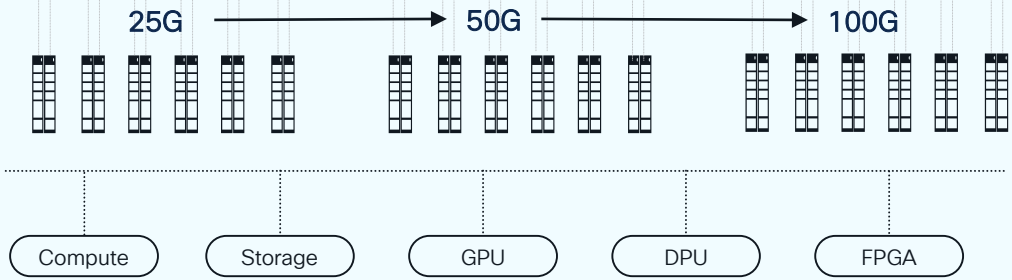
- Significant pluggable DCI (DWDM coherent)
- Open Line System

Switch fabric



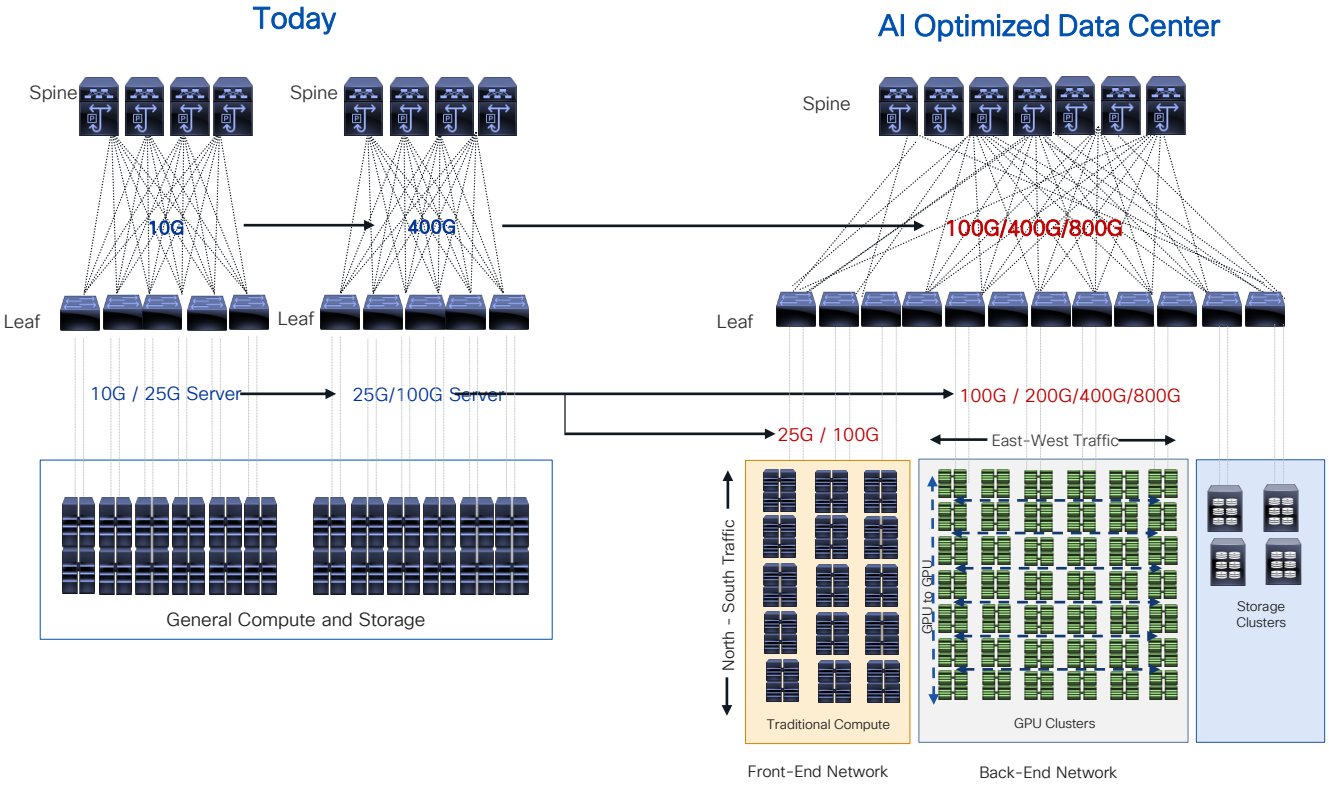
- Switch silicon bandwidth growing due to higher Radix and faster Serdes speeds
- Switch ASIC throughput growing: 6.4 Tbps to 12.8 Tbps to 25.6 Tbps to 51.2 Tbps
- Optics increasing from 40Gbps to 100G Gbps to 400Gbps to 800Gbps

Access



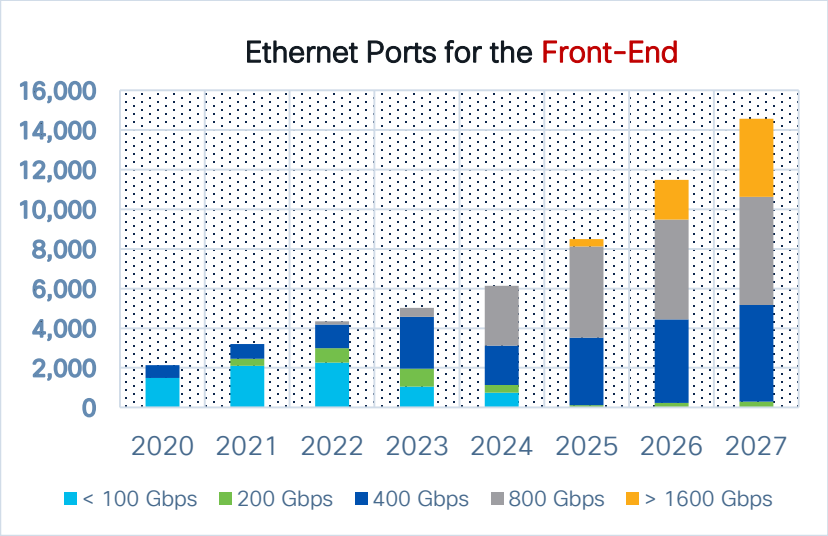
- Server network connectivity evolves with server processor upgrade cycles as data center traffic grows
- Server port speed is transitioning from 1/10 Gbps to 25 Gbps to 100 Gbps
- Storage, GPU, DPU, FPGA driving connectivity bandwidth, PCIe speed increase

Moving from a General-Purpose Compute Design to an AI/ML Optimized Infrastructure

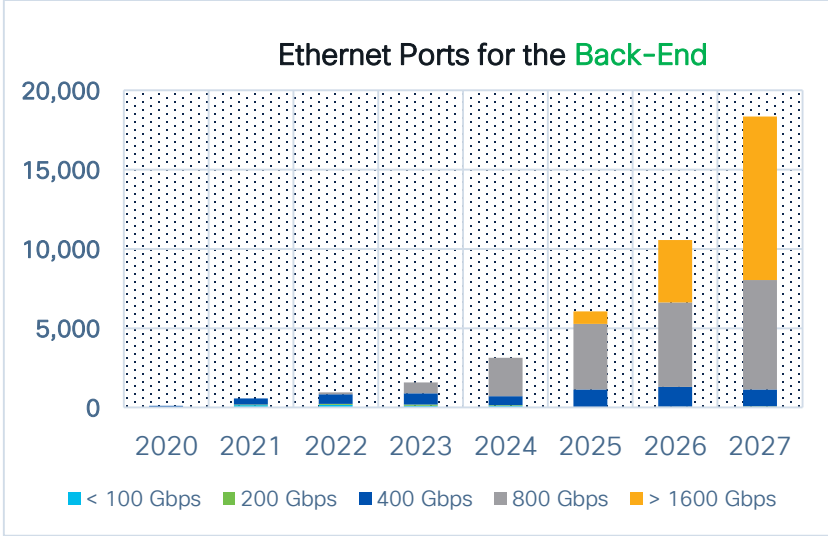


- **AI/ML workloads:** pushing data centers to evolve their network architecture
- **AI-Specific Networking**: a dedicated Back-End network for AI workloads to isolate them from other data center traffic and ensure low-latency communication.
- **Back-End AI/ML clusters:** consists of hundreds to thousands of AI/ML accelerators, CPUs, storage devices, Switches, and Network Interface Cards (NICs) connected to GPUs
- **High-Speed Interconnects:** Backend network requires high speed 100G/200G or 800G optics to connect servers and network switches. These high bandwidth connections are essential for handling the data generated by AI workloads

Ethernet Speed Transitions in AI Networks



Switch ports deployed in the front-end connectivity with Ethernet to grow from 4.3 M in 2022 to 14.6 M in 2027



Majority of the switch ports in AI back-end Networks to be 800 Gbps in 2025 and 1600 Gbps in 2027, showing a very fast migration to the highest speeds available in the market.

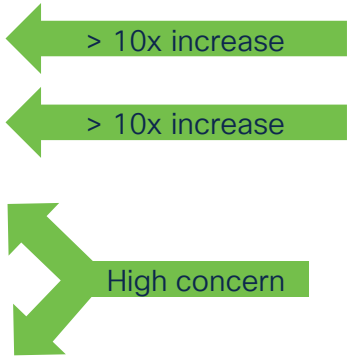
Dell'Oro: 95N16_AR_AI_Networks_For_AI_Workloads_Forecast_Tables_4Q23

Network operator top of mind(s)

	Traditional DC	AI/ML
Increasing capacity and sustainability	●	●
Preserve investments in existing optics infrastructure and cabling	●	●
Simplify operations and management of optical links	●	●
Preparing for capacity expansion	●	●

AI/ML is a disruptive event for equipment design

		Traditional Front-end DC	AI/ML Back-end DC
Rack Bandwidth (ToR/MoR)		3.2T-12.8T	>> 100T
Rack power		~10 kW	100 kW+
Packet Loss impact (reliability)		Low importance	Critical importance
Latency importance	Absolute	Low	Low
	Tail	-	High



Optimal switch and interconnect design is affected by these requirements

400G/800G/1.6T use cases

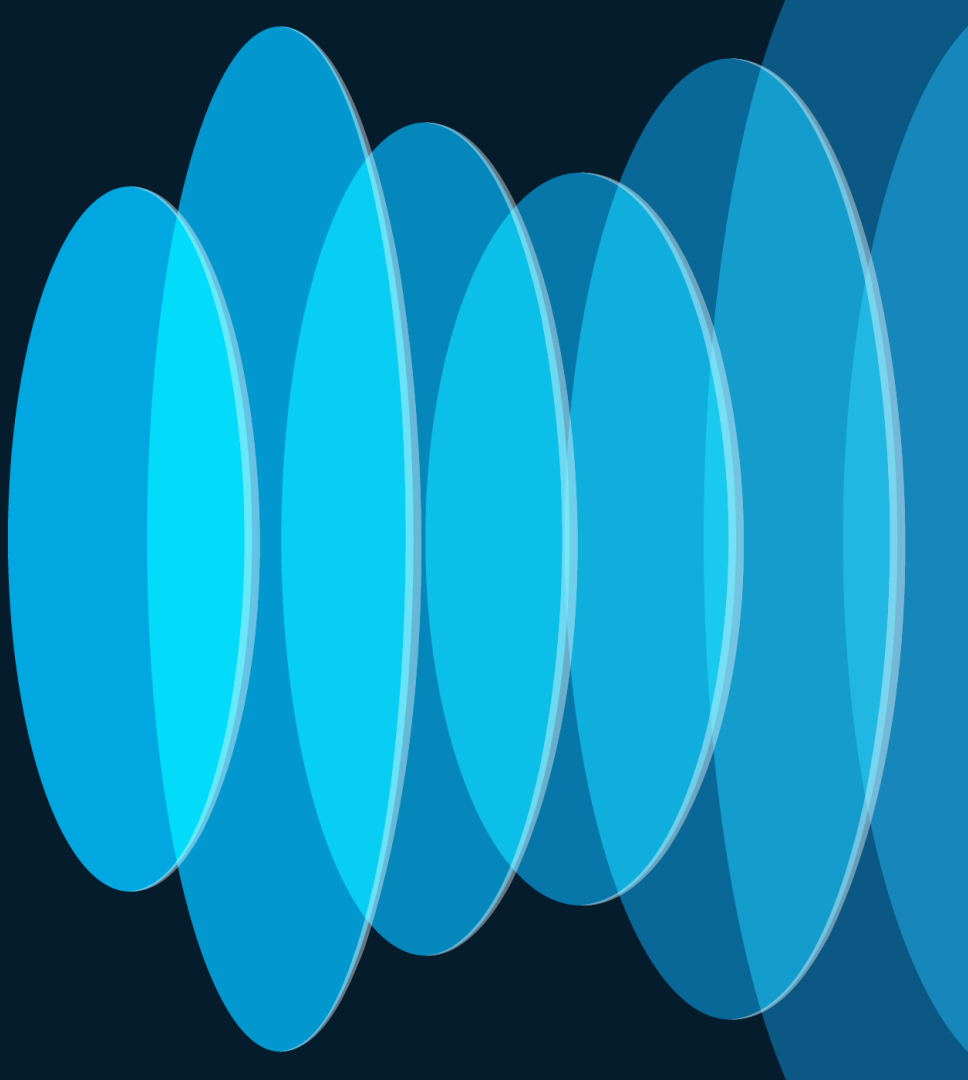
Cloud service providers

Telco service providers

Enterprise

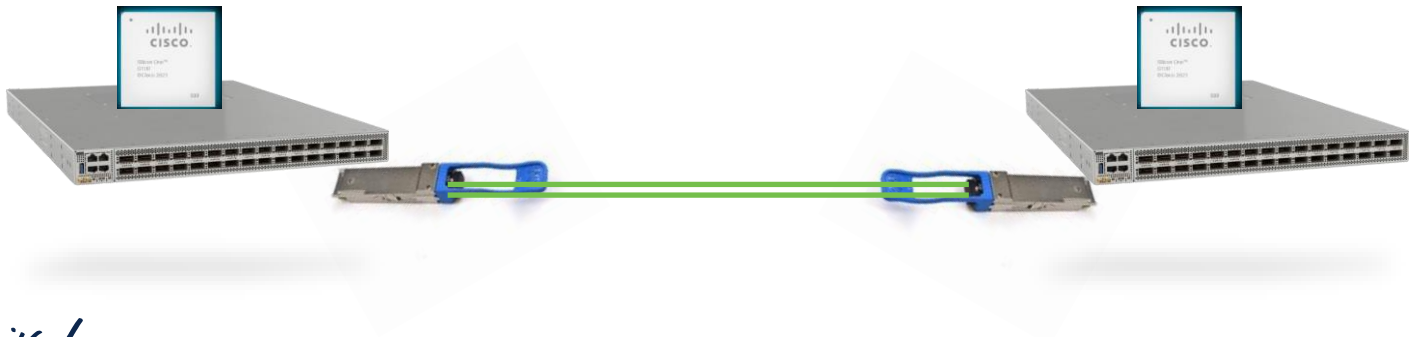
Media networks

How does AI/ML inflection point affect optics?



The function of pluggable optics

- The sole function of Optics is to extend the interfaces from one piece of equipment to another. The ASIC inside is driving the interface.
- Therefore, it is the ASIC capabilities and roadmap which primarily matter, and the role of optics is to keep up – without causing too many issues



Interconnects for an AI/ML world

AI/ML is a disruptive event for traditional networking

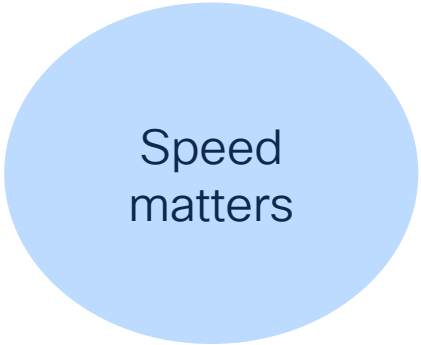
		Traditional Front-end DC	AI/ML Back-end DC
Rack Bandwidth (ToR/MoR)		3.2T-12.8T	>> 100T
Rack power		~10 kW	100 kW+
Packet Loss impact (reliability)		Low importance	Critical importance
Latency importance	Absolute	Low	Low
	Tail	-	High

Lots of interconnect
- → Speed matters
- → Power matters
- → Density matters

Massive rack density increase
- → Power matters
- → Copper cables matter
- → Density matters
- → Thermal solutions matter

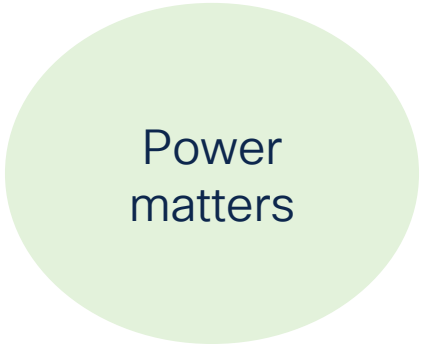
Job completion time (JCT)
- →Link BER performance critical.
- →Tail latency most important

Priorities for interconnect in an AI/ML world



Interfaces
800 GbE → 1.6 TbE

SERDES/ λ
100 Gb/s → 200 Gb/s



Lower W/Gb

Copper Cables

Linear pluggable

*compute plus networking

Rack density

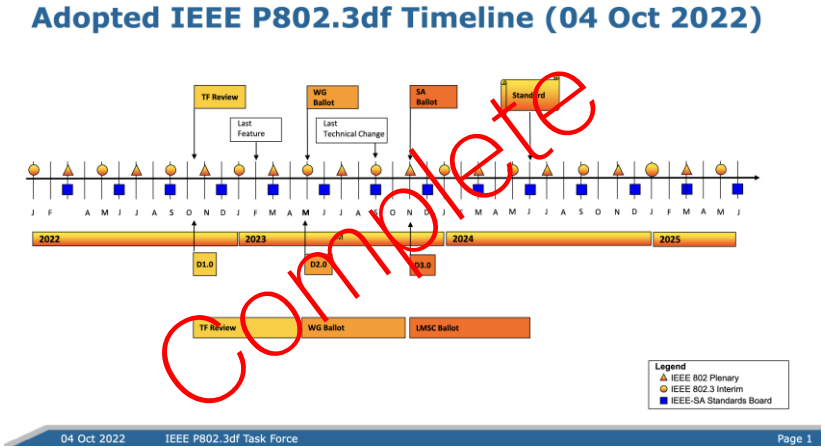
Switch density

Liquid cooling

800 GbE & 1.6 TbE Timelines

100 Gb/s based (P802.3df)

- 800 GbE based on an 8x 100 Gb/s approach (optical and electrical)

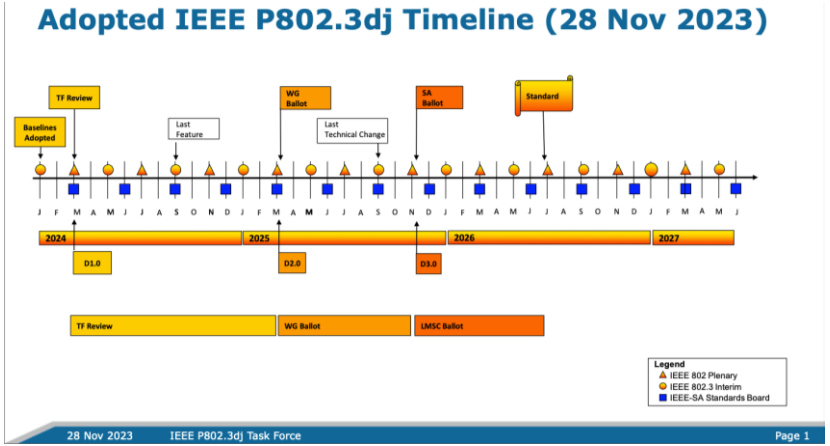


Complete

https://www.ieee802.org/3/df/proj_doc/timeline_3df_221004.pdf

200 Gb/s based (P802.3dj)

- 800 GbE (4x 200 Gb/s) and 1.6 TbE (8x 200 Gb/s) (optical and electrical)



https://www.ieee802.org/3/dj/projdoc/timeline_3dj_231128.pdf

What will be coming with 802.3dj

Summary IEEE P802.3dj Progress @ End of Mar 2024 Plenary – PMDs (& AUIs)

Ethernet Rate	Assumed Signaling Rate	AUI	Backplane	Cu Cable	SMF 500m	SMF 2km	SMF 10km	SMF 20km	SMF 40km
200 Gb/s	200 Gb/s	200GAUI-1 C2C C2M	200GBASE-KR1	200GBASE-CR1	200GBASE-DR1	200GBASE-FR1			
400 Gb/s	200 Gb/s	400GAUI-2 C2C C2M	400GBASE-KR2	400GBASE-CR2	400GBASE-DR2	400GBASE-DR2-2			
800 Gb/s	200 Gb/s	800GAUI-4 C2C C2M	800GBASE-KR4	800GBASE-CR4	1.800GBASE-DR4 2.800GBASE-FR4-500	1. 800GBASE-DR4-2 2. 800GBASE-FR4	800GBASE-LR4		
	800 Gb/s						800GBASE-LR	800GBASE-ER1-20	800GBASE-ER1
1.6 Tb/s	100 Gb/s	1.6TAUI-16 C2C C2M							
	200 Gb/s	1.6TAUI-8 C2C C2M	1.6TBASE-KR8	1.6TBASE-CR8	1.6TBASE-DR8	1.6TBASE-DR8-2			

Coherent optics

Copper cables (~1m)

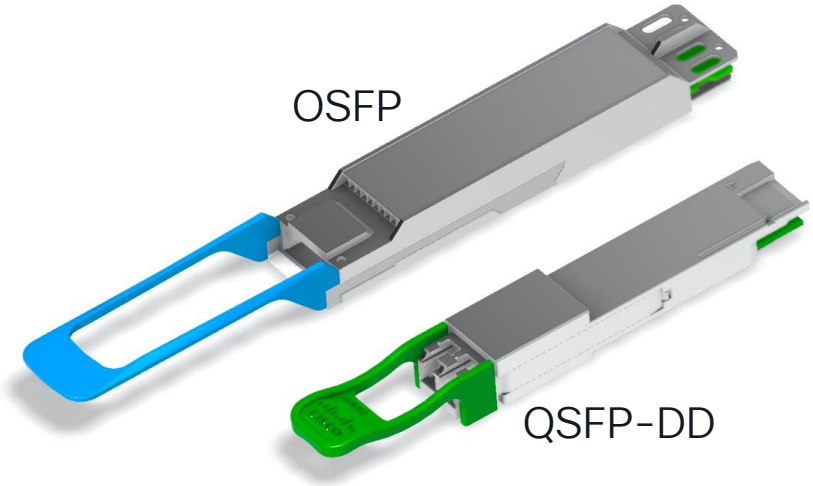
IMDD optics

https://www.ieee802.org/3/dj/projdoc/Baseline_Status_24_0314.pdf

Highest rate and breakout rates done in same project

- 1.6 TbE, 800 GbE, 400 GbE, 200 GbE

Pluggable Optical Modules: QSFP-DD or OSFP

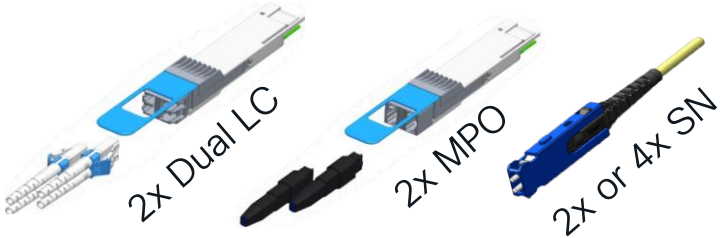


400G/800G/1.6T variants

Both variants support all the technical requirements:

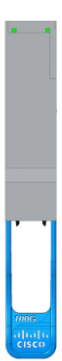
- 32 ports in 1 RU is feasible
  - 64 ports in 2RU is feasible
- QSFP-DD is compatible with QSFP
- QSFP-DD enables better system co-design which has thermal advantages

Breakout optical connector options[1]



[1] only QSFP-DD shown but similar on OSFP

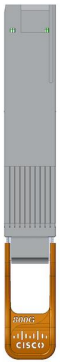
QSFP-DD Extends to 1.6T



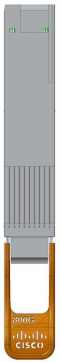
100 Gb/s
25 Gb/s SerDes
QSFP28

400 Gb/s
50 Gb/s SerDes
QSFP56-DD

400 Gb/s
100 Gb/s SerDes
QSFP112



800 Gb/s
100 Gb/s SerDes
QSFP-DD800



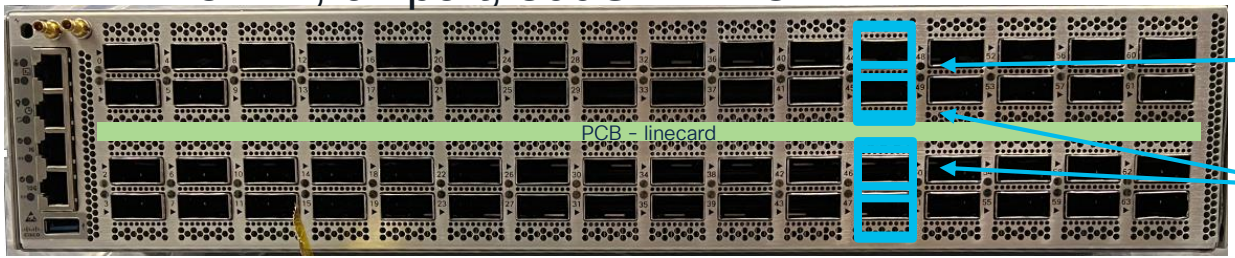
1600 Gb/s
200 Gb/s SerDes
QSFP-DD1600

Complete backwards compatibility

Powerful value proposing to enable seamless network growth and investment protection. Technically superior solution.

800G Optical Modules: QSFP-DD or OSFP

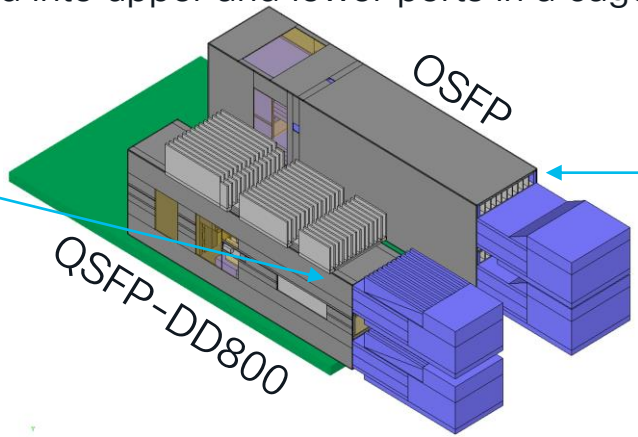
51.2T, 64 port, 800G in 2RU



Showing two modules inserted into upper and lower ports in a cage. Heatsink differences are the key.

Module slides under heatsinks.

- Heatsink design can change and optimize per equipment design

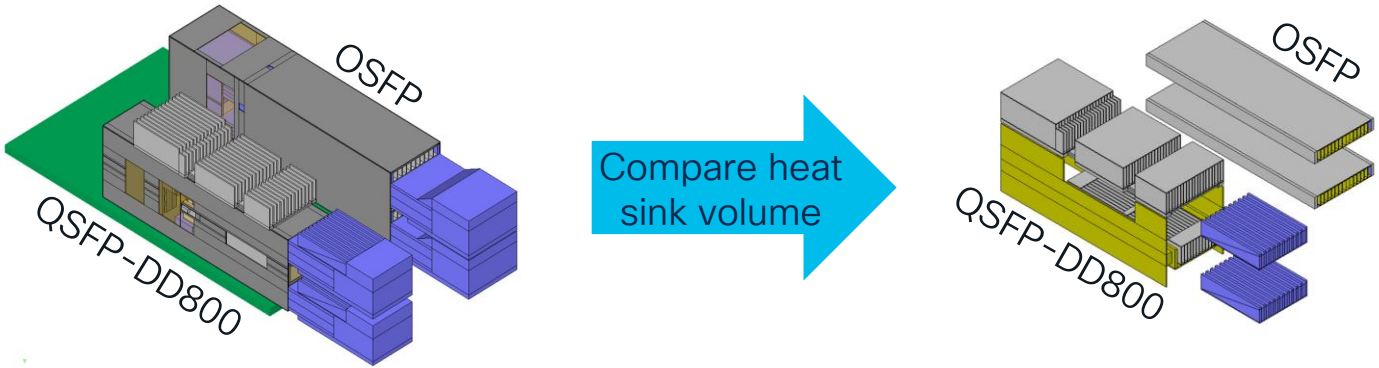


Heatsink is integrated with module

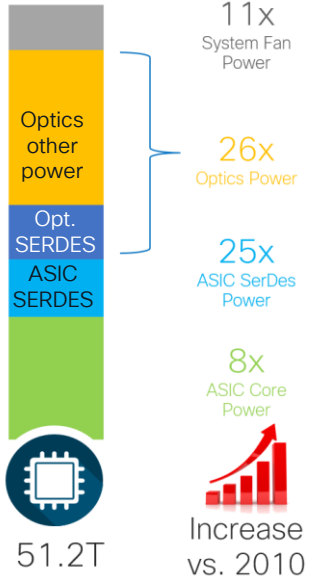
QSFP-DD – avoids the “risk” of the integrated heatsink

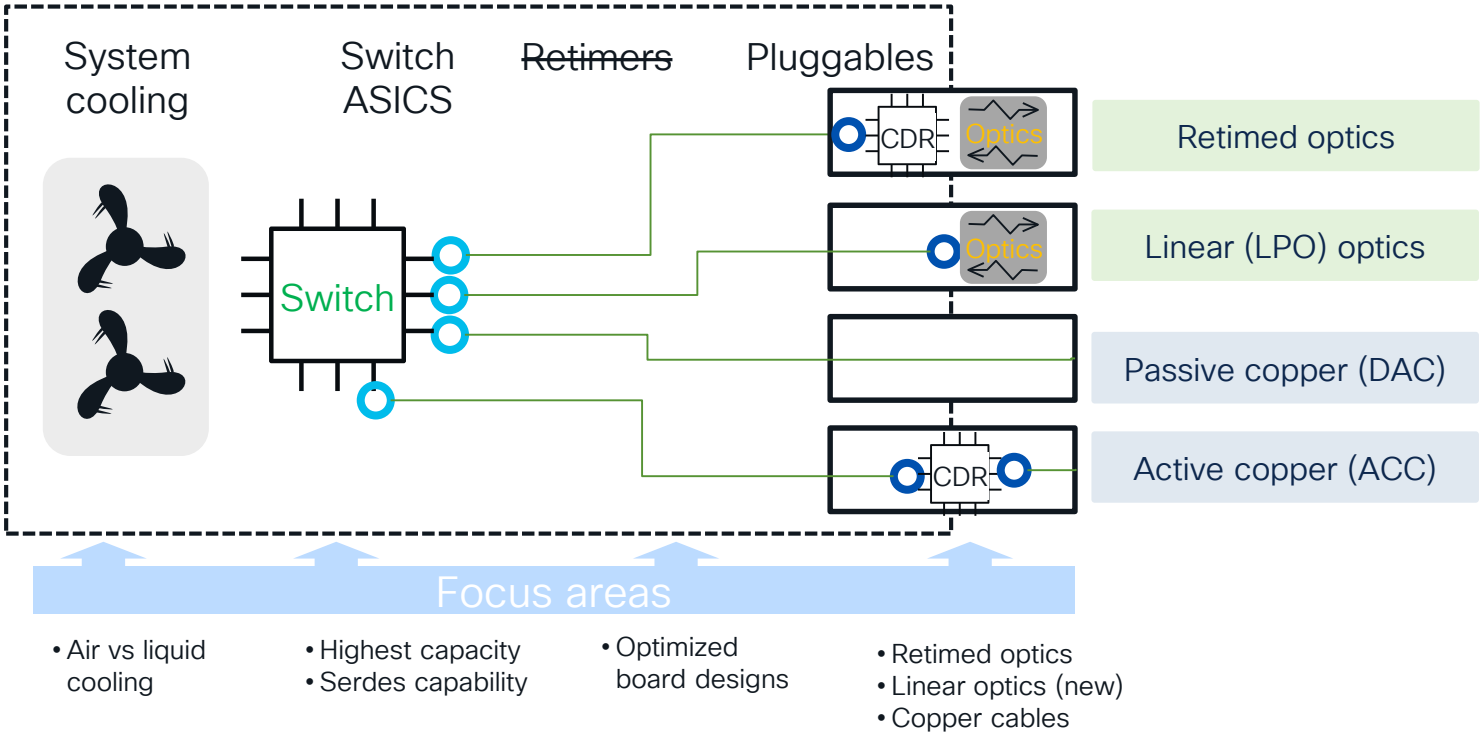
QSFP-DD’s flat top (and bottom) case design: allows **systems to optimize riding heat sinks**.

- *This fundamentally is the BREAK of that intuitive perspective*
- In an air-cooled system, the best way to cool anything is to **increase the heat sink fin area**.
- QSFP-DD is **innovation friendly**: yields **constant improvements** to cage design, faceplate design, heatsink design, air flow balance etc. Friendly to ASIC cooling
- While an integrated heat sink seems better, a QSFP-DD design can have nearly 3x the heatsink volume. *OSFP heatsink was defined in ~2015 and can’t be changed*



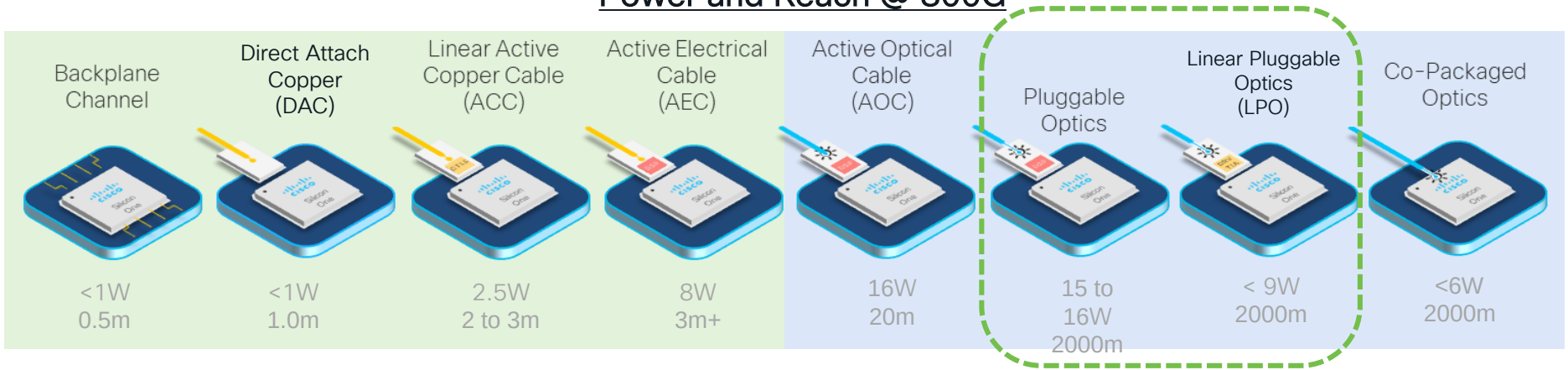
Interconnect and power





High-Density Interconnect options: Power vs. Reach Tradeoff

Power and Reach @ 800G



Linear Pluggable Optics (LPO)

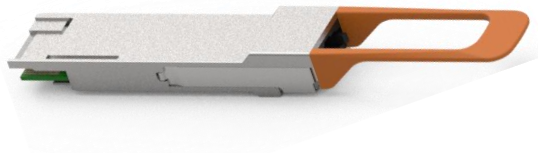
- very attractive as a solution for a piece of the power challenge
- But it needs an ecosystem – and that isn't settled yet.

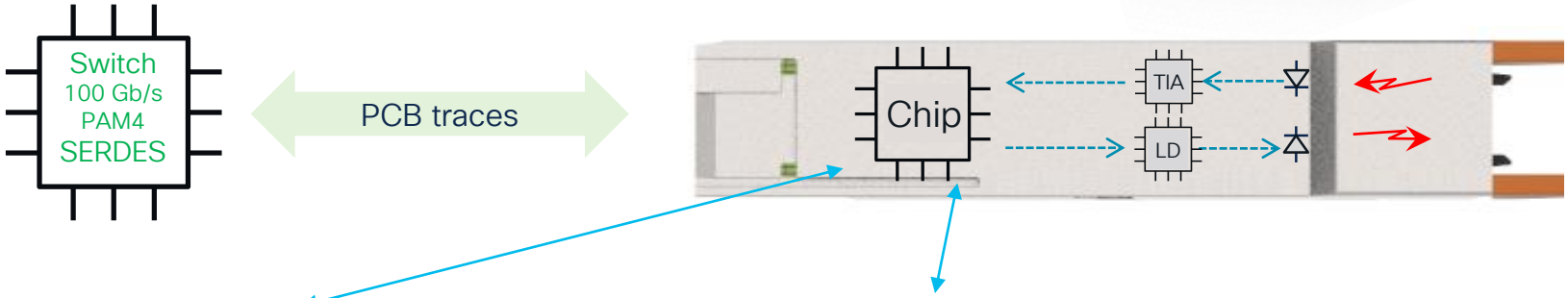
LPO-MSA

Newly formed MSA with expertise from across the industry

Accelink, AMD, Arista, Broadcom, Cisco, Eoptolink, Hisense Broadband, Innolight, Intel, Macom, Nvidia, Semtech

Linear pluggable optics (LPO)





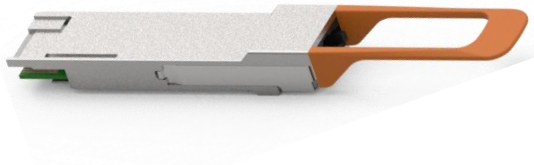
Current (Retimed)

- DSP has full digital equalization for both electrical & optical signals
- Enables broad interoperability
  - Host/Port/Module
- Full telemetry & loopbacks possible
- FEC Monitoring or Partitioning possible
- But this adds power

LPO (Non-retimed)

- Power reduction due to absence of digital equalization
- Performance is based on quality of every component in link as they are concatenated
  - Optics, Serdes, PCB, connectors
  - Varies port to port
  - Some linear gain and equalization in module
- Interoperability becomes more challenging
- Loss of some telemetry, monitoring or loopback

How does LPO work?

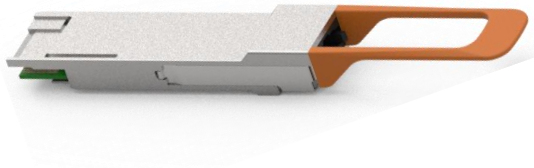


ASIC Serdes are the key

- Very capable ASIC serdes Tx and Rx equalization capabilities
- Serdes are designed and capable to drive linear CR channel (copper cable)
- CR channels are specified (IEEE) with lower host loss ports
- LPO modules need to work in every switch port, so they need to compensate for the extra host loss in those high loss ports
- LPO modules have limited linear gain and equalization that can compensate for host channel
- Linear components are key – SiPhotonics helps.

This is the big change!

What is Half-linear (LRO)

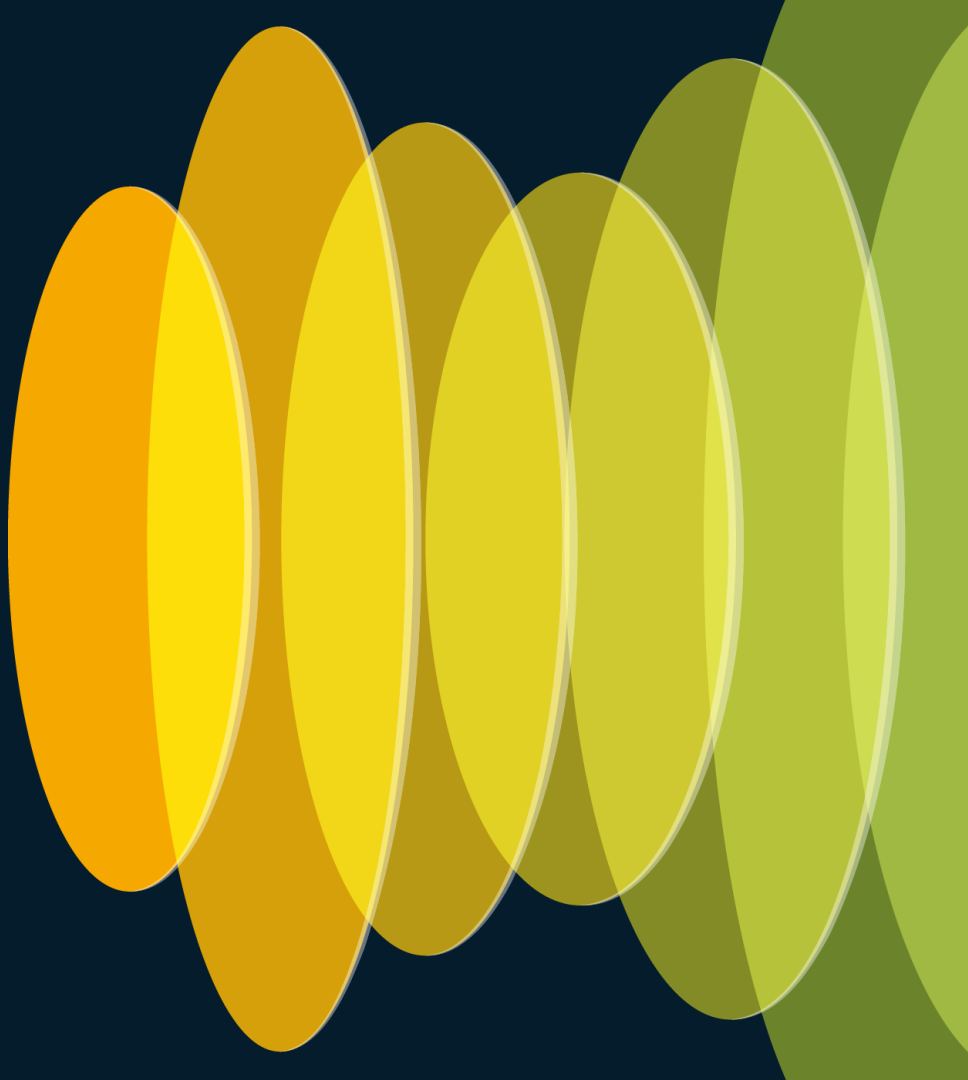


Half Linear (LRO, sometimes called TRO)

- Pitched as a way of halving the problem. Lower power half of DSP used.
- Reduces ability of host Tx ASIC serdes' to optimize the full end-to-end link performance.
- Measurements so far show that link performance is ~1-2 orders of magnitude worse than LPO.

For AI/ML link performance (BER margin) is very important. Any bit error can force a packet retransmit which delays the AI/ML job completion time.

Current state of art



It’s all about ASICs, SerDes, and Optics

Cisco 8201-32F | 32 ports of 400GE in 1U

640G | 1.28T | 3.2T | 6.4T | 12.8T | 25.6T | 51.2T | 102.4T | 204.8T

← 204.8T is 16x the bandwidth of 12.8T

320x increase in switching bandwidth over 16 years

Switching Capacity

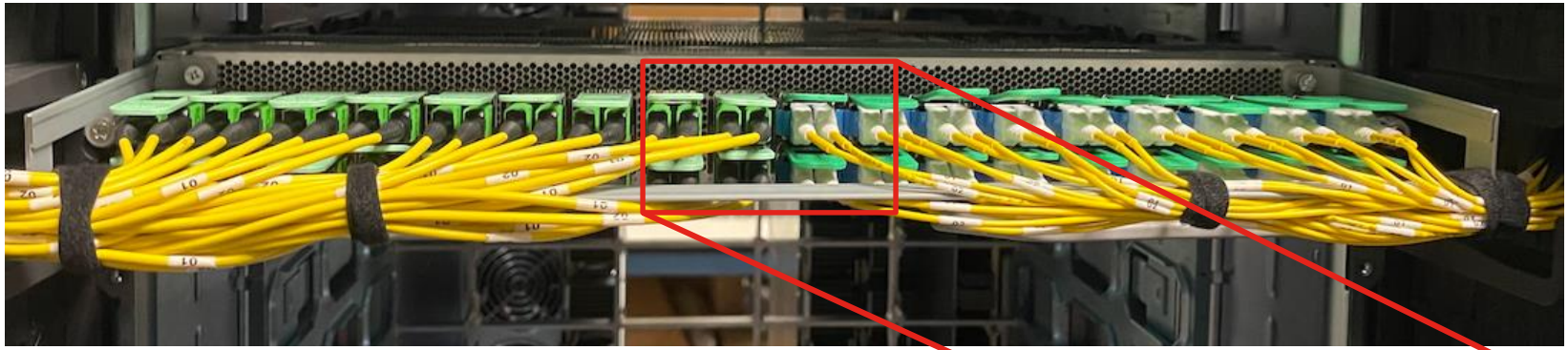
Cisco One ASICs

Year	2010	2012	2014	2016	2018	2020	2022	~2024 ?	~2026 ?	in 16 years ...
ASIC # SerDes	64	128		256		512			1024	16x
Serdes Gbps	10G NRZ		25G NRZ		50G NRZ		100G PAM4	200G PAM4		20x
Optics/Port	QSFP+		QSFP28		QSFP-DD56		QSFP-DD800	QSFP-DD1600		5 pkgs
Port Gbps	40G		100G		400G		800G	1.6T		40x
Ports/Router	16	32		64	32	64	64	64	128	

ASIC density continues to redefine how products are built.
Gates & GHz. SerDes & Interconnect. Optics & wavelengths.

Credit: adapted from slide by Rakesh Chopra & John Chapman
https://blogs.cisco.com/sp/co-packaged-optics-and-an-open-ecosystem

32p 1RU 25.6T Systems using QSFP-DD800



1RU 32 port
QSFP-DD800

2x400G
Dual MPO
Dual LC

Cisco Silicon One G200: Retimed vs. LPO System Power Comparison

Cisco demonstrated LPO operation at OFC2024

<u>Set-up:</u>

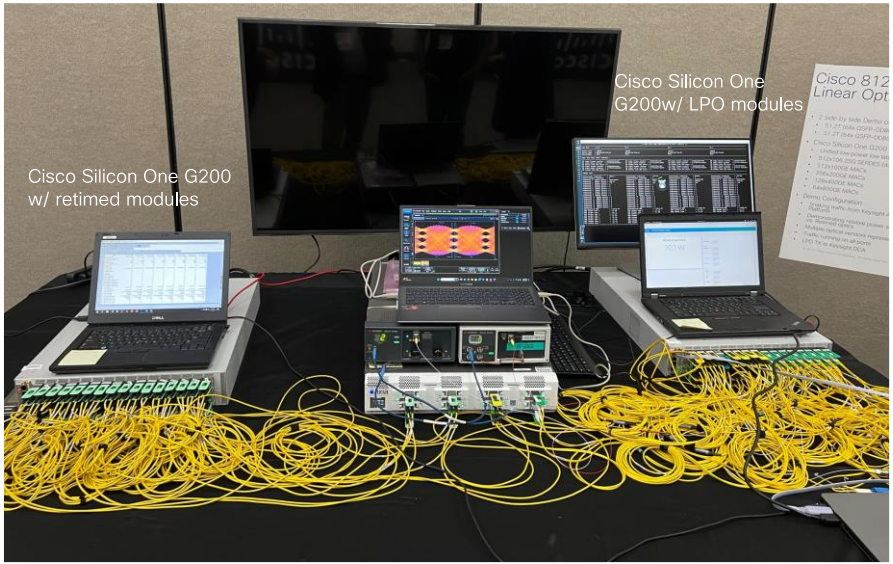
Two identical Silicon One 51.2Tb 64-port G200-based switches:

- 100% retimed optics
- 100% LPO optics

<u>Result:</u>

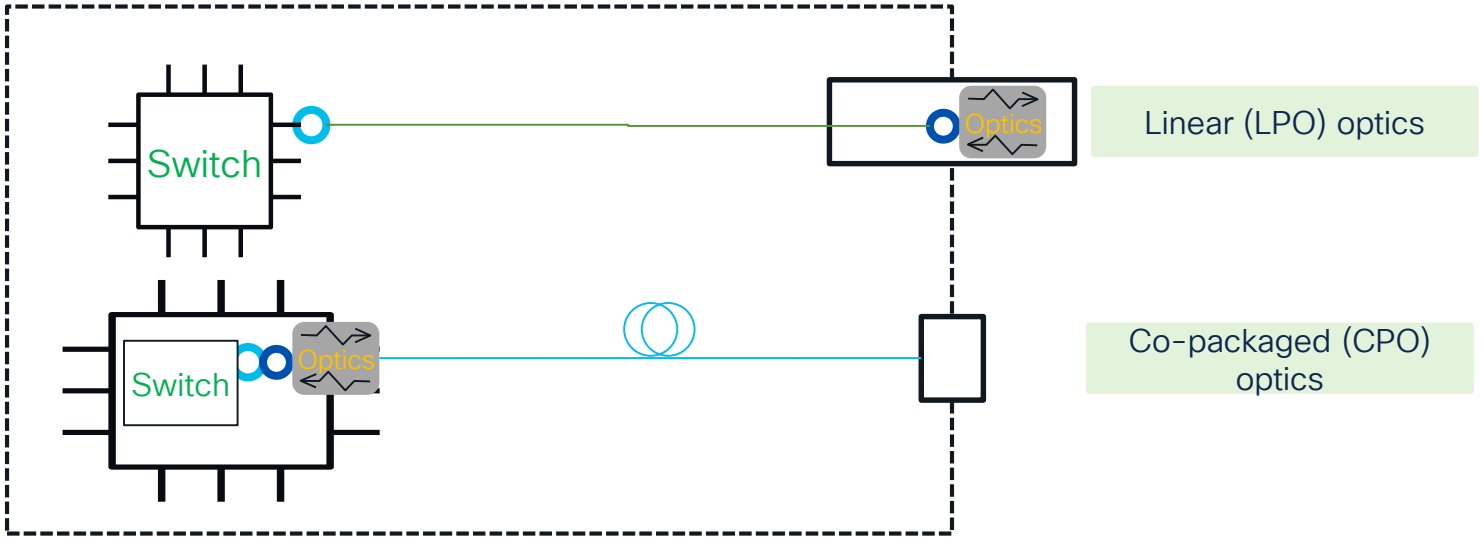
Both switches ran full traffic on all ports

Overall power reduction: ~ 700W



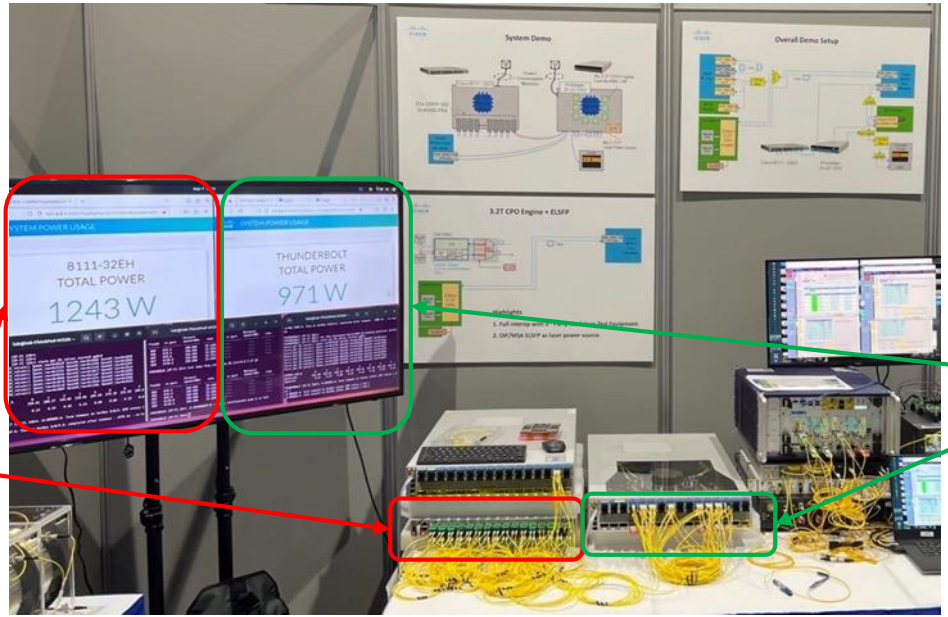
CPO & LPO

Co-packaged optics (CPO) and Linear Pluggable Optics (LPO) are two implementation variants of the same idea – reduce ASIC to optics power/DSP



Cisco Silicon One: OFC 2023 Demo – 25.6T
Co-packaged Optics vs Retimed

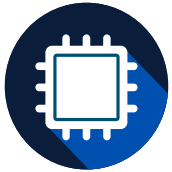
CPO system achieved ~ 22% total system power reduction



25.6T Ethernet Switch
64 ports @ 400G QSFP-DD

25.6T Ethernet Switch
Enabled by CPO and using ELSFP external light sources

800G does not necessarily mean 800 GbE



25.6T

SERDES

100G

X256

Fixed Configs

1 RU	32p @ 800G
	QSFP-DD800, $OSFP_{800G}$
2 RU	64p @ 400G
	QSFP-DD, OSFP

25.6T ASIC roadmap and system density requirements drive to 800G ports

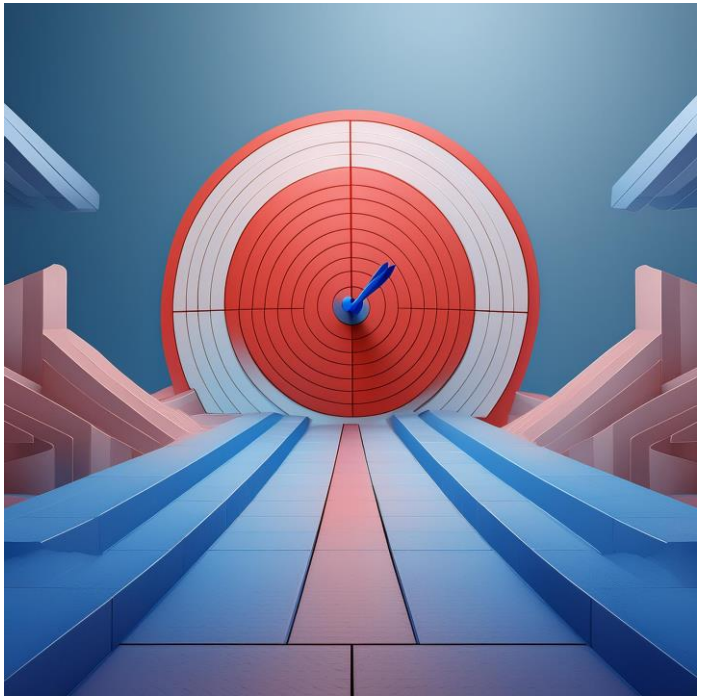
- Dominant interest in 800G module today is to support 2x 400 GbE breakout
- No immediate network need for 800 GbE
  - Although IEEE working on it (more later)
- 800G modules have the same issues:
  - Thermals and Signal integrity
  - Backwards compatibility?

Greater density requires advanced cooling

If rack capacities grow from ~10kW → >100kW something needs to remove the heat. Two solutions under development:

- Cold plate liquid cooling
  - Liquid circulated through pipes and cold plates which are attached to key components (ASIC, optics) to efficiently remove heat
- Immersion liquid cooling
  - Equipment is immersed in special non-conductive liquids to very efficiently remove heat

Key takeaways

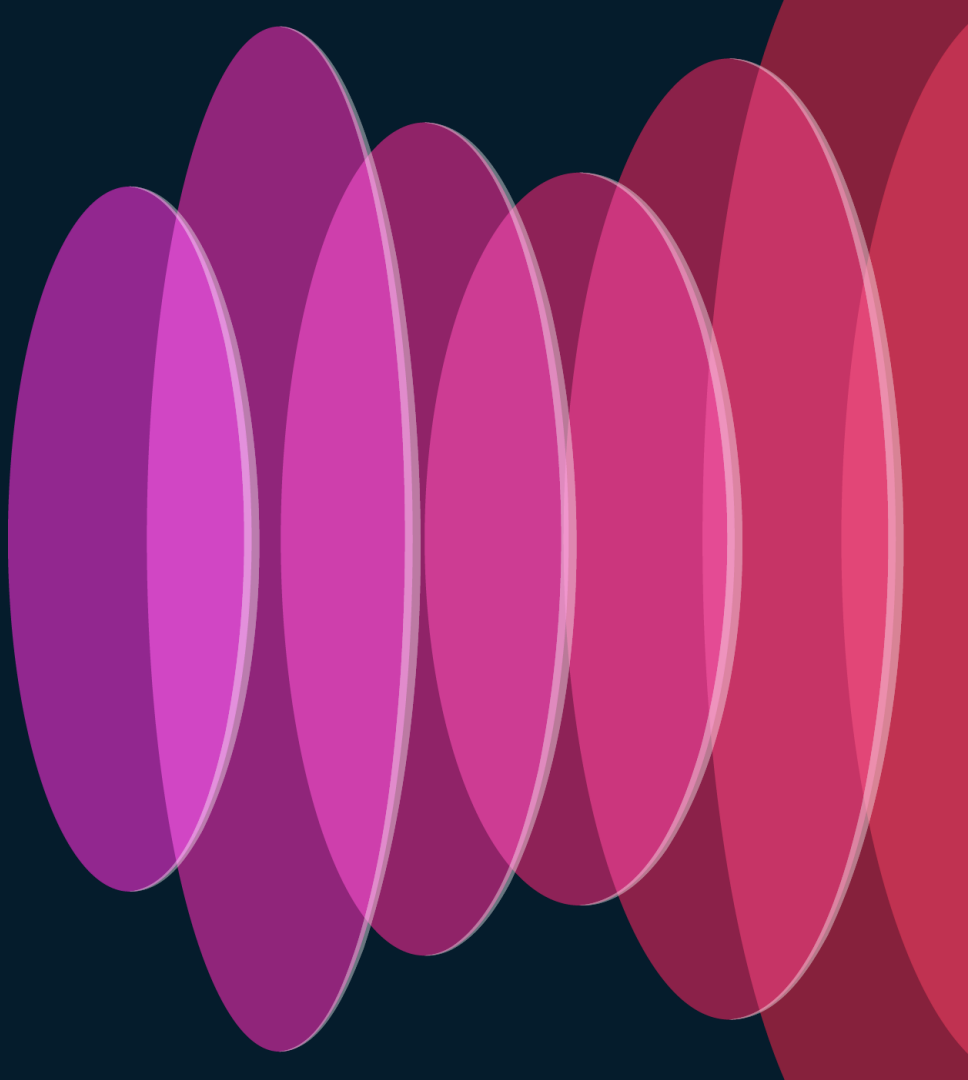


(AI generated pic)

Requirements around AI/ML are driving interconnect right now

- Faster speeds – 800G → 1.6T → …
- Lowering power is key
  - Copper cables still live @ 200 Gb/s
  - Novel implementations like LPO / CPO
  - Faster speeds yield better energy efficiency (W/Gb)
- Advanced cooling strategies are being considered
  - Liquid cooling and Immersion cooling
- AI networking will dominate the future design models for data centers

Scaling out 400G & 800G



Is 400G & 800G ready for broader adoption?

(AI generated pic)

The early progress of optics & interconnects is steered by the needs and use cases of the early adopters.

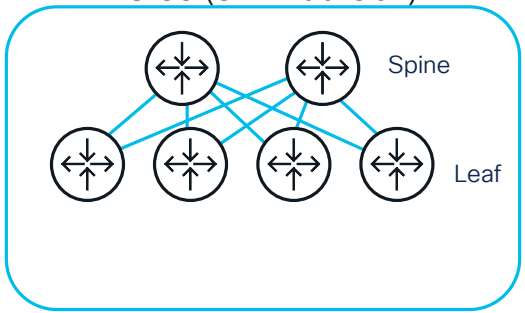
All adopters benefit from the economies of scale and technology maturity that results.

Looking at 400G/800G today we already have:

- 400 GbE & 800 GbE Ethernet defined
- Standardization of DCI interconnect (400ZR, 800ZR)
- Denser switches and optics, breakout deployments

Why move to higher speeds?
400G → 800G example (same is true for 100G→ 400G)

25.6T user capacity using multiple switches with 12.8T ASICs (32x 400 GbE)



50 Gb/s ASIC IO (SerDes)

64 ports of 400GbE
(256 ports of 100 GbE)

~3000 Watts
26,280 kWh/year

=

25.6T user capacity using single switch with 25.6T ASIC (32x 800 GbE)

100 Gb/s ASIC IO (SerDes)

32 ports of 800G
(64 ports of 400 GbE
256 ports of 100 GbE)

~400 Watts
3,504 kWh/year

=

Up to 87%
Energy Savings

83% less space/fans

400G & 800G Today

Client (aka IMDD) (Short reach)		Coherent (Long Reach)
400 GbE, 800GbE	Mature standards	400G (OIF, Open ZR+)
100 Gb/s	Mature technology	16QAM @ 60 Gbd, 56G Serdes
800 GbE @ 200 Gb/s	Next Gen technology	800ZR: 16QAM @ 118 Gbd, 112G Serdes
Incl. Breakout	Mature Deployment	400ZR, Bright 400ZR+

Most 400G/800G Standards are done

Cisco led many of these efforts

☑ Complete

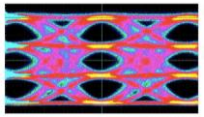
Cisco-led

Standards	IEEE 802.3bs ✓ IEEE 802.3cd ✓ IEEE 802.3cm ✓ IEEE 802.3cn ✓ IEEE 802.3ct ✓ IEEE 802.3cu ✓ IEEE 802.3ck ✓ IEEE 802.3db ✓	400 GbE & 200 GbE MAC & Initial Interfaces 50 GbE MAC & Interfaces (also 100 GbE & 200 GbE PMDs) 400 GbE MMF (BiDi and SR8) Extended reach (40km) 50 GbE, 200 GbE, 400 GbE 100GbE Coherent 80km 100G-FR, 100G-LR, 400G-FR4, 400G-LR4-6 100GE serdes 100/200/400GE MMF (100Gb/s short wavelength)
	OIF400ZR ✓ OIF800ZR OIF1600ZR Open ZR+ ✓	400 GbE Coherent 120km 800 GbE Coherent 120km 800 GbE Coherent 120km Long Distance 400G
	802.3df ✓	200G/400G/, 800G Ethernet Task Force @ 100Gb/s per lane
	802.3dj	200G/400G/800G/1.6T Ethernet Task Force @ 200Gb/s per lane
	802.3dk	Greater than 50 Gb/s Bidirectional Optical Access PHYs Task Force.
MSAs*	LPO MSA	100G Linear pluggable
	100G Lambda MSA ✓	100G-FR, 100G-LR, 400G-FR4, 400G-LR4
	QSFP-DD MSA ✓	400G/800G/1.6T Form factor
	OSFP MSA ✓	400G/800G/1.6T Form factor
	400G-BiDi MSA ✓	400 GbE MMF BiDi

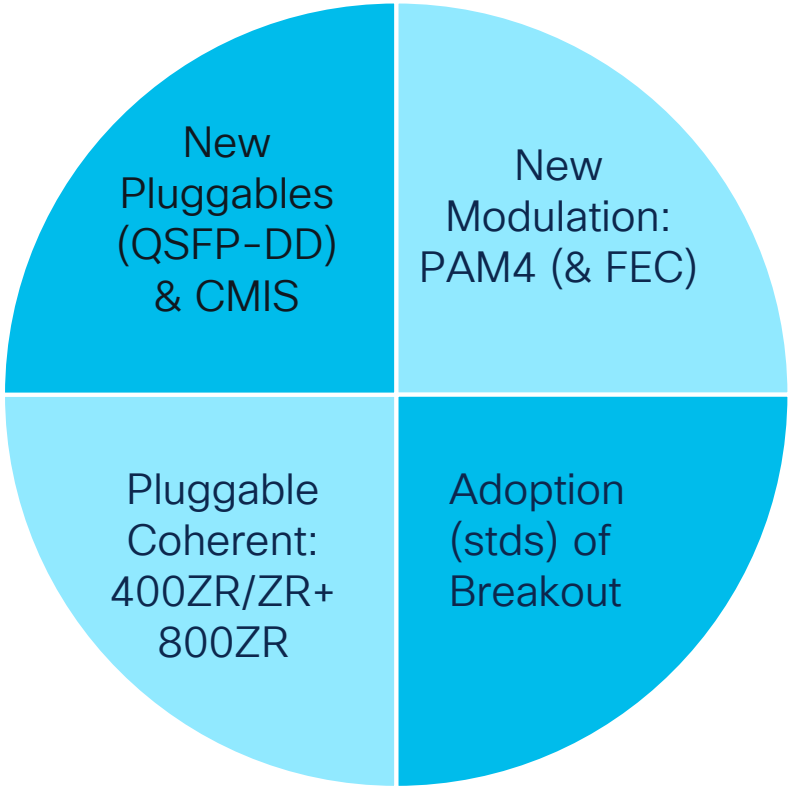
Flexibility of 400G / 800G pluggable modules

- Copper cables
- Multimode Fiber – 100m
- Single Mode Fiber inside DC – 500m & 2km
- Single Mode Fiber Campus – 10 km
- Outside plant, DCI – 100-1300 km

Key elements of 400G & 800G optics



New Pluggables with 8-wide connector to support 400G and 800G ports

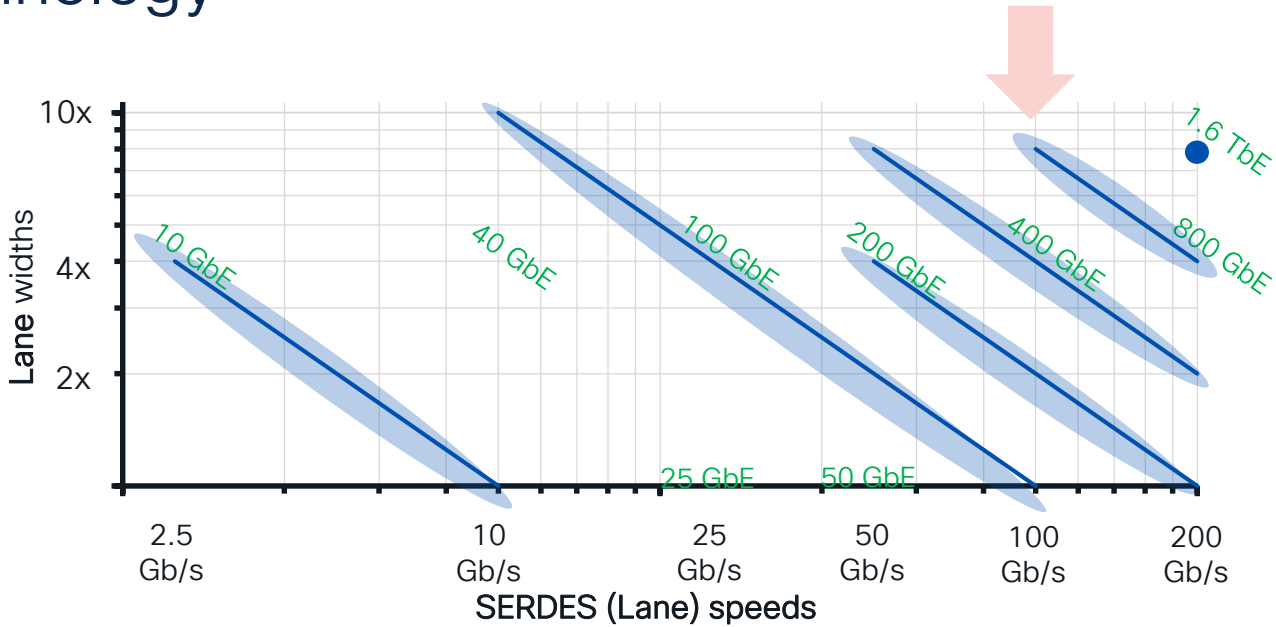


Higher speed interfaces adopted PAM4 modulation.
Ubiquitous use of FEC.

Long reach coherent without any system port density reduction
→ Routed Optical Networking

Pluggable modules supporting multiple lower speed interfaces

400 GbE / 800 GbE – Uses mature 100 Gb/s technology



Technology maturity of 100 Gb/s technology enables:

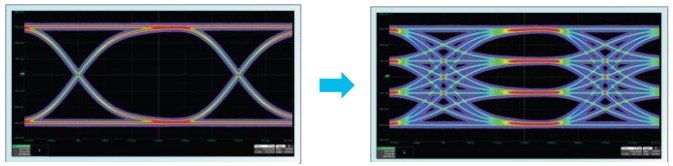
- Excellent performance
- Excellent interoperability
- Broad adoption

Technologies used for 400G, 800G optics

Enabling higher performance but lowering cost

Advanced Modulation

Client optics



NRZ
1 bit/sec/symbol

PAM4
2 bit/sec/symbol

Coherent optics

Coherent QPSK
2 bit/sec/symbol

Coherent 16-QAM
4 bit/sec/symbol

Using more complex modulation allows us to increase the data rate (Gb/s) without increasing the signaling speed (Gbaud)

Forward Error Correction (FEC)

Allows correction of errors at receiver

Enables use of relaxed specs (saves $) to get same performance or enables much higher performance.

Usually embedded in Ethernet switch or module DSP ASIC

Common Management Interface Spec (CMIS)

Standardized now in OIF – ubiquitous adoption for 400G modules and above

Host (switch/router)

High speed
electrical interface
(OIF CEI, IEEE, etc)

Pluggable Module

Optical/Copper
(IEEE, MSAs, OIF)

Management
- configure
- initialize
- monitor

Often overlooked, CMIS standardizes management between modules and hosts

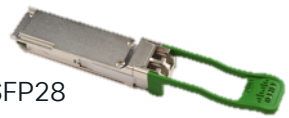
Consistent definition for configuration, initialization, monitoring, telemetry, firmware update etc.

100G Single Lambda Optics

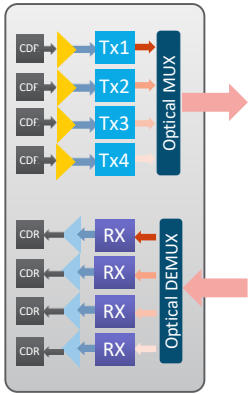
Journey to 400G

100G Lambda MULTI-SOURCE AGREEMENT

Before



QSFP28

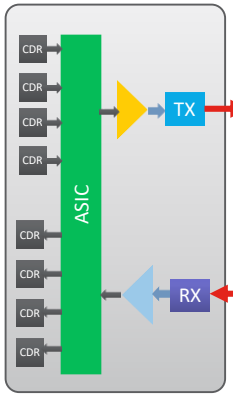


4x25Gb/s electrical interface

4x25Gb/s optical interface (multiplexed on a duplex fiber pair)

Now

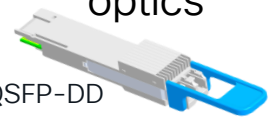
QSFP28



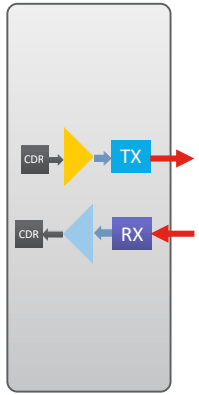
4x25Gb/s electrical interface

1x100Gb/s optical interface (on a duplex fiber pair)

400G/800G optics



QSFP-DD



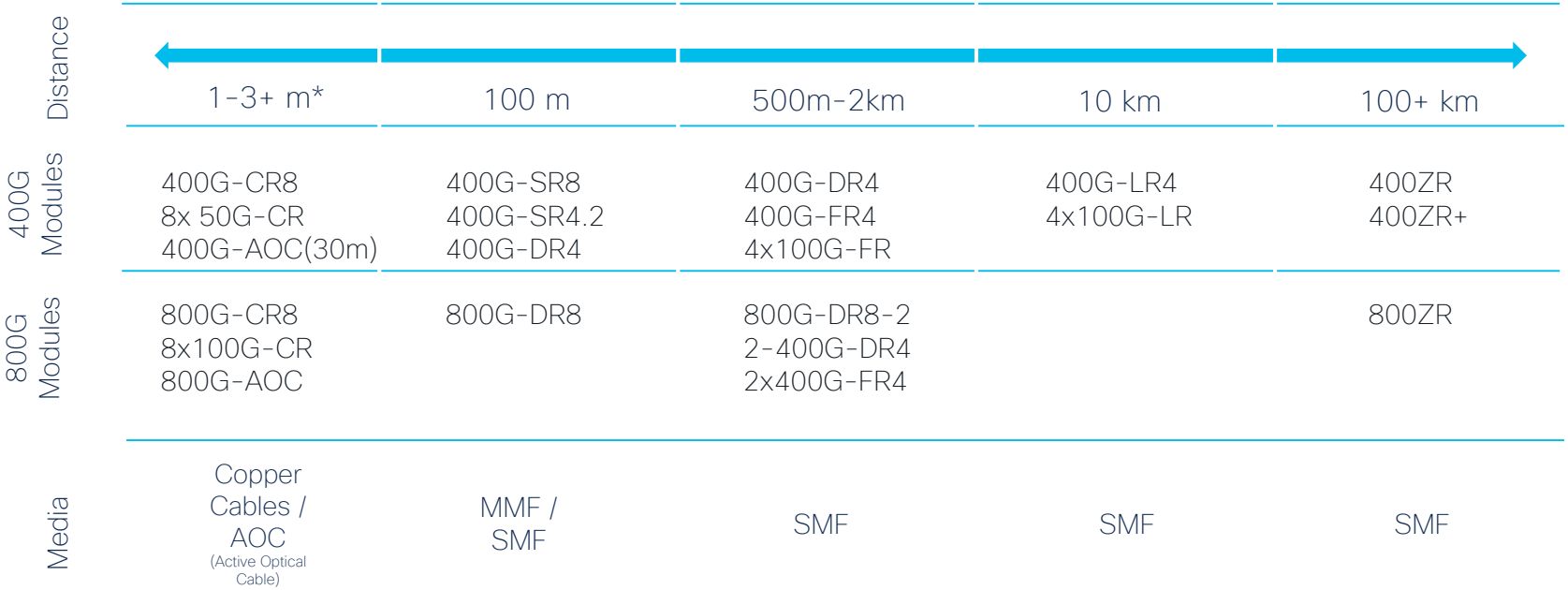
nx100 Gb/s nx50 Gb/s electrical interface

nx100Gb/s optical interface (on a duplex fiber pair)

QSFP-100G-SR4	100m	500m	QSFP-100G-DR
QSFP-100G-SM-SR	2km	2km	QSFP-100G-FR
QSFP-100G-LR4	10km	10km	QSFP-100G-LR
QSFP-100G-ER4L	40km	25km	QSFP-100G-ERL

Forward compatible optical network interface!

400 GbE & 800 GbE modules and use cases
Portfolio of interfaces available for use



* Meets or exceed IEEE specs. Depends on ASIC serdes

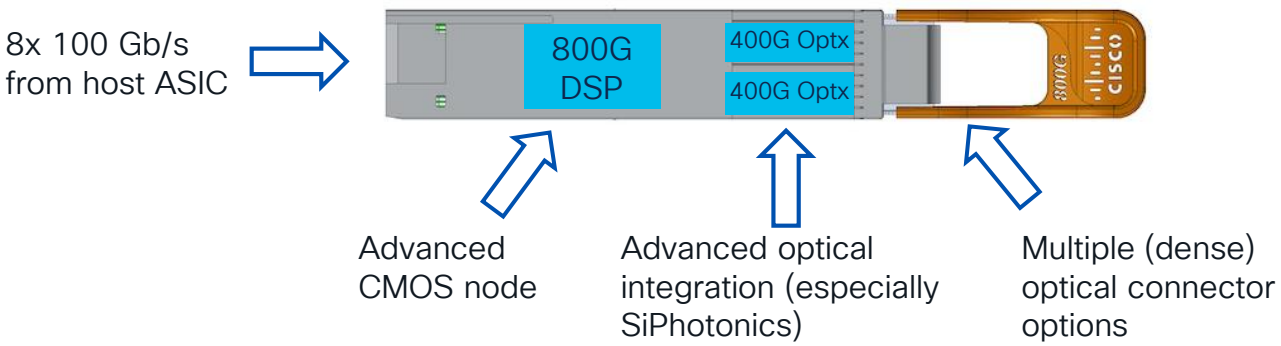
QSFP-DD800 supporting dense 400 GbE (aka breakout)

QSFP-DD

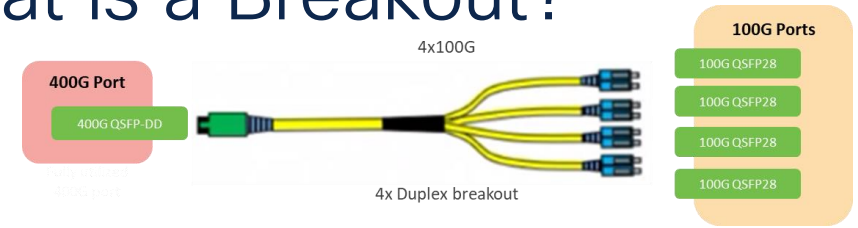
Both 400G & 800G form factor enables an economical way to implement breakout to lower speed Ethernet interfaces. This maximizes

- Cost and density on one end of link
- Compatibility and return on investment with existing equipment

Example: 800G module

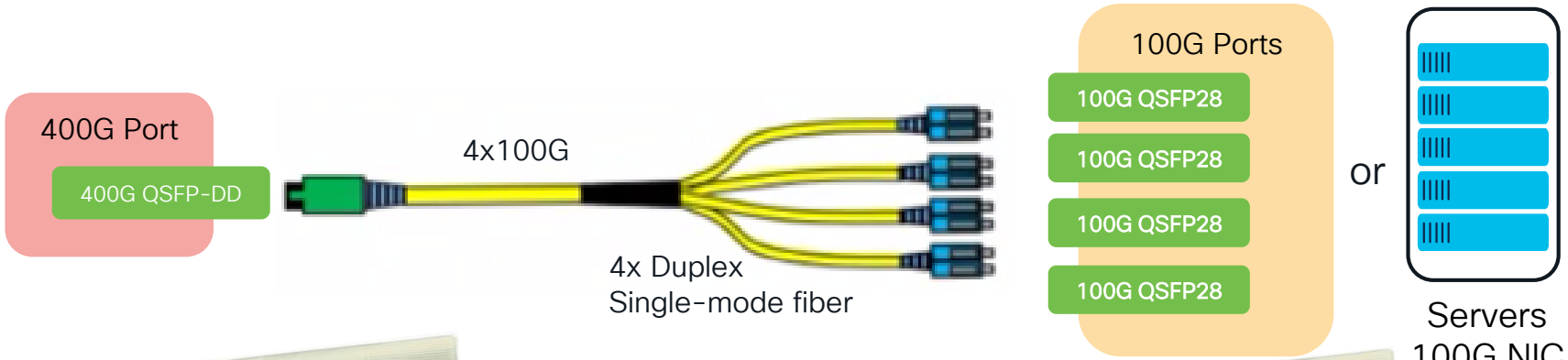


What is a Breakout?



- Breakouts take advantage of ports with multiple optical lanes for both the Tx and Rx
  - Optical lanes in this context means pairs of fibers
  - e.g. 400G DR4 optical connector has 4 pairs of fibers, each pair can be configured as a 100G-DR
- A breakout is when a single port is configured as multiple lower speed interfaces
- Breakout transceivers generally use MPO connectors which have multiple fibers for both the Tx and Rx
  - The port controls how the module will be configured either for breakout or non breakout operation
  - For 2x breakout, module's support regular LC connectors too
- Breakouts can also be done for copper cables and AOCs
  - Cables and AOCs are fixed for either breakout or non-breakout applications

Breakout enables operational efficiency



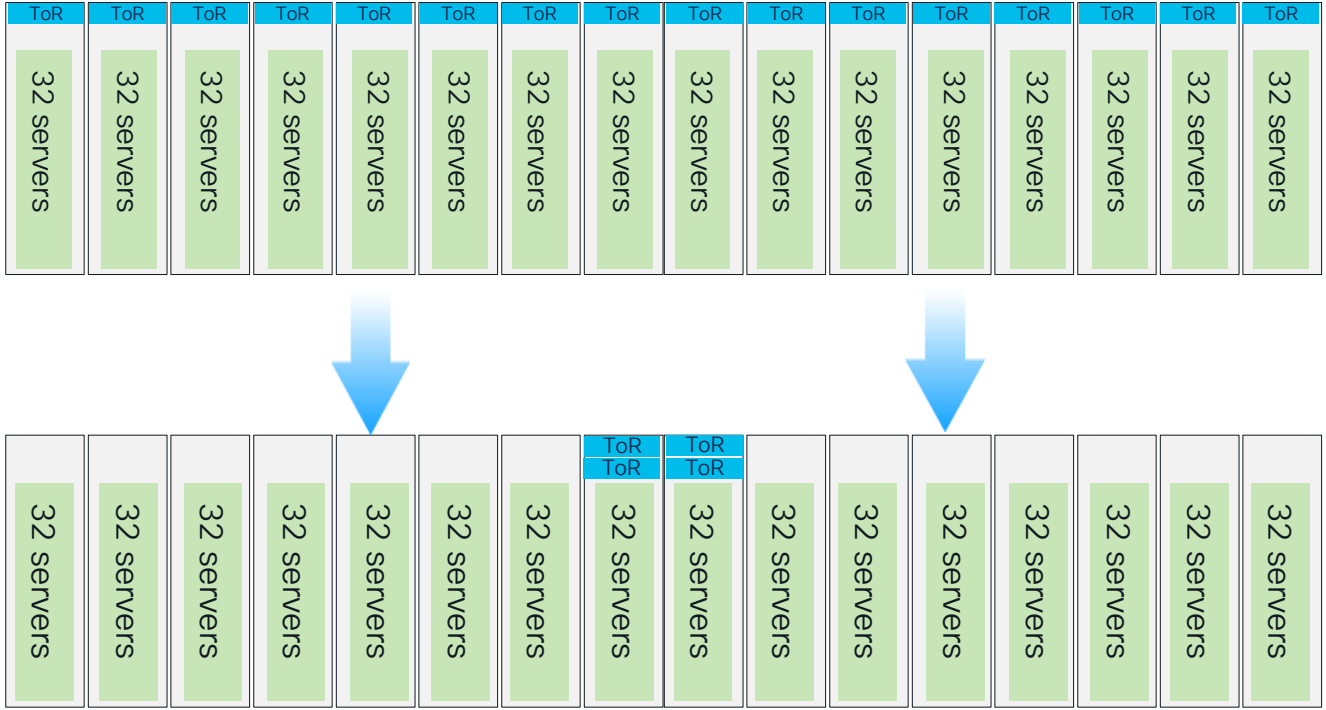
400G DR4 optics
Power: 12W
Space: 1 port

100G DR optics
Power: 4.3W x 4 = 17.3W
Space: 1 port x4 = 4 ports

Savings: 8 logical 100G ports, 5 Physical ports
→ 8 logical 100G ports 3 physical Ports, 5.3W power savings

Breakouts Promote Sustainability

32 1RU servers per cabinet w/ 16 cabinet row



Non Breakout Architecture
- 16 TOR = 811W x 16 = 12976W
- 512 DAC 0.5W per end = 512W
- Total power 13,744W

Breakout Architecture using 64 port switches
- 4x switch: 1324W x 4 = 5296W
- 128 DR4 modules 9W x 128 = 1152W
- 100G-DR modules 3W x 512=1536W
- Total Power 7984W

40% savings in power

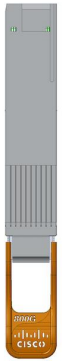
Implementing Dense 400 / 800 GbE

QSFP and QSFP-DD allow wide range of deployment options between switch and compute equipment with compatible modules

100 Gb/s
25 Gb/s SerDes
QSFP28

400 Gb/s
50 Gb/s SerDes
QSFP56-DD

400 Gb/s
100 Gb/s SerDes
QSFP112

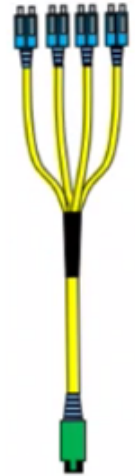


800 Gb/s
100 Gb/s SerDes
QSFP-DD800

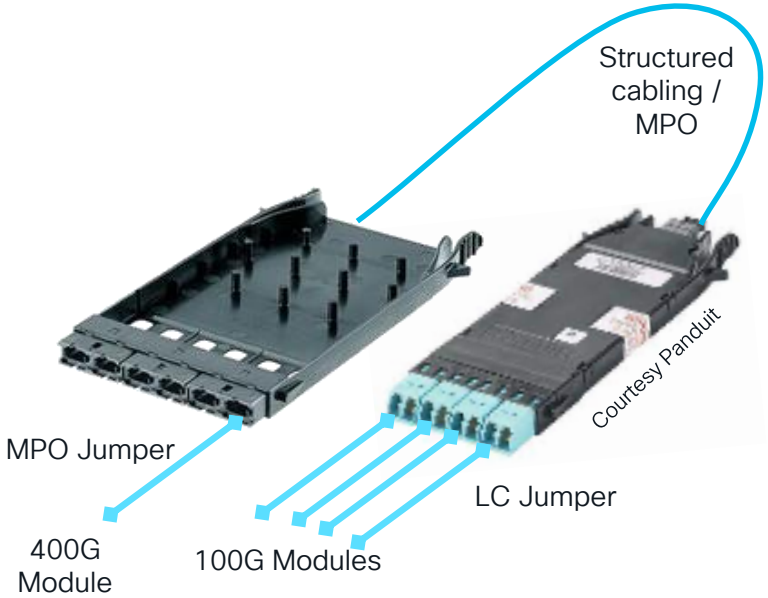
100 GbE	1x	4x	4x	8x
200 GbE	-	2x	2x	4x
400 GbE	-	1x	1x	2x
800 GbE	-	-	-	1x

Deploying Breakout

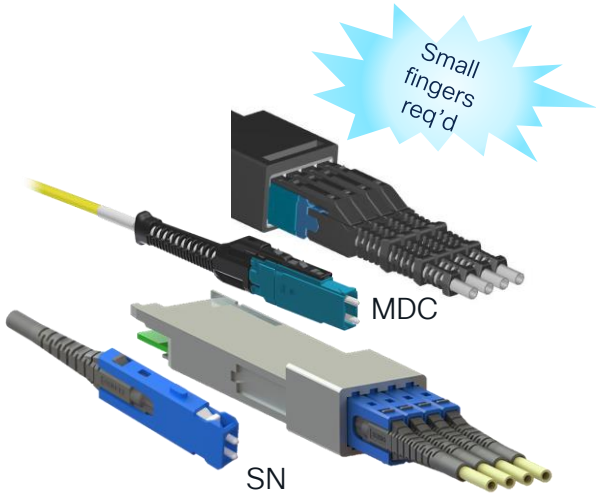
- Multiple options exist



Breakout cables



Using structured cabling and breakout cassettes



Dense VSFF connectors mostly used in patch panels

Network Operator Hurdles for Breakouts

- Reliability – What if one lane fails, do I have to interrupt non-affected traffic?
  - Nearly all systems design include redundancy or fail-over paths
  - Cisco optical modules have world class reliability
    - Typical field failure rate <<100ppm
  - Breakout modules do not have higher field failure rates than single lane modules

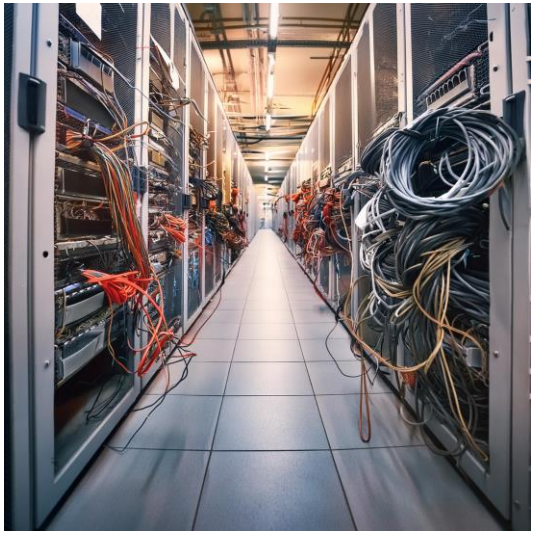
- Fiber management – Managing more connections in front of the switch will be a nightmare
  - Cisco sells a high-density patch panel system that eases fiber management
    - Supported by Cisco!

- Why transition to a higher BW switch now?
  - Manage migration to higher speed with the latest features on a port-by-port basis.
  - Ports are backwards compatible and can accommodate existing optics
  - Use breakout modules to connect to lower speed ports while network evolves
  - Use higher speed modules when needed as the rest of the network is upgraded
  - Reduce power consumption for equivalent bandwidth

Managing Breakout Cabling: Patch Panels

Recommended to use high density breakout patch panels

Can address any concerns about fiber management with breakouts



When you say “Breakout”, is this what you are thinking? (AI generated pic ☺)

Example using Cisco patch panels

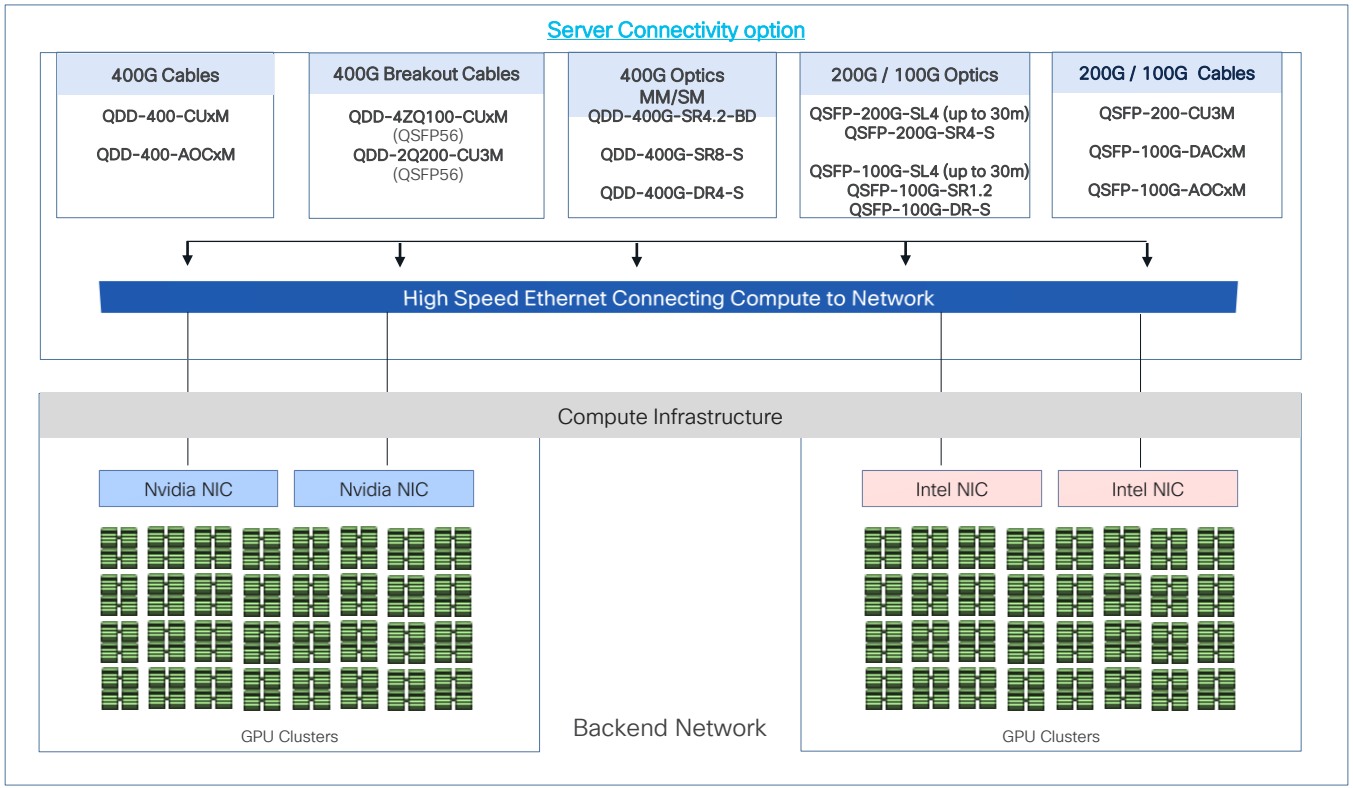
- 24 RU’s of patch panels
- 432 jumpers

Supporting :

- Cisco 8812 Chassis (172 Tbps)
- 12 Line Cards
- 36 Ports of 400G
- 4x100G Modules broken out
- 1728 x 100G connections

Maintain the same chassis and migrate each port to higher speed interfaces as needed

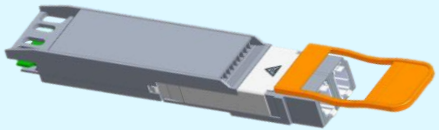
AI/ML Optics Connectivity Options for Compute



- A need for high-speed optics in compute to support high bandwidth GPU clusters
- Fewer servers per rack due to power and heat constraints (GPU)
- More inter-rack communication due to less servers in the rack
- Active Optical Cables (AOCs) and Pluggable Short Link (SL) optics for reaches of up to 30m
- 400G and 100G BIDI for reaches exceeding 30m
- Use of Direct Attach Cables (DAC) for smaller clusters

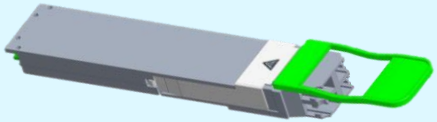
Pluggable form factors – a wealth of options

OSFP



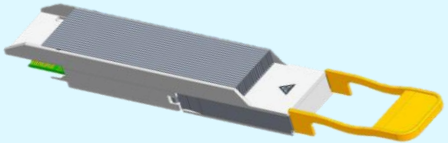
400G/800G/1.6T

OSFP-RHS



400G/800G/1.6T

OSFP-XD

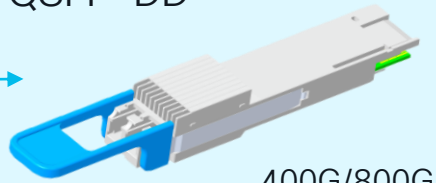


1.6T

Switches

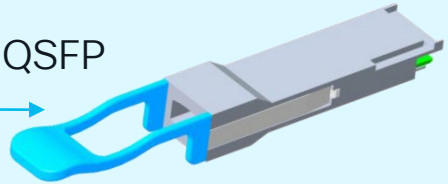
Compute

QSFP-DD



400G/800G/1.6T

QSFP



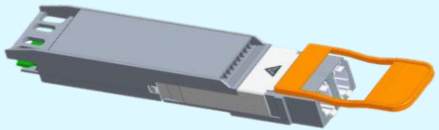
100G/200G/400G
QSFP28, QSFP56, QSFP112

Port compatibility

Everything within the square is compatible with each other

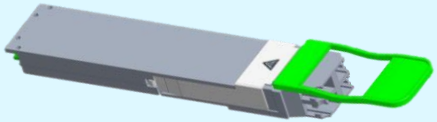
Pluggable form factors – interface interoperability

OSFP



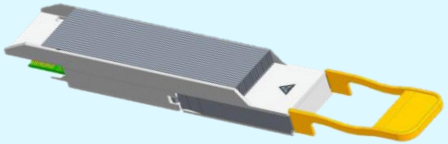
400G/800G/1.6T

OSFP-RHS



400G/800G/1.6T

OSFP-XD

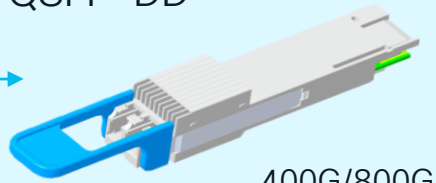


1.6T

Ethernet plug and play as normal.

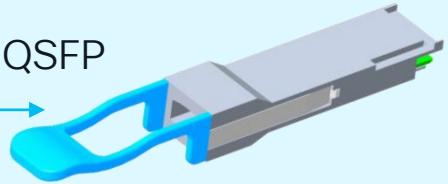
Breakout as well!

QSFP-DD



400G/800G/1.6T

QSFP

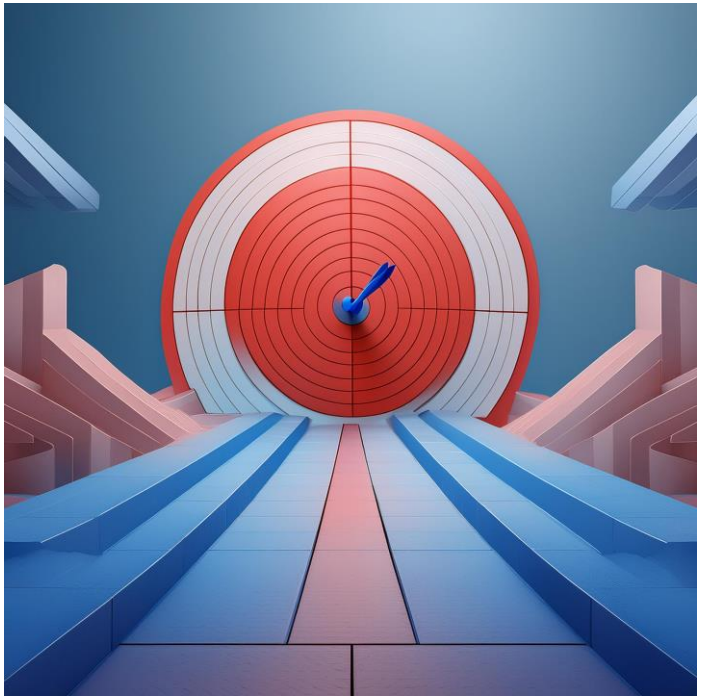


100G/200G/400G
QSFP28, QSFP56, QSFP112

Ethernet Interface compatibility

Port compatibility
Everything within the square is compatible with each other

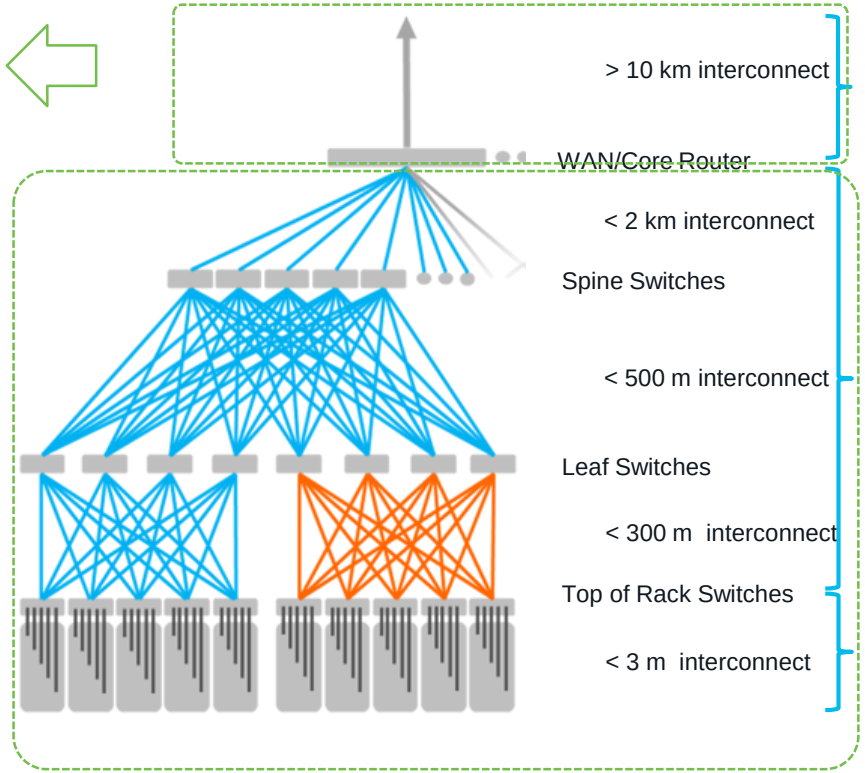
Key takeaways – client (IMDD) reaches



(AI generated pic)

- 400G & 800G Standards are mature. Ethernet interop is guaranteed regardless of form factor
- Optics and copper interconnect available in all reaches
- Cisco has a wide variety of high-density transceiver and cabling breakout solutions
- Sustainable and simplified deployment options are available

Optics deployment use cases



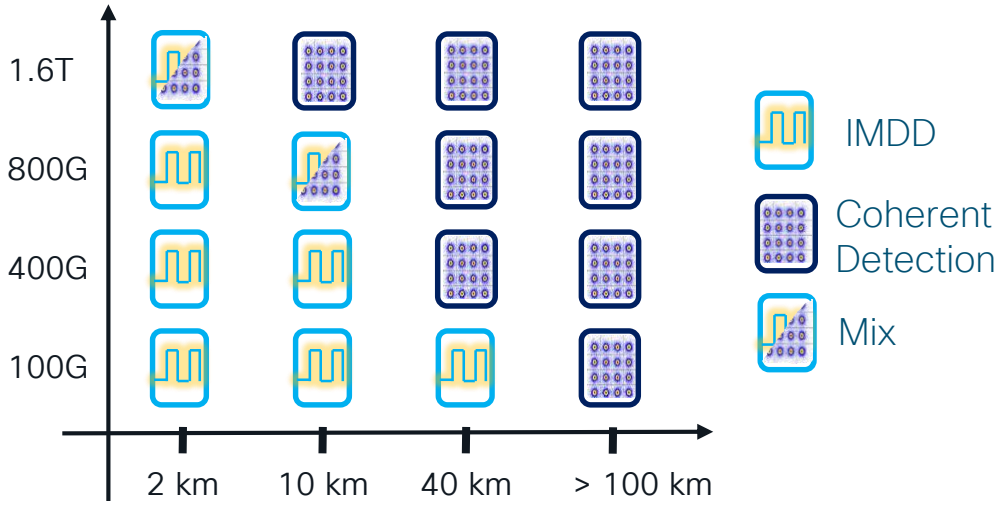
Inter DC Interconnect (Regional Networks)
- Mostly serviced by coherent optics

Intra DC Interconnect (Front End Data Center Networks)
- Shorter distances < 2km
- Primarily serviced by direct detect optics
- Highest data center volumes (millions of units per year)

Server racks
- Mostly direct attached copper

Coherent optics addresses higher reaches (and speeds)

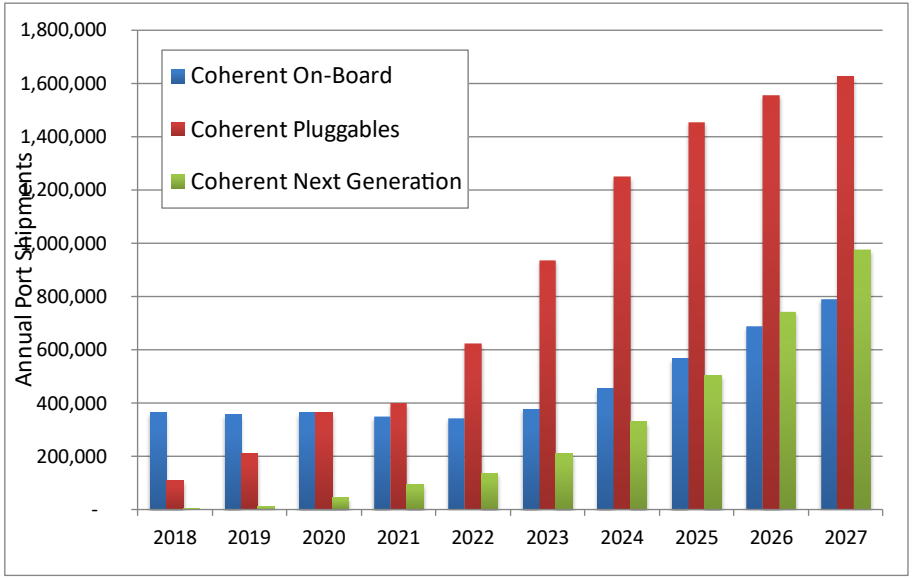
- IMDD and Coherent technology will continue to be used
  - 800G, 1.6T and beyond
  - Coherent pushing towards shorter reaches. Not only in DWDM but also Grey applications
  - Focus shifting from performance enhancements ($$$) to interoperable interfaces and pluggables
- Coherent multivendor Interop more and more prevalent
  - 100G, 400ZR/ZR+, 800ZR/ZR+



As speed increases, the crossover distance between IMDD & coherent decreases

Coherent Optics shifting towards pluggable

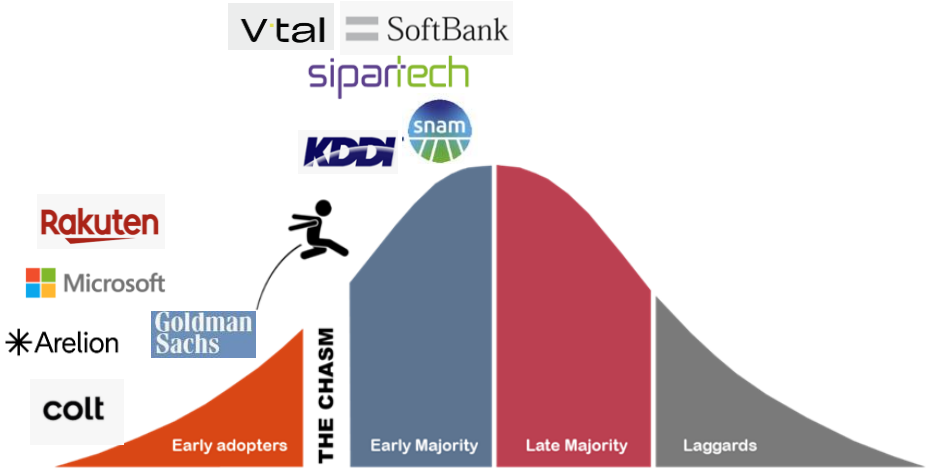
- Historically, embedded solutions have been a majority of coherent ports
- CFP2 gained share, but still a minority
- 400G pluggable has flipped the script
  - Support for router form factors (QSFP-DD)
  - Interoperability
  - Reduced performance gap



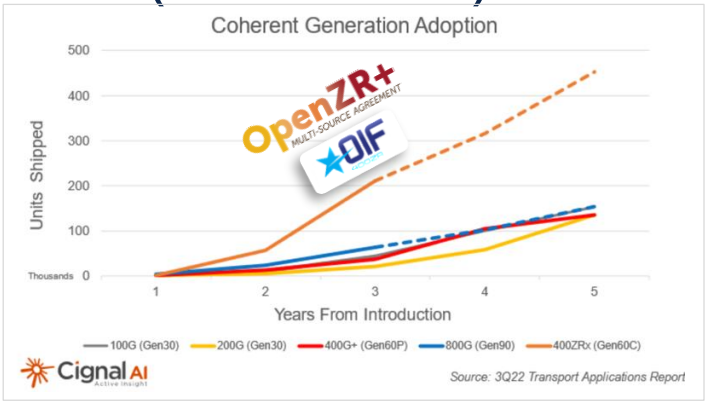
Going forward, Embedded designs will benefit from alignment with pluggable

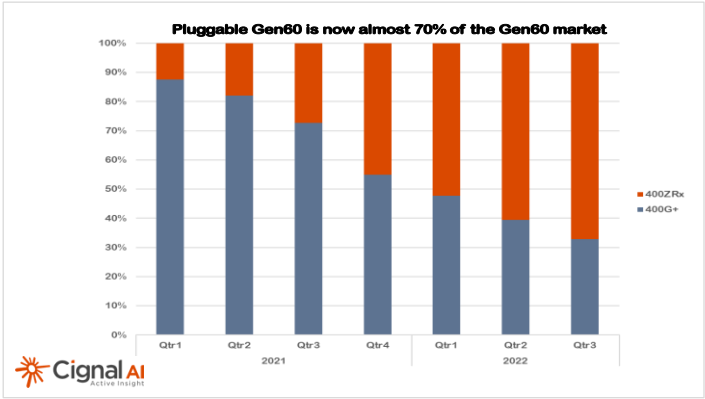
Market adoption: 400G Pluggable (aka DCO)

- 400G pluggable coherent optics (400ZRx) is the fastest adopted coherent technology
- 400G ZRx represents 70% of total deployed 400G coherent interface
- Open line system is accelerating the adoption of 400G DCOs as alien wavelengths without reengineering the photonics layer



Source: Cignal AI



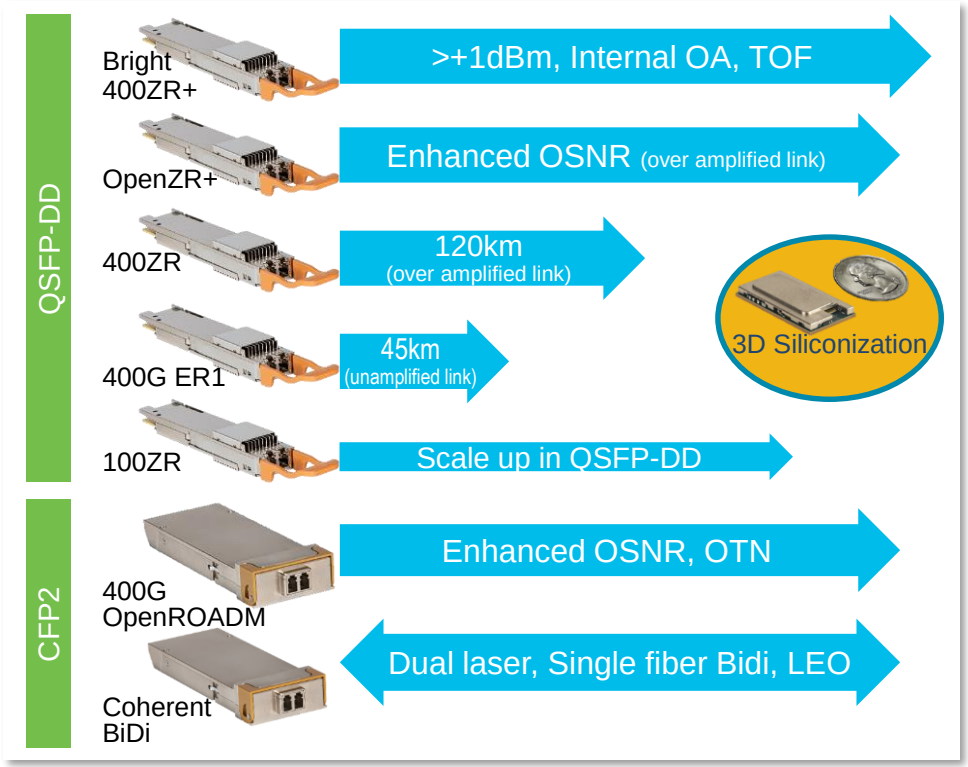


Source: 3Q22 Transport Application Report

Broad 400G MSA compliant pluggables portfolio

Addressing wide range of hyperscale, service provider and enterprise applications

- Anchored by industry standards--many driven by Acacia/Cisco
- Leading 400G pluggable deployments (>**1/4 million ports** shipped)
- Bright 400ZR+ ramping; >10k units shipped
  - 25M+ service hrs w/ no field failures



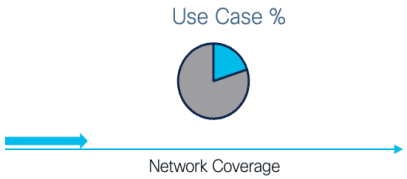
400G QSFP-DD Digital Coherent Optics Portfolio

ER1

Point to Point
Intra-data center, campus interconnect, core-to-edge router

Lowest Cost
Based on fixed laser with simple point-to-point connectivity

Short Reach
Up to 45KM for unamplified at 13dB

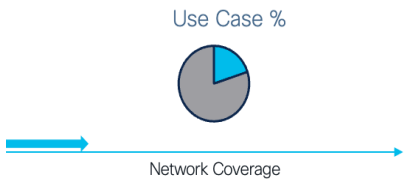


ZR

Point to Point
Web, Data Center Interconnect, Non-SP/SP router interconnect

Low Cost
Lowest cost 400G DCO option for very simple designs

Short Reach
Up to 120KM for P2P amplified links

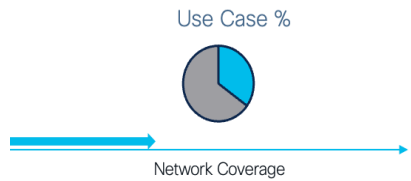


ZR+

Data Centric
Web Scale, DC Interconnect, Non-SP/SP Router Interconnect

Cost Optimized
Essential power and features only to optimize for cost

Simple Features
Designed for open line systems that balance power levels; high-performance forward error correction

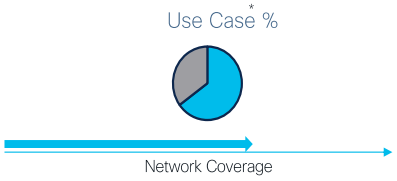


Bright ZR+

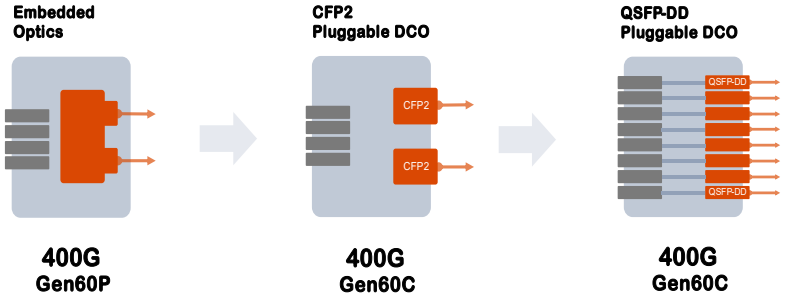
Transport Centric
Service Providers, Routed Optical Networking

High TX Power
+1dbm for difficult spans; interop with brownfield transponder & legacy line systems

Advanced Features
TOF, OTN and L1 Encryption features; integrated optical amplifier

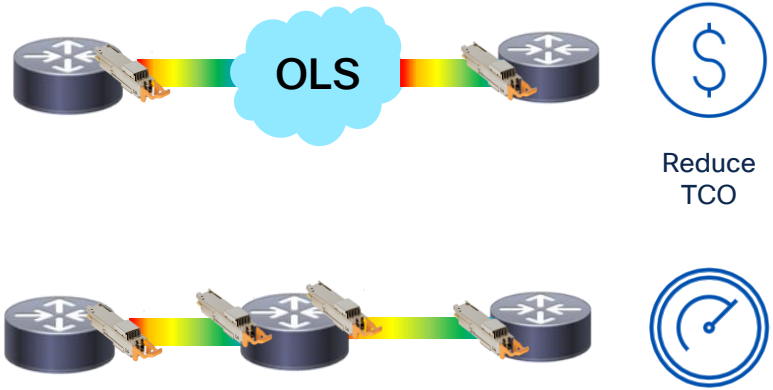


Pluggable coherent – Enabling efficient network architectures



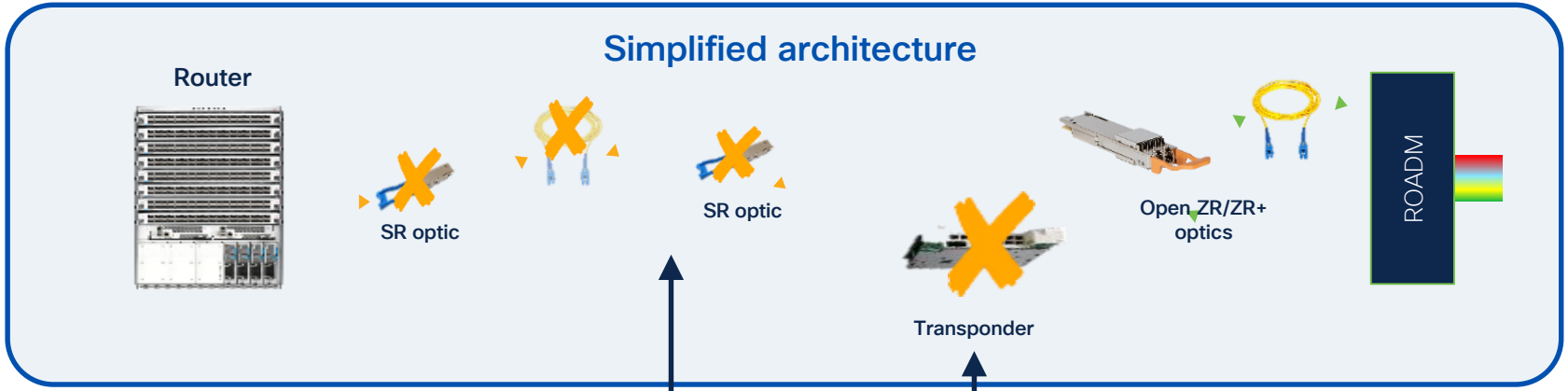
Density is improving many folds; power per bit coming down dramatically* ….

Routed Optical Networking (RON)



*Source: Cignal AI in PONC 2023

Coherent optics enables savings – power, space cooling

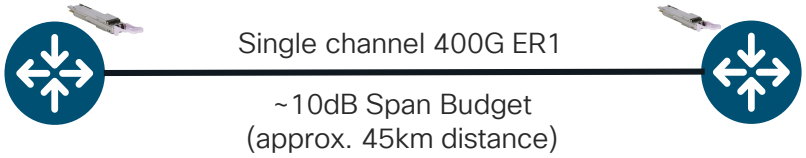


- Eliminate transponder and associated shelf/chassis
- Eliminate client optics

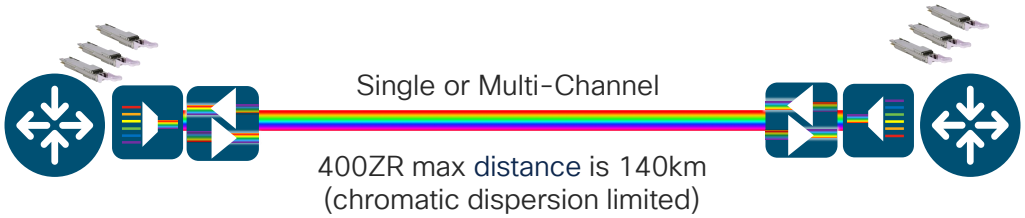
"Simplicity is the ultimate sophistication."

Leonardo da Vinci

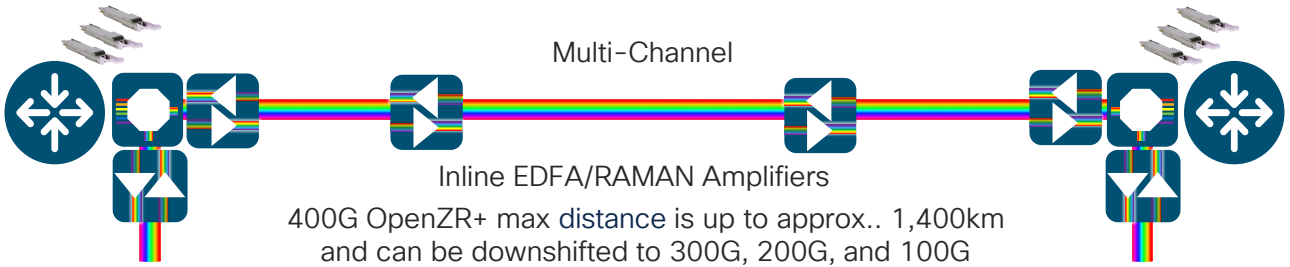
DCO Applications



Unamplified, Dark Fiber



Amplified, Point-to-point DWDM

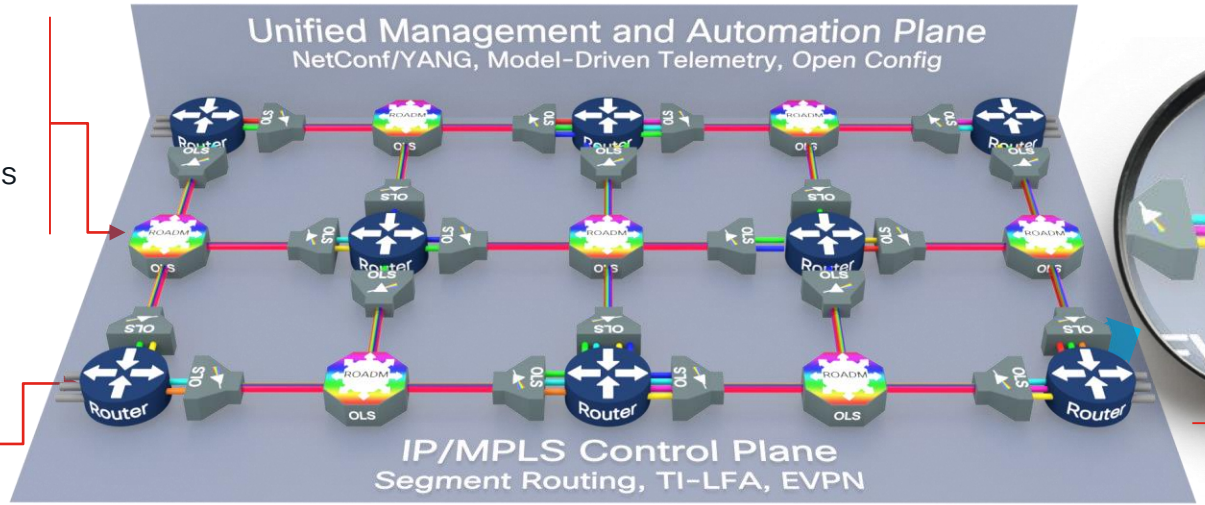


Full featured DWDM
(ROADMs, ILAs)

Coherent pluggable enabling Routed Optical Networking

400/800ZR/ZR+
QSFP-DD DCO
Pluggable Optics

75GHz min. width
DWDM: ROADM,
FOADM or Terminals



L3 VPN
L2 VPN
TDM
(via L1 PLE*)

400/800GbE Capable
Router: Modular, Fixed

*PLE: Private Line Emulation

Routed Optical Network (RON)

Use cases and deployment models

TCO*	46%
CapEx*	35%
OpEx*	57%

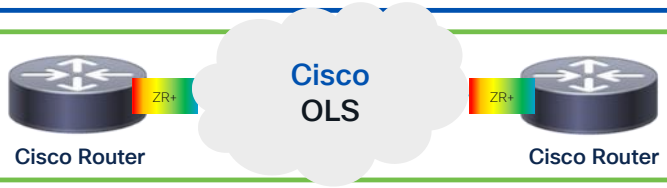
* ACG Research 2023

Different use cases and deployment models

1 QDD-ZR/ZR+ over third-party OLS

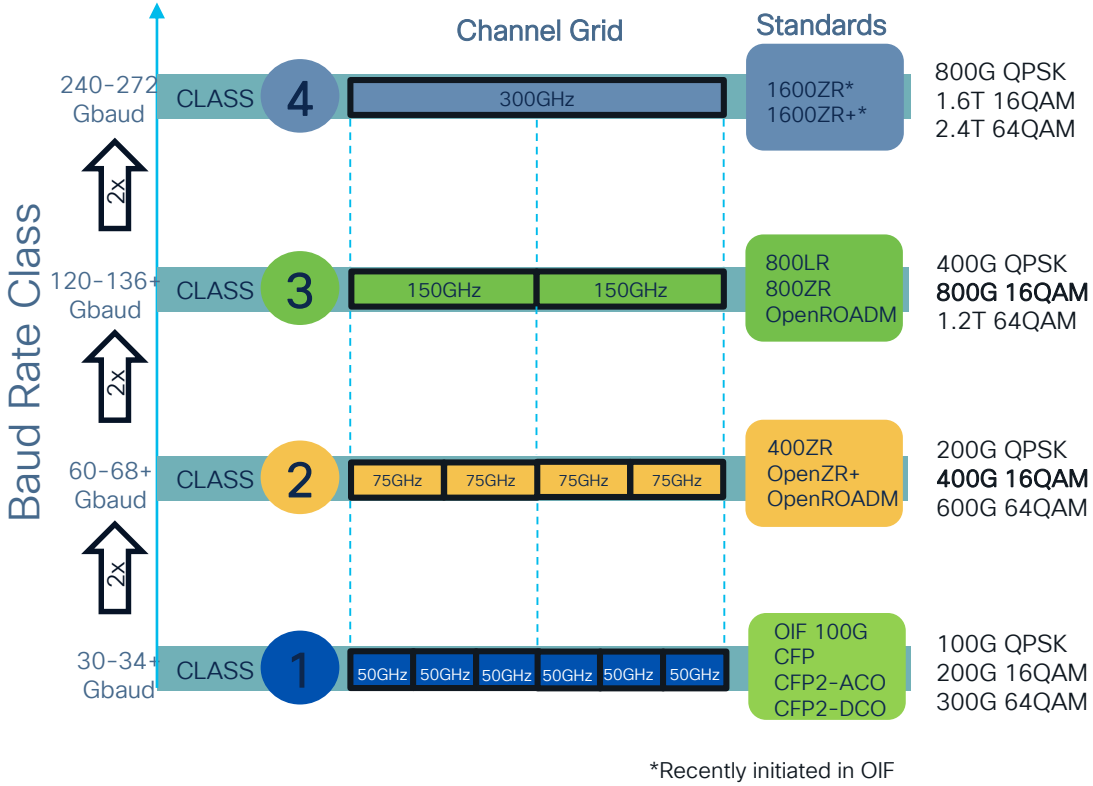
2 QDD-ZR over dark fiber or passive DWDM

3 QDD-ZR+ over Cisco OLS



>75% of the RON deployments are over third-party line system!

Evolutions of coherent transceivers



*Recently initiated in OIF

- Doubling the baud rate has always been lowest cost, lowest power solution for doubling data rate.
- Faster baud rate = wider (& fewer) channels
  - 400G (Class 2) utilized 75 GHz channel spacings
  - 800G will move to 150 GHz spacings
  - 1.6T anticipated to move to 300 GHz spacings

Key features needed for 800G Coherent module

Multi-haul capable
- 2-4 bits/symbol

High Tx Optical Pwr capable

Client Electrical Traffic
- 800 GbE support
- 100/200/400 GbE muxing support

- Low power consumption
- CMIS compliance

Transmission modes
- OIF 800ZR
- **800 ZR+** with Interoperable **PCS**

800ZR Class 3 ~120+Gbaud 800G Standard Interop
800 ZR+ Class 3 ~ 130+Gbaud 800G higher OSNR Interop PCS Transmission

Next-Gen Coherent Optics

Routed Optical Networking (RON) building blocks

Module Portfolio

- 800ZR in QSFP-DD/OSFP. - 120km
- 800ZR+ in QSFP-DD/OSFP/CFP2
- 400G Long-haul in QSFP-DD and CFP2

Low Power 4nm DSP – Delphi

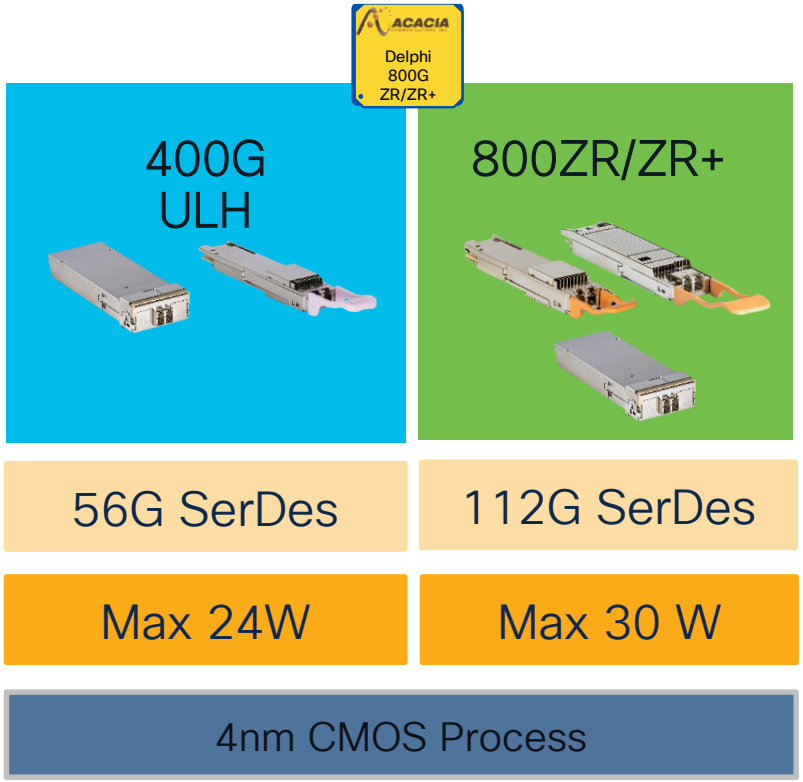
Up to 130Gbd Optics

400G ULH Pluggables:

- Designed for existing 400G platforms
- Integrated EDFA and TOF
- Flexible line side modes with PCS

800ZR/ZR+ Pluggables:

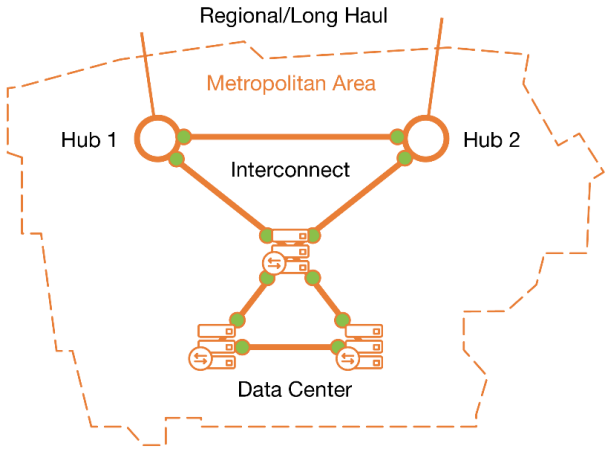
- Comply to OIF 800ZR
- Comply to 800ZR+ OpenROADM



Coherent 800G Enables Next Gen Speed and Distance Migration

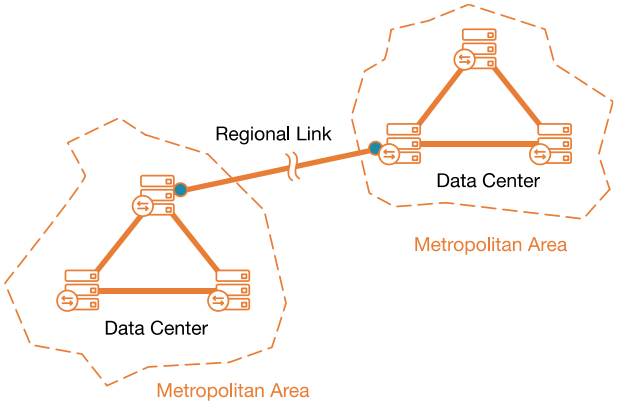
400ZR to 800ZR

- Migrate P2P DCI/Metro architecture to 800G



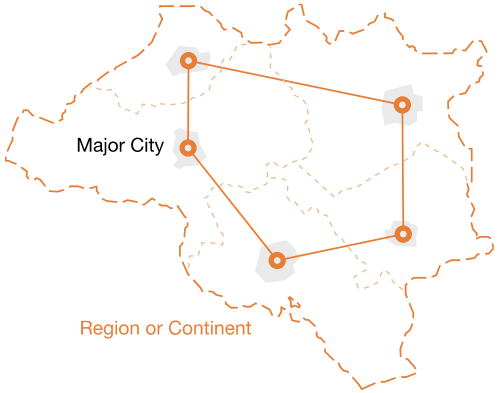
400G ZR+ to 800G ZR+

- Migrate regional links to 800G



400G Ultra Long Haul*

Adds ULH capability to 400G pluggables portfolio



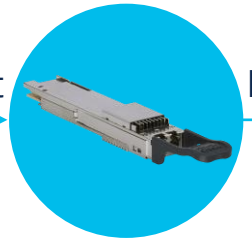
* 800G modules will have advanced 400G modes

Coherent Standardization @ 400G

400G Pluggable Coherent – 100+ km to 1000+ km

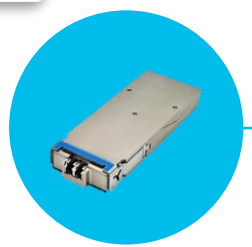
OIF 400ZR

Simplified functionality
Edge optimized FEC (C-FEC)

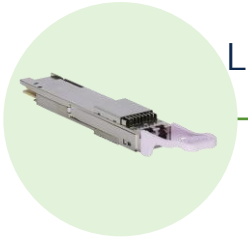
400GE → Client →  → Line → *400ZR – 120km*

Open ROADM

Flexible Client Mapping +
High Performance FEC (oFEC)

100G
200G
300G
400G
Ethernet
OTN
FlexO
→ Client →  → Line → *100G*
200G
300G
400G
OpenROADM

OpenZR+ MULTI-SOURCE AGREEMENT

100GE
200GE
400GE
Nx 100GE
(Muxponder Mode)
→ Client →  → Line → *100G*
200G
300G
400G

Simplified functionality +
High performance FEC (oFEC)

High performance pluggable modules
Multi-vendor interoperability

Cisco Blog: OpenZR+ Delivers 400G Coherent Optics for Multi-Haul Networks

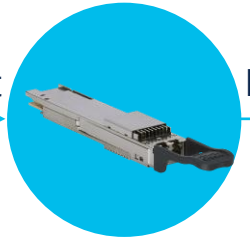
Coherent Standardization @ 800G

800G Pluggable Coherent – 10+ km to 1000+ km

OIF 400ZR

Simplified functionality
Edge optimized FEC (C-FEC)

800GE
2x400GE

Client →  → Line

800ZR – 120km

Open ROADM

Flexible Client Mapping +
High Performance FEC (oFEC)

100G
200G
300G
400G
Ethernet
OTN
FlexO

Client →  → Line

100G
200G
300G
400G
600G/ 800G
OpenROADM

Interoperable PCS

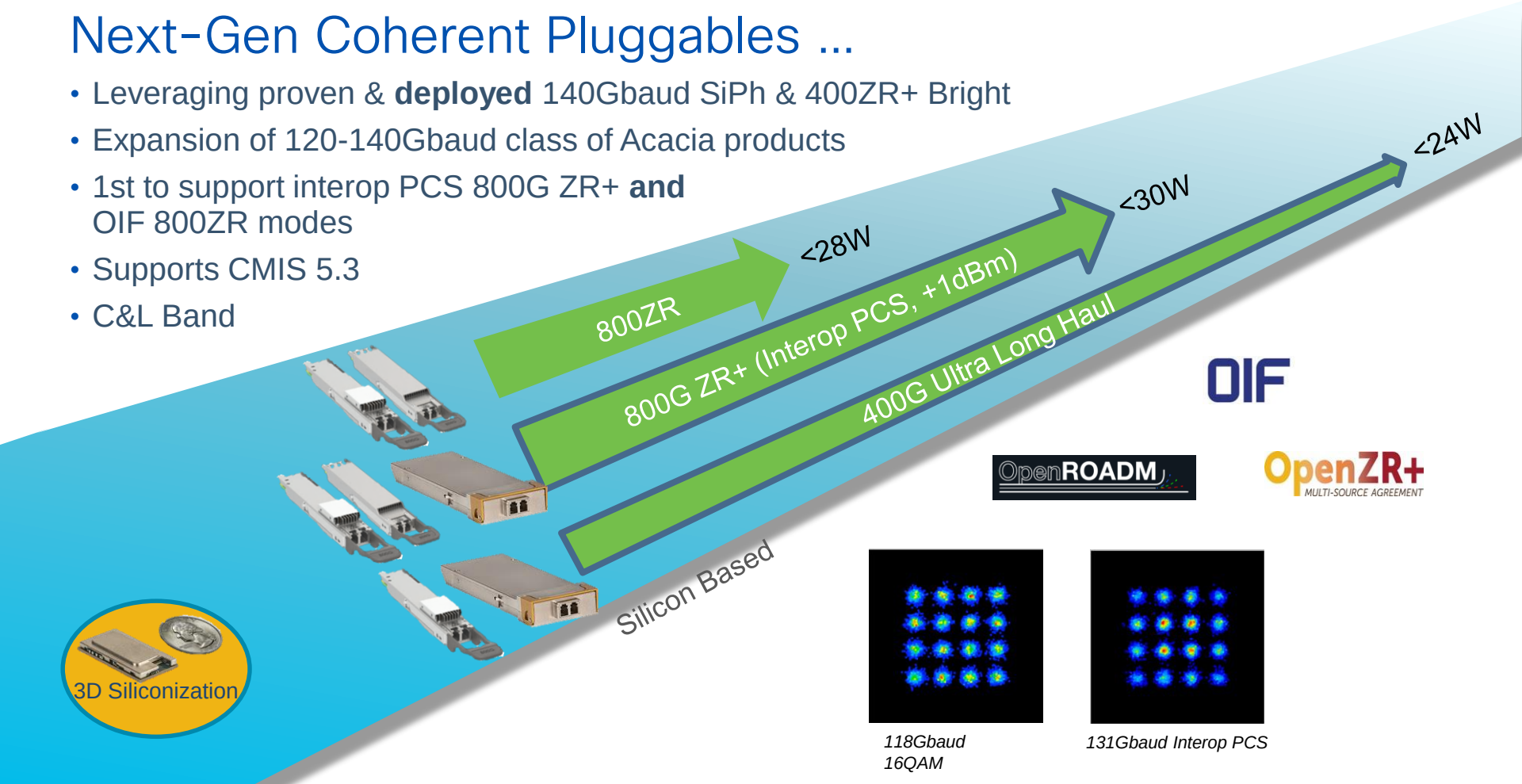
Open ROADM is taking the industry lead for 800ZR+

Cisco's Delphi-based solutions will be industry first compliant Open ROADM solution

First definition of interoperable Probabilistic Constellation Shaping (PCS) – advanced technique gives additional performance improvement

Next-Gen Coherent Pluggables …

- Leveraging proven & **deployed** 140Gbaud SiPh & 400ZR+ Bright
- Expansion of 120-140Gbaud class of Acacia products
- 1st to support interop PCS 800G ZR+ **and** OIF 800ZR modes
- Supports CMIS 5.3
- C&L Band



OIF

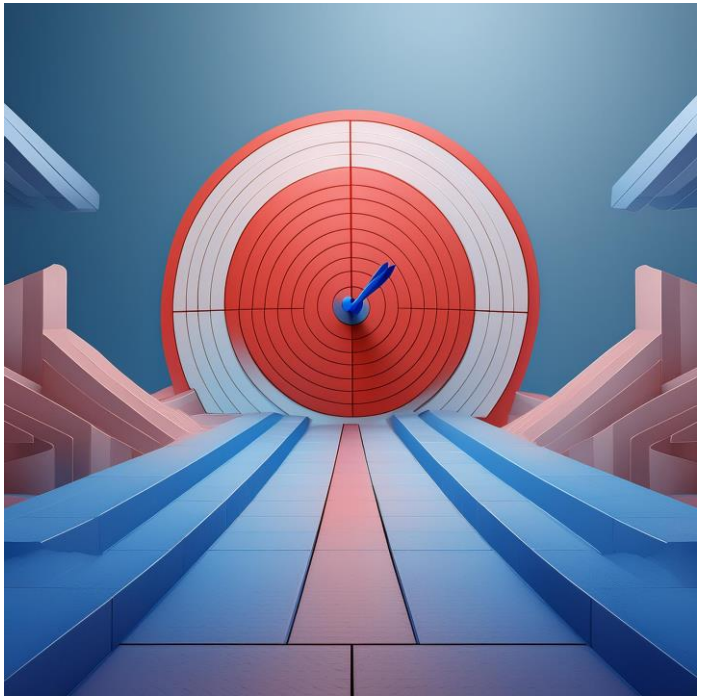
OpenROADM

OpenZR+ MULTI-SOURCE AGREEMENT

118Gbaud 16QAM

131Gbaud Interop PCS

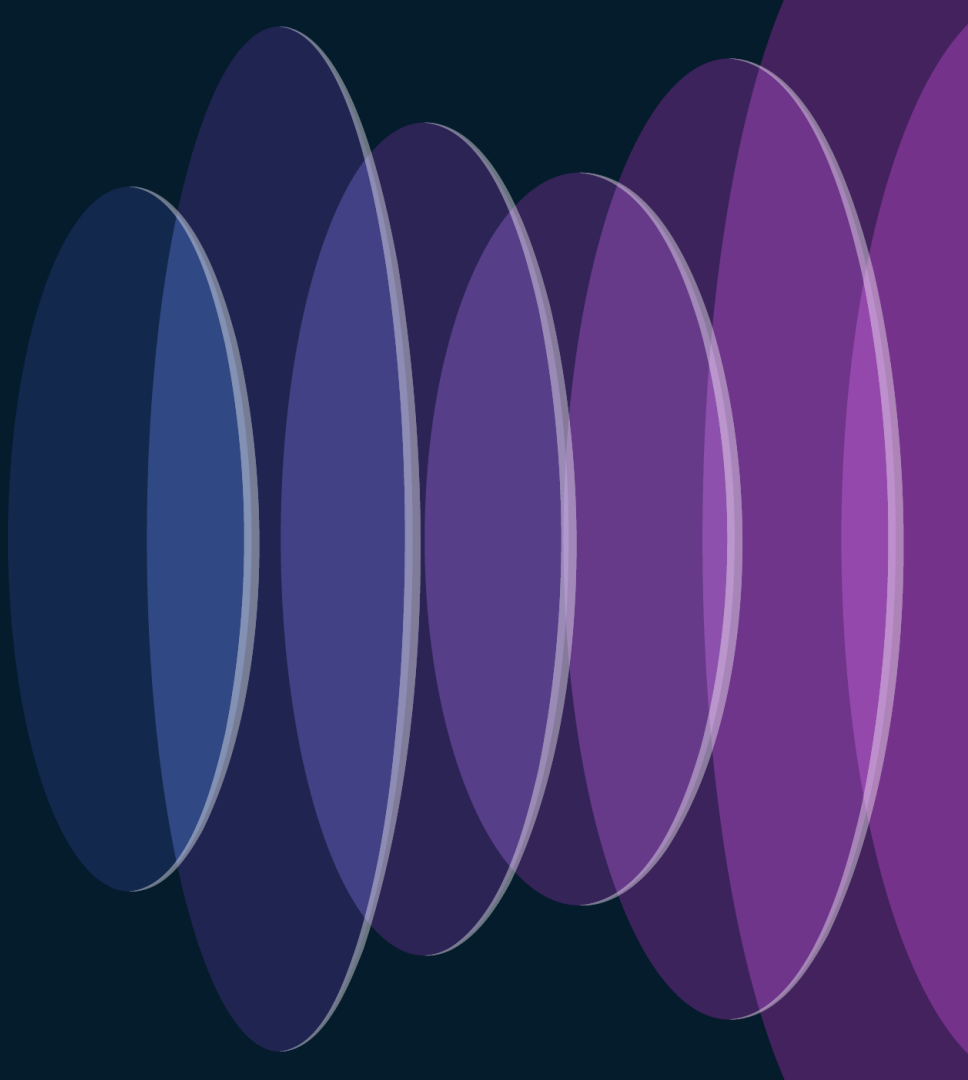
Key takeaways – coherent reaches



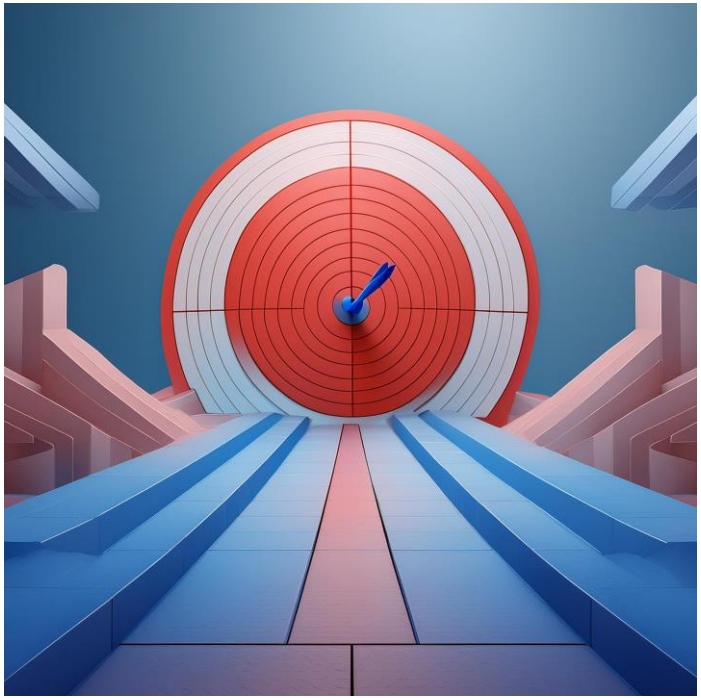
(AI generated pic)

- Pluggable solutions are leading option for coherent interfaces
- 400G coherent solutions (and standards) are mature with lots of deployment options
- 800G coherent is coming soon – interoperable PCS standardization improves performance
- Routed optical networking: improved sustainability, operational efficiency & cost efficiency

Wrap up



Key takeaways – thank you for your attention



(AI generated pic)

Optics and Interconnect are responding to demands from AI/ML growth

New speeds: 800G, 1.6T
New approaches: Linear pluggable (LPO), Co-packaged optics (CPO), Liquid cooling

IMDD: 400G is mature & 1st gen 800G uses same tech. Breakout solutions offer efficiency

Coherent: 400G is mature, 800G available soon. Pluggable & RON adoption are growing

Complete Your Session Evaluations

Complete a minimum of 4 session surveys and the Overall Event Survey to be entered in a drawing to **win 1 of 5 full conference passes** to Cisco Live 2025.

Earn 100 points per survey completed and compete on the Cisco Live Challenge leaderboard.

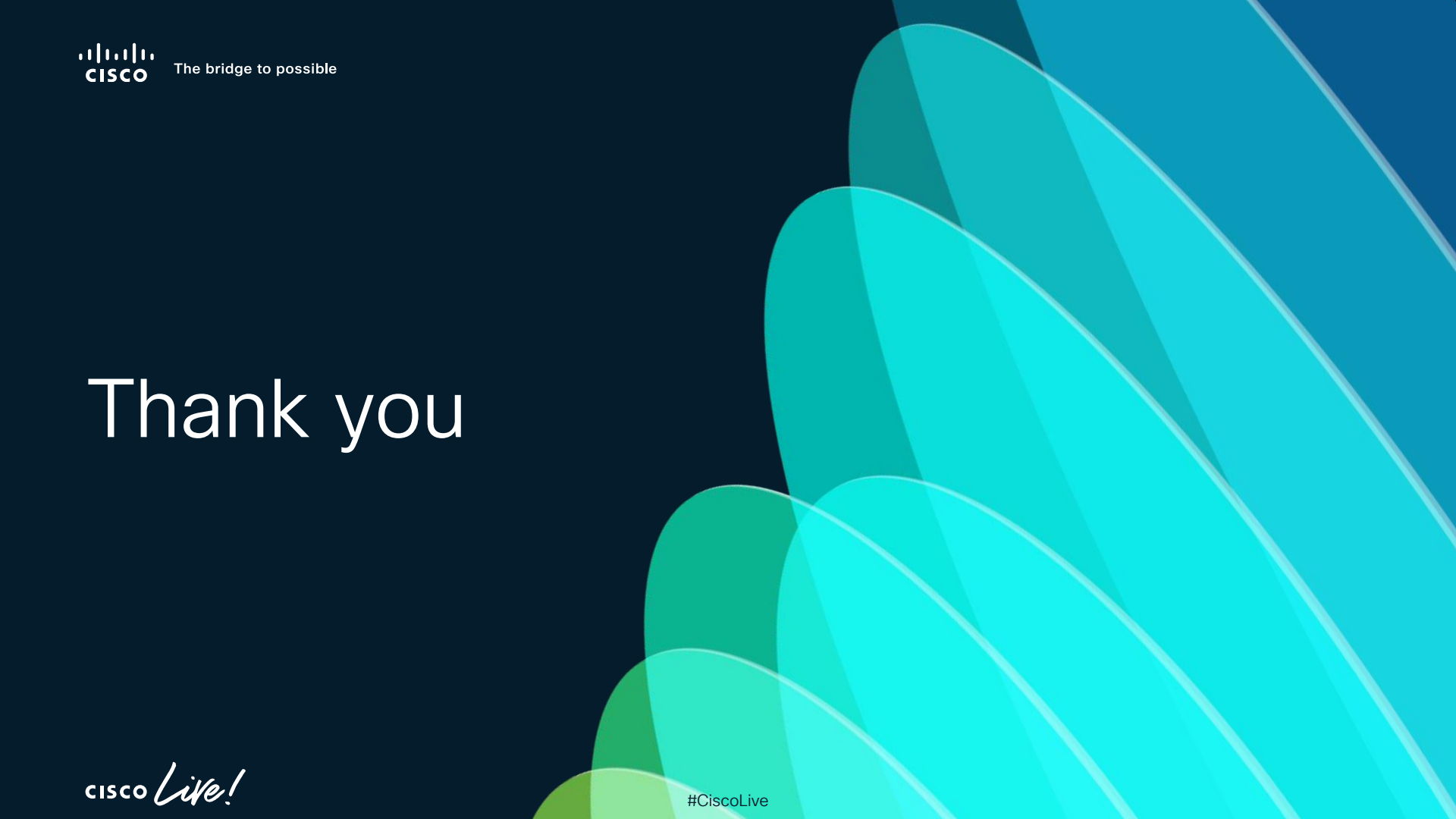
Level up and earn **exclusive prizes!**

Complete your surveys in the **Cisco Live mobile app.**

Continue your education

- Visit the Cisco Showcase for related demos
- Book your one-on-one Meet the Engineer meeting
- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs
- Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

Contact us at: Webex Space: "BRKOPT-2699: 400G, 800G, and Terabit Pluggable Optics: What You Need to Know"


The bridge to possible
Thank you
cisco Live!
#CiscoLive