# Let's go cisco live! #CiscoLive



## 400G to Terabit Optics

What you need to know!

Mark Nowell, Cisco Fellow Errol Roberts, Distinguished Architect

BRKOPT-2699



## Cisco Webex App

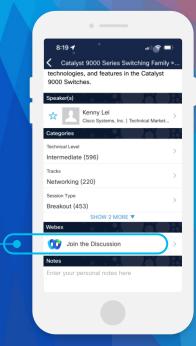
#### Questions?

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 9, 2023.



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



## Agenda

- The current state of 400G market adoption
- Successfully deploying 400G
  - Considerations and challenges
- Going beyond 400G
  - New Implementations
  - New Technology
- Where are the standards going?
- Conclusion

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



## Market dynamics





## Network operator top of mind



Increasing capacity and sustainability



Preserve investments in existing optics infrastructure and cabling



Simplify operations and management of optical links

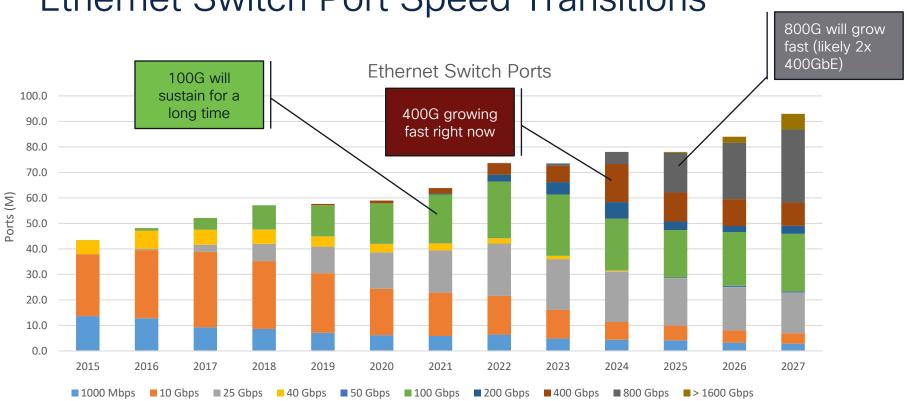


Preparing for capacity expansion





**Ethernet Switch Port Speed Transitions** 



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



BRKOPT-2699

## Market segments adopting higher speeds



#### Hyperscalers

100G/400G/800G fabrics
Al/ML compute clusters
Disaggregation



#### Webscalers

Scale-out fabrics
25/50/100G server NICs
Vendor NOS supporting open,
API-based automation



400G/800G



#### **Enterprise**

High performance IO AI/ML compute clusters Automation/ Monitoring



#### Media providers

Fabric for Media (IPFM)

8K uncompressed video driving
100G endpoints

Need for 400G uplinks



#### Telco service providers

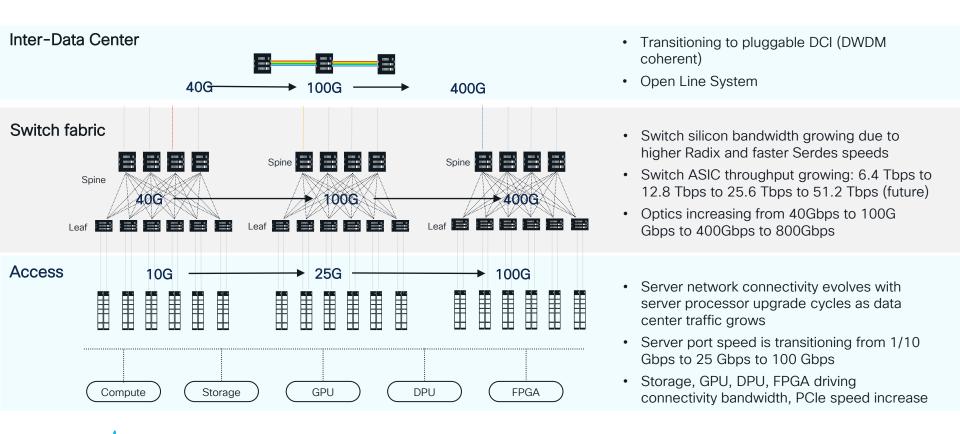
100G/400G fabrics

Space constrained SP DC and edge locations

Ready for NFV/5G adoption cycle

cisco life!

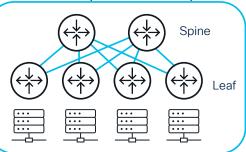
## Speed evolution in the data center





## 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)



Up to 87%Energy Savings

83% less space/fans

100 Gb/s ASIC IO (SerDes) 32 ports of 800G (64 ports of 400 GbE 256ports of 100 GbE)

~400 Watts 3,504 kWh/year



## Current state of 400G optics





## Cisco and 400 GbE Industry Activities

Standards	IEEE 802.3bs	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 <b>✓</b> / 802.3cw disco	400 GbE Coherent 120km / 400 GbE Coherent 80km
	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*	100G Lambda MSA 🔽	100G-FR, 100G-LR, 400G-FR4, 400G-LR4
	QSFP-DD MSA	400G/800G/1.6T Form factor
	OSFP MSA Illulia	400G/800G/1.6T Form factor
	SFP-DD MSA ✓	100G Form factor
	DSFP MSA.	Alternative 100G Form Factor (Mobile)
	400G-BiDi MSA	400 GbE MMF BiDi

<sup>\*</sup> Multi-Source Agreements - new ones all the time. Not all get wide industry adoption

## Flexibility of 400G pluggable modules

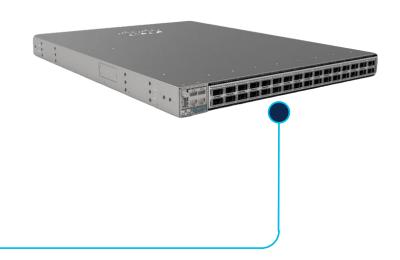
Copper cables

Multimode Fiber – 100m

Single Mode Fiber inside DC – 500m & 2km

Single Mode Fiber Campus – 10 km

Outside plant, DCI – 100-1000 km





BRKOPT-2699

## 400 GbE modules and use cases Full portfolio of interfaces available for use

Distance					
Dist	3+ m	100 m	500m-2km	10 km	100+ km
Optics	400G-CR8 8x 50G-CR 400G-AOC(30m)	400G-SR8 400G-SR4.2 400G-DR4	400G-DR4 400G-FR4 4x100G-FR	400G-LR4 4x100G-LR	400ZR 400ZR+
Media	Copper Cables / AOC (Active Optical Cable)	MMF / SMF	SMF	SMF	SMF



## What innovations did 400G bring?

New pluggable required to support 400G ports (8-wide)

New Pluggables (QSFP-DD)

New Modulation: PAM4 (& FEC) Higher speed interfaces adopted PAM4 modulation. Ubiquitous use of FEC.

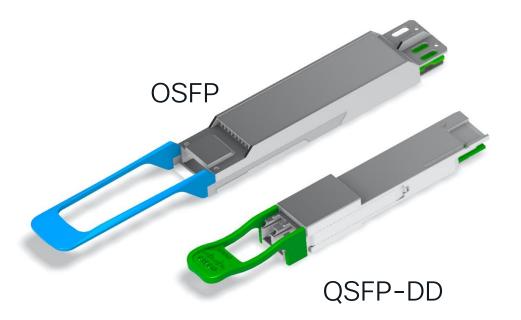
Long reach
coherent without
any system port
density reduction
→ Routed Optical
Networking

Pluggable Coherent: 400ZR/ZR+ Adoption (stds) of Breakout

Pluggable modules supporting multiple lower speed interfaces



## 400G Optical Modules: QSFP-DD or OSFP

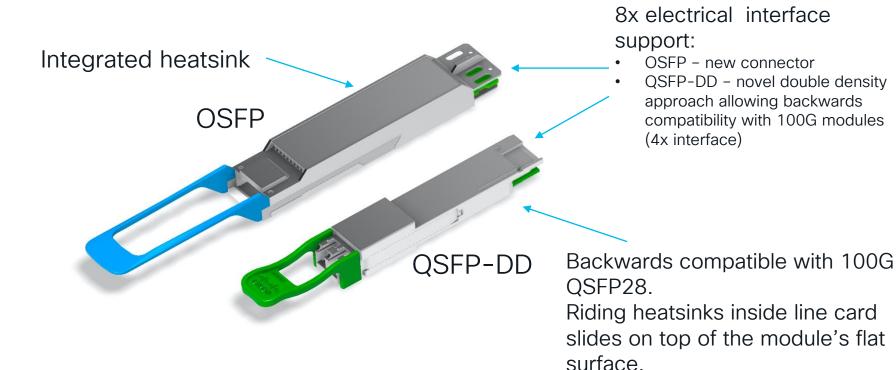


Initiated by Cisco, QSFP-DD was proven to address all the technical and market requirements for a successful 400 GbE roll-out.

QSFP-DD is supported by every system vendor and module vendor. Drives economies of scale.

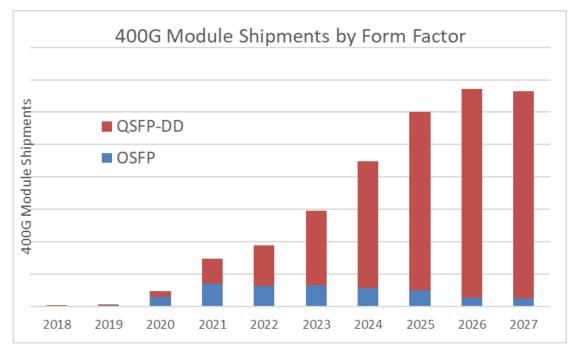
Ultimately, both modules meet all technical requirements (power, cooling, signal integrity). Differences are in alignment to market needs.

## 400G Optical Modules: QSFP-DD or OSFP





## Market share and forecast: QSFP-DD vs OSFP



LightCounting Mega Datacenter Report Database , Aug 2'22

#### **QSFP-DD Dominates 400G**

Two phases in module adoption:

- 1. Early (Hyperscaler) adoption
- 2. Broad market adoption

QSFP-DD's backwards compatibility is the key factor to explain wide adoption



BRKOPT-2699

## Technologies used for 400G optics

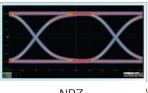
Enabling higher performance but lowering cost

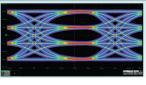
#### Modulation

Client optics

Coherent

optics





NRZ 1 bit/sec/symbol



Coherent QPSK 2 bit/sec/symbol

PAM4 2 bit/sec/symbol



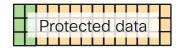
Coherent 16-QAM 4 bit/sec/symbol

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

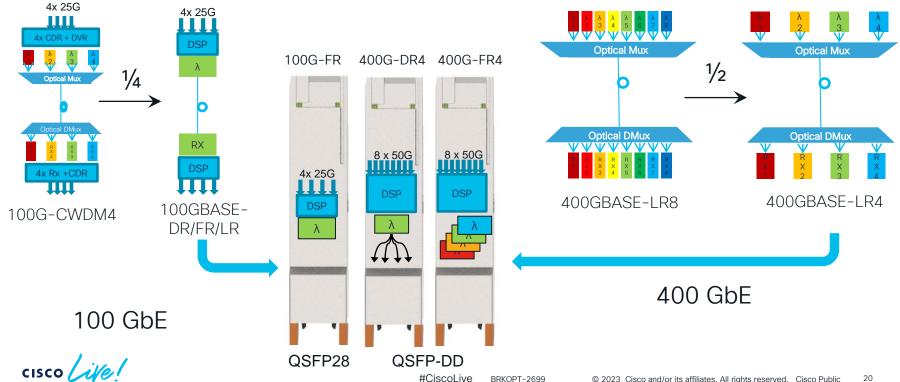
Enables use of relaxed specs (saves \$) to get same performance or enables much higher performance.

Usually embedded in Ethernet switch ASIC



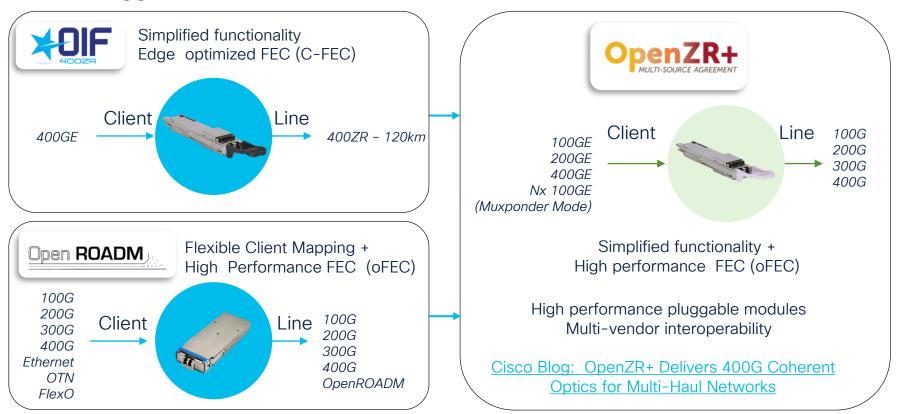
## 100 Gb/s per wavelength optics

Simplifying 100 GbE and 400 GbE



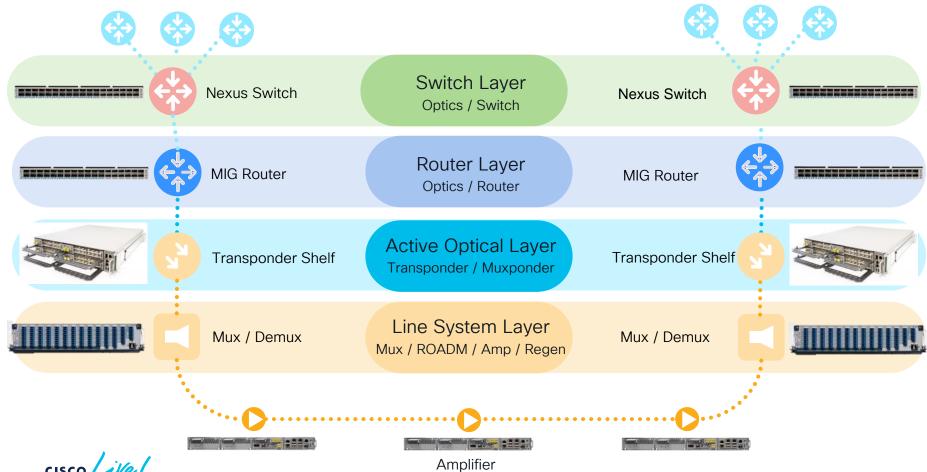
### Standardization

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

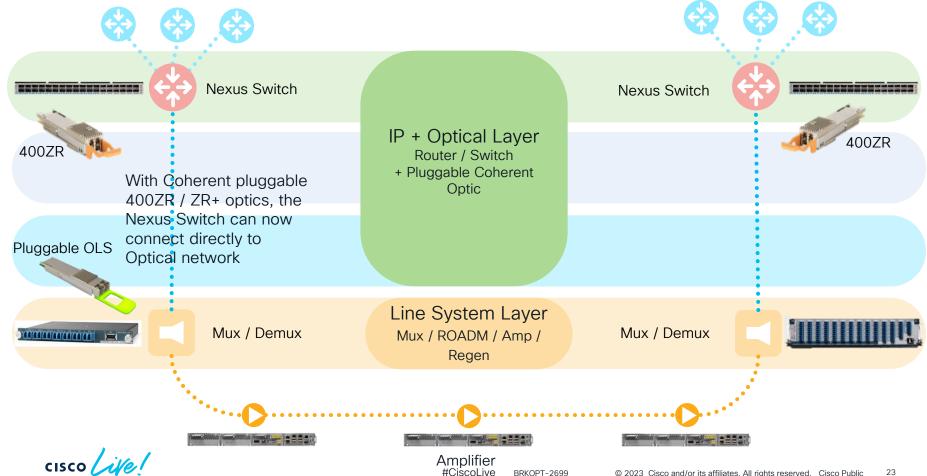




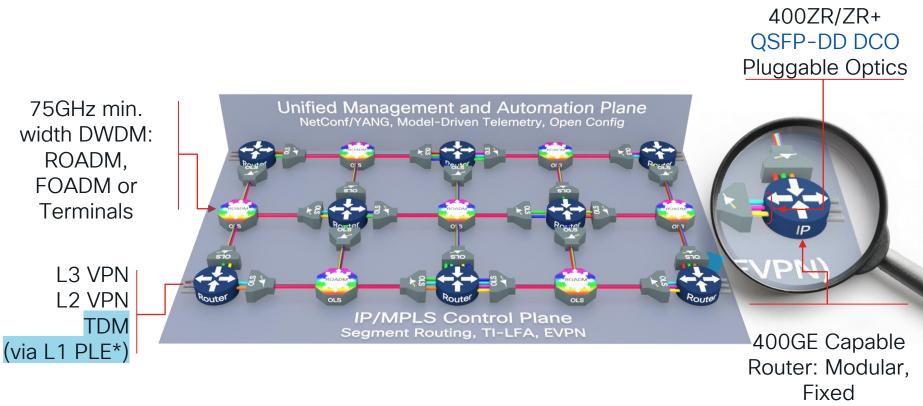
### Traditional DCI Network Architecture



#### Coherent Pluggable enables a simplified Network Architecture



## Coherent pluggable enabling Routed Optical Networking

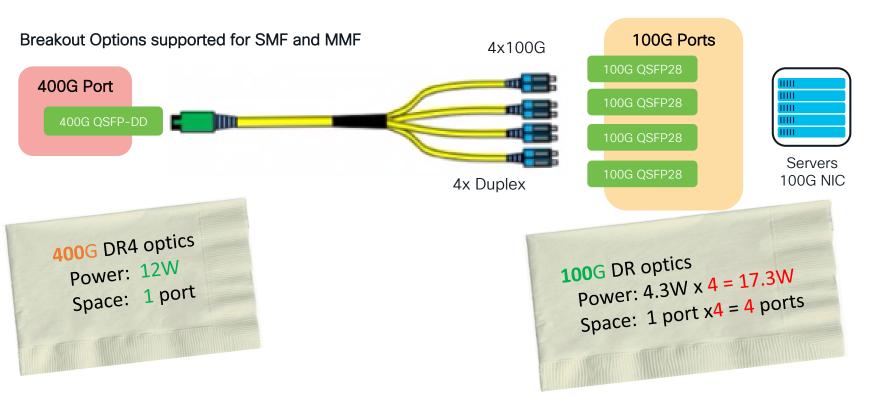


\*PLE: Private Line Emulation



## Breakout: 400G to 100G connectivity

Maximize port efficiency + forward compatibility with 100G single lambda

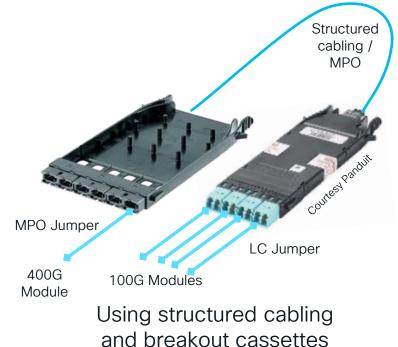


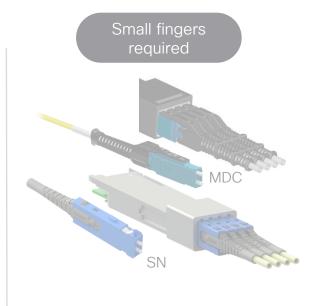


## **Deploying Breakout**

Multiple options exist

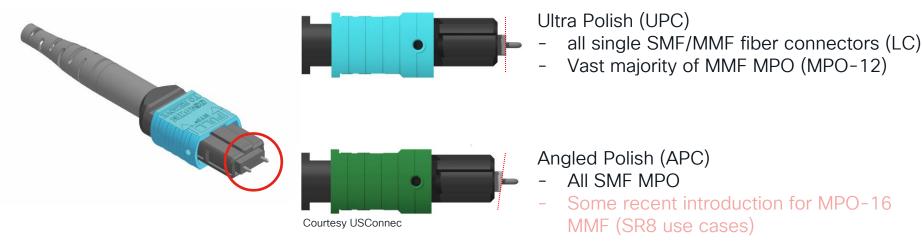






New dense VSFF connectors in module nose

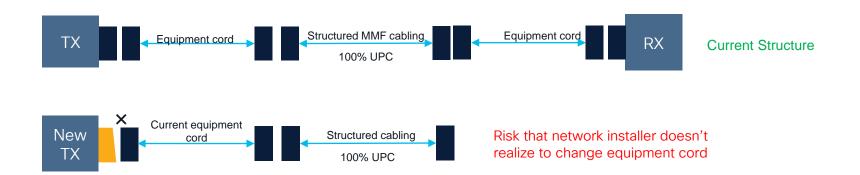
Deployment considerations:
Multi-fiber (MPO) connectors: Angled (APC) vs flat polish (UPC)



Some recent 400G MMF specs defined use of MPO APC. Awareness will prevent deployment issues



## APC vs UPC deployment usage in MMF



- Incorrect mismatch of PC and APC results in out of spec fiber plant (air gap). Damage risk exists
- MMF modules with APC is limited to MPO-16 based connectors and are being used in Hyperscaler environments for SR8 optics
- Everywhere else uses MPO-12 based connectors using UPC



## 400G summary

400G pluggable technology is mature today. Standards are complete.

400G brought a lot of innovation that will be extended into next gen

- New QSFP-DD form factor(s) capable of supporting high density at all reaches. Backwards compatible with 100G QSFP28
- High-speed PAM4 optics. Higher integration, lower cost
- Coherent pluggable: 400ZR and 400ZR+ Enables Routed Optical Network architectures
- Mainstream adoption of breakout
  - 400G module as 4x100G (SMF) or 8x 50G (MMF). DAC too.
  - Some deployment considerations with breakout

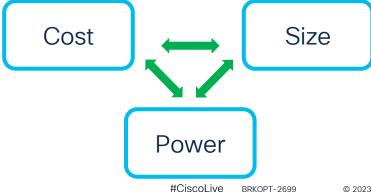


## Beyond 400G pluggable and 400 GbE



## The function of optics

- The only function of Optics is to extend the interfaces from one ASIC/Switch to another
- Therefore, it is the ASIC roadmaps which primarily matter, and the role of optics is to keep up - without causing too many issues
- Unfortunately, optics does have challenges that affect what can be built:

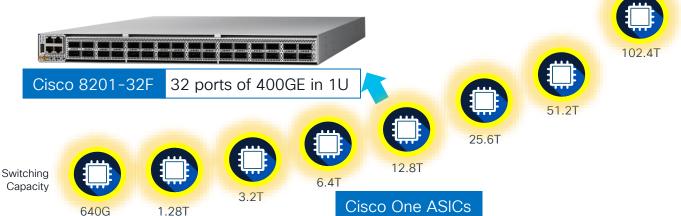




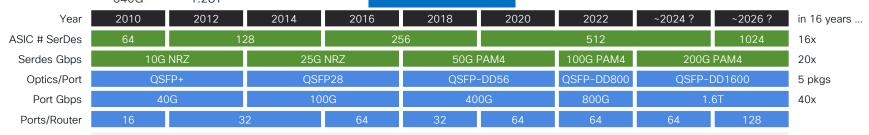
## It's all about ASICs, SerDes, and Optics



← 204.8T is 16x the bandwidth of 12.8T



320x increase in switching bandwidth over 16 years



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



## Optics (ports) follow ASIC roadmap



0)		
onfig	1 RU*	32p @ 800G
	I RU"	QSFP-DD800, OSFP <sub>800G</sub>
0	2 RU 64p @ 400G QSFP112	64p @ 400G
×e		QSFP112

	1 RU	32p @ 1.6T
	I RU	1.6T pluggable
	2 RU	64p @ 800G
		QSFP-DD800, OSFP <sub>800G</sub>

1 RU	32p @ 3.2T
	3.2T pluggable
2 RU	64p @ 1.6T
ZRU	1.6T pluggable



## 800G does not necessarily mean 800 GbE



25.6T

SERDES

-ixed Configs

100G

X256

	1 RU	32p @ 800G
	I KU	QSFP-DD800, OSFP <sub>800G</sub>
	2 RU	64p @ 400G
		QSFP112

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

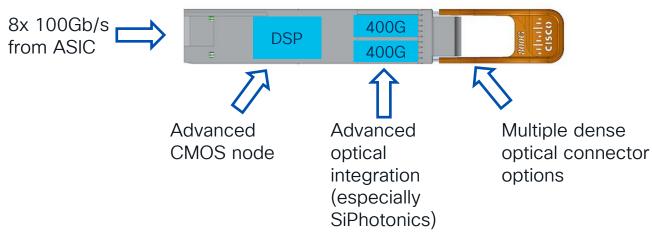
- Dominant interest in 800G module today is to support 2x 400 GbE breakout
- No immediate network need for 800 GbE
  - Although IEEE working on it (more later)
- 800G modules have the same issues:
  - Thermals and Signal integrity
  - Backwards compatibility?



## QSFP-DD800 supporting dense 400 GbE (aka breakout)

#### **QSFP-DD**::

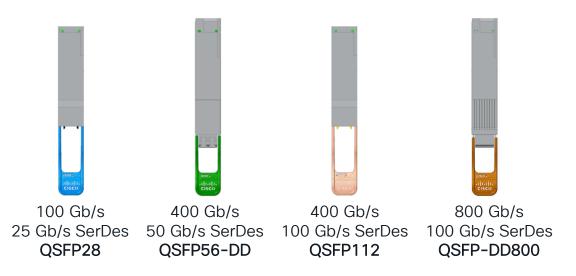
- 800G form factor enables an economical way to implement 400 GbE
  - Maximize the return on investment on the 400 GbE building blocks





## Implementing Dense 400 GbE

The QSFP module story continues



Powerful backwards compatibility



#### QSFP-DD800

#### **Thermals**

30W confirmed. No issues with any variant

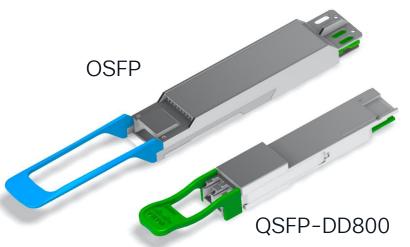
#### Electrical

· 100 Gb/s confirmed

#### Integration

- Continued CMOS node migration:
   7nm → 5nm
- · SiPhotonics continues to mature

#### 800G Optical Modules: QSFP-DD or OSFP (again)



Both variants support all the technical requirements:

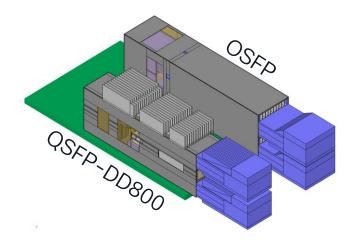
- 32 ports in 1 RU is feasible
- Electrical signal integrity @ 8x 100 Gb/s
- Thermal cooling capabilities up to 30W

Breakout optical connector options<sup>1</sup>





#### 800G Optical Modules: QSFP-DD or OSFP (again)



Showing two modules inserted into upper and lower ports in a cage.

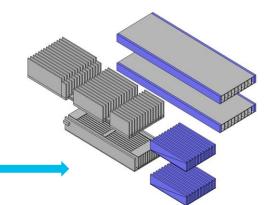
External heatsinks as part of cage (QSFP-DD) or integrated (OSFP)



#### 800G Optical Modules: QSFP-DD or OSFP (again)

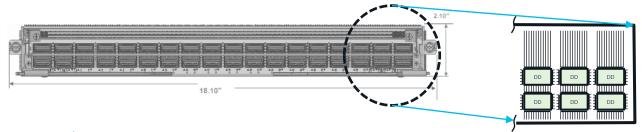
Removing modules and cages to just show the heatsink differences.

QSFP-DD based system can provide a much higher heatsink volume. Easier to cool.



Heatsink tradeoffs: QSFP-DD allows heatsinks to scale with thermal design need

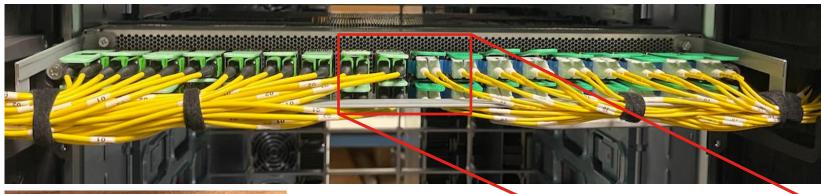
OSFP design was developed before coherent pluggable modules or 800G/1600G were considered viable



Every linecard can be optimally designed to match user needs



#### 25.6T Systems using QSFP-DD800





1RU 32 port QSFP-DD800

2x400G Dual MPO Dual LC



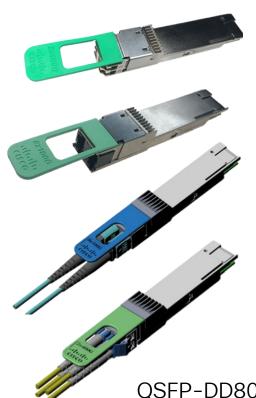
#### Market adoption forecast for 800G modules

QSFP-DD800 **OSFP 800** 400G adoption X Market alignment **Backwards** Only to OSFP 400 Compatibility Thermal performance Fixed Adaptable Technical Electrical performance details equivalent 2x optical connector support

Expect QSFP-DD800 broad adoption



#### **QSFP-DD800** Ready for the Next Generation



- Port is backward compatible to QSFP+, QSFP28, QSFP56, QSFP112, QSFP-DD
- Support 2x400G, 8x100G designs
- QSFP-DD will support over 30W of power dissipation
- Flexibility of a riding heatsink
- Increased heatsinking area for more power handling
- Wide variety of optical connector options
- Supports Dual MPO-12 & Dual Duplex LC to leverage existing cabling
- Availability of all variants is in progress
  - 8x 100 GbE (SMF and Copper Cables)
  - 2x 400 GbE (SMF and Copper Cables)
  - MMF still under investigation

QSFP-DD800 leverages QSFP the industry's cornerstone high speed form factor for the next generation of networking equipment

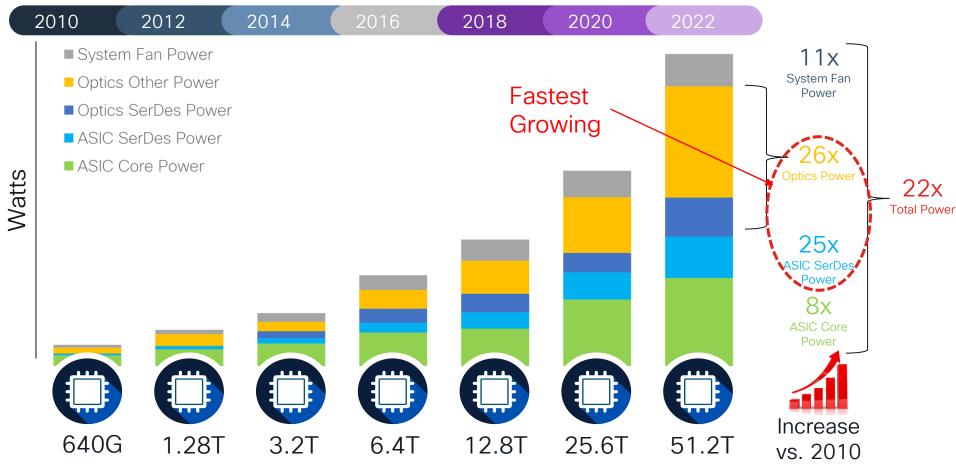
What's next for optical interconnect?

#### Going faster has some challenges

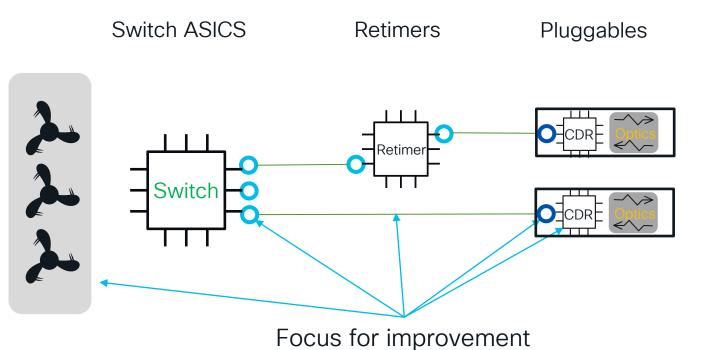


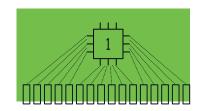


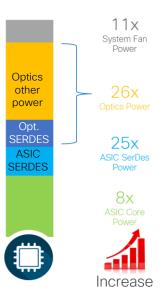
#### Relentless Advancement – 80x BW over 12 Years



#### Interconnect and power





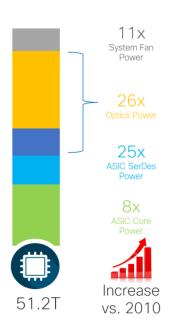




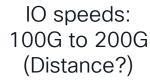
51.2T

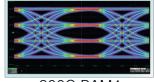
vs. 2010

#### Addressing > 800G



Growing power remains a challenge





200G PAM4



Is this the time for Co-packaged optics?

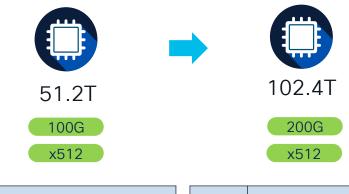


Can pluggables continue?



#### Requirements for 1.6T pluggable

2x800 GbE, 4x400 GbE



1 RU	32p @ 1.6T	
1 RU	OSFP-XD <sub>1.6T</sub>	_
2 DI I	64p @ 800G	
2 RU	QSFP-DD800, OSFP <sub>800</sub>	

1 RU	32p @ 3.2T
TRU	OSFP-XD <sub>3.2</sub>
2 RU	64p @ 1.6T
Z RU	QSFP-DD1600 OSFP <sub>1.6T</sub>

Same system design density and speed requirements with 200Gb/s SERDES.

 QSFP-DD and OSFP roadmaps will extend to 1.6T

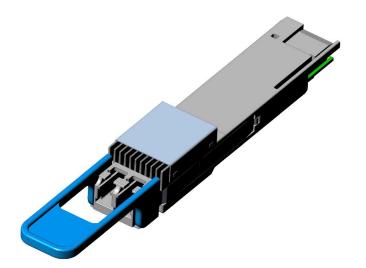
New entrant (just to add to the mix)

- OSFP-XD
  - New 16x100G connector
  - Despite name not compatible with OSFP
  - Solves the 51.2T 1RU "problem" but market shifting to 2RU
  - Some good design work that may be useful in future



#Ciscol ive

#### The good news: Clear path to QSFP-DD1600



QSFP-DD MSA already working on QSFP-DD1600

Module form factor allows innovation on:

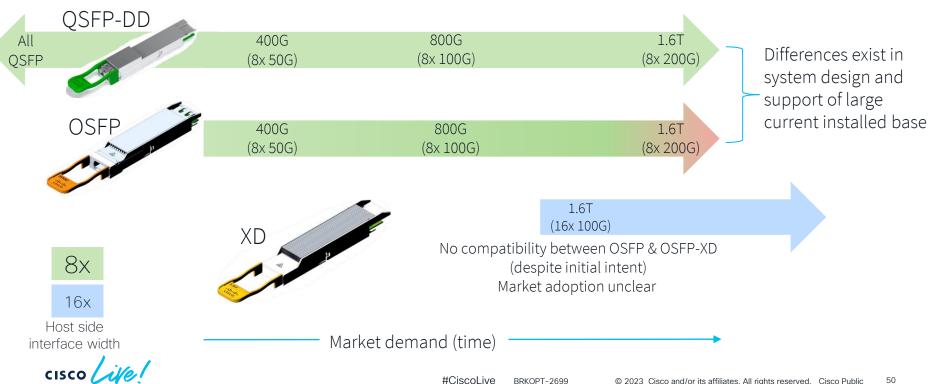
- Connector optimization
- Cage design
- Heatsink and thermal design
- System design

QSFP28 → QSFP-DD → QSFP-DD1600



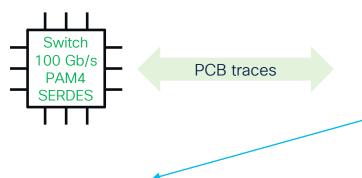
#### Path ahead

Pluggable optics roadmap continues and extends beyond 800G



#### Linear drive optics





Current (Retimed)

- DSP fully equalizes electrical & optical signals
- Enables broad interoperability
  - Host/Port/Module
- Full telemetry & loopbacks possible
- FEC Monitoring or Partitioning possible
- But this adds power

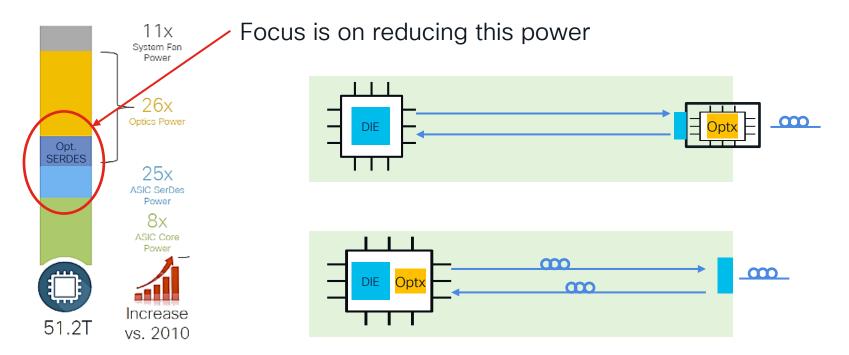
#### Linear (Non-retimed)

Chip

- Power reduction due to absence of equalization
- Performance is based on quality of every component in link (engineered link)
  - Optics, Serdes, PCB, connectors
  - Varies port to port
- Interoperability isn't "broad" any more
- No telemetry or loopback



#### Beyond Pluggable: Co-Packaged Optics





#### Beyond Pluggable: Co-Packaged Optics



Cisco 25.6T CPO system demonstrated @ OFC'23

• 25-30% power reduction

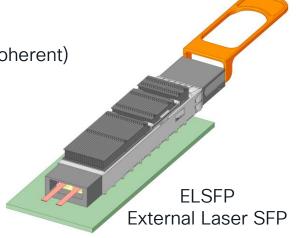


Fiber connectors

Optical power source (ELSFP)

#### Co-Packaged Optics

- Co-packaged (CPO) and Near-packaged (NPO) variants can provide power reductions. Silicon Photonics required.
- Some early industry standardization in flight. Goal to establish interoperable components → new "ecosystem"
- System configs would likely be:
  - 100% optics
  - 50% co-packaged optics / 50% pluggable ports (copper cables, coherent)
- Operational considerations:
  - Pluggable external laser power sources (reliability, thermals)
  - Reduced overall power
  - Reduced port flexibility



#### Maintaining reach at higher speeds



Direct Detect 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 DWDM interop, OIF 400ZR, OpenZR+, OpenROADM, ITU SG15/Q11





#### Road ahead for Coherent MSA pluggable

Today **Future** 

Expanding applications for 400G pluggables



50G SerDes 16-22W

Next-Gen 800G and 400G LH pluggables



50G/100G SerDes

22-30W

1.6T Coherent for DC, carrier networks



100G/200G SerDes

25-40W



Where are the standards going?



#### Cisco's Focus: Compatibility and Standards

- Network operators require compatibility as they transition and adopt next generation technologies
- Cisco focuses on standard interfaces as much as possible to enable these smooth transitions



#### Standards Landscape

The interconnect industry is working to standardize what is needed. Standardization required to ensure interoperability

#### Optical Internetworking Forum

- Extending the 400ZR work
- 800ZR in definition (targets same application as 400ZR)
- New: 800LR shorter reach coherent (10km)
- Module management (CMIS)

#### IEEE 802.3df and 802.3dj

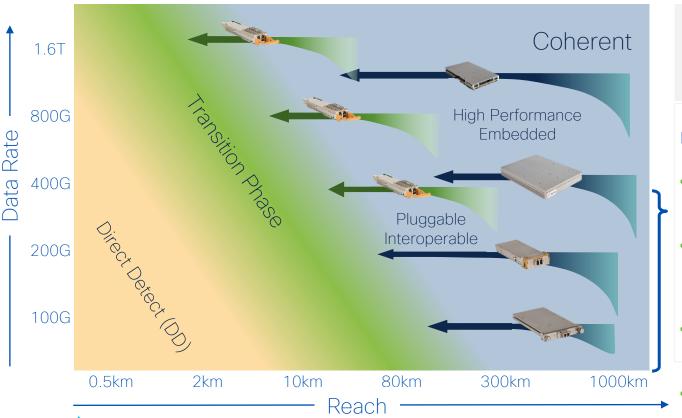
 The foundational Ethernet standards for 800 GbE and 1.6 TbE

#### Form factors

- Form factor MSAs as we have covered already:
  - QSFP-DD, OSFP-XD, OSFP
- OIF is working on copackaging implementation definitions



#### Coherent Technology Trends - Shorter Reaches



Edge applications driving transition towards shorter reach coherent solutions & industry standardization

Industry Momentum for Pluggable DCO Driven by Edge and Access

- OIF
  - Metro DCI 400ZR standard
  - •800LR 10KM in works
- IFFF
  - 100GbE & 400GbE beyond 80km reach
  - •800 GbE -10km & 40km
- CableLabs
  - Adopted 100G & 200G
     Coherent Access Standard
- Open ROADM
  - 100G/200G/300G/400G

#### IEEE 802.3df/3dj are working on 800 GbE and 1.6 TbE

Task Force underway. Initial specs being adopted.

#### Major themes:

- 1) 100 Gb/s based. (P802.3df)
  - 800 GbE based on an 8x 100 Gb/s approach (optical and electrical)
- 2) 200 Gb/s based (P802.3dj)
  - 800 GbE (4x 200 Gb/s) and 1.6 TbE (8x 200 Gb/s) (optical and electrical)
  - Will include 200 GbE and 400 GbE variations (useful for breakout)
- 3) Coherent solutions will be defined for 10 km and 40km
- 4) New standard reach 2km parallel SMF



#### 802.3df & 802.3dj: Adopted Objectives

Ethernet Rate	Assumed Signaling Rate	AUI	Cu Cable	MMF 50m	MMF 100m	SMF 500m	SMF 2km	SMF 10km	SMF 40km	
200 Gb/s	200 Gb/s	Over 1 lane	Over 1 pair			Over 1 Pair	Over 1 Pair			
400 Gb/s	100 Gb/s						Over 4 Pair			
	200 Gb/s	Over 2 lanes	Over 2 pairs			Over 2 Pair				
800 Gb/s	100 Gb/s	Over 8 lanes	Over 8 pairs							
	200 Gb/s	Over 4 lanes	Over 4 pairs			Over 4 pairs	<ol> <li>Over 4 pairs</li> <li>Over 4 λ's</li> </ol>			
	800 Gb/s							Over 1 pair	Over 1 pair	
1.6 Tb/s	100 Gb/s	Over 16 lanes								
	200 Gb/s	Over 8 lanes	Over 8 pairs			Over 8 pairs	Over 8 pairs			ļ

#### Technology Reuse

802.3df

Leverage existing or work-inprogress 100 Gb/s per lane (e.g. 3cu, 3ck, 3db) to higher lane counts

802.3dj

Develop 200 Gb/s per lane electrical signaling for 1/2/4/8 lane variants of AUIs and electrical PMDs

Develop 200 Gb/s per optical fiber for 1/2/4/8 fiber based optical PMDs and 4 lambda WDM optical PMD

Coherent signaling technology



#### 800 GbE/1.6 TbE Timelines

#### IEEE 802.3df and 802.3dj timelines

- 8x 100 Gb/s variants: Specifications mature now. Std complete mid-2024
- 4x 200Gb/s variants: Baselines adopted later this year, Std complete 2026
- Coherent variants: Baselines later this year, Std complete 2026

OIF is making good progress on standardizing 800G coherent for:

- ZR (~100km) estimated completion 1H'24
- LR (~10km) estimated completion 1H'25



Wrap up



#### Summary

- 400G forced a lot of innovation. Today the market is growing & the technology is mature
  - Wide adoption QSFP-DD pluggable form factor for non-coherent and coherent interfaces. Enables new network architectures (routed optical networking)
- 800G is building on that innovation fully using 100 Gb/s technology. Market is starting to deploy.
  - Power is a dominant challenge, Breakout becoming the norm (2x 400 GbE etc.)
  - Reuse existing fiber infrastructure
  - Continued backward compatibility with QSFP-DD800
- Beyond 800G prediction of the death of pluggables is premature
  - 1.6T → QSFP-DD1600 will continue to support.
  - New implementations under development (co-packaged, linear).
- IEEE defining next phase of foundational specifications today 800 GbE & 1.6 TbE



#### Fill out your session surveys!



Attendees who fill out a minimum of four session surveys and the overall event survey will get **Cisco Live-branded socks** (while supplies last)!



Attendees will also earn 100 points in the **Cisco Live Challenge** for every survey completed.



These points help you get on the leaderboard and increase your chances of winning daily and grand prizes



# 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



#### Thank you



### Cisco Live Challenge

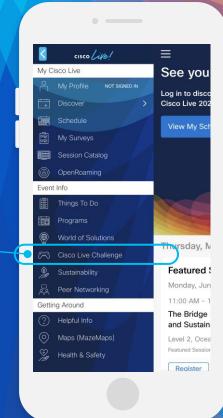
Gamify your Cisco Live experience! Get points for attending this session!

#### How:

- Open the Cisco Events App.
- 2 Click on 'Cisco Live Challenge' in the side menu.
- 3 Click on View Your Badges at the top.
- 4 Click the + at the bottom of the screen and scan the QR code:







## Let's go cisco live! #CiscoLive