Improve your data center network efficiency with 400G optics

Cisco Knowledge Network, January 26, 2023

Errol Roberts, Distinguished Architect
Ray Nering, Product Manager
Yoav Schreiber, Marketing Manager

Agenda

- Growth drivers
- Technology enablers
- Strategic decisions
- Use cases
- Path to 800G and beyond
- Summary and next steps

Increasing demands on data centers

300%

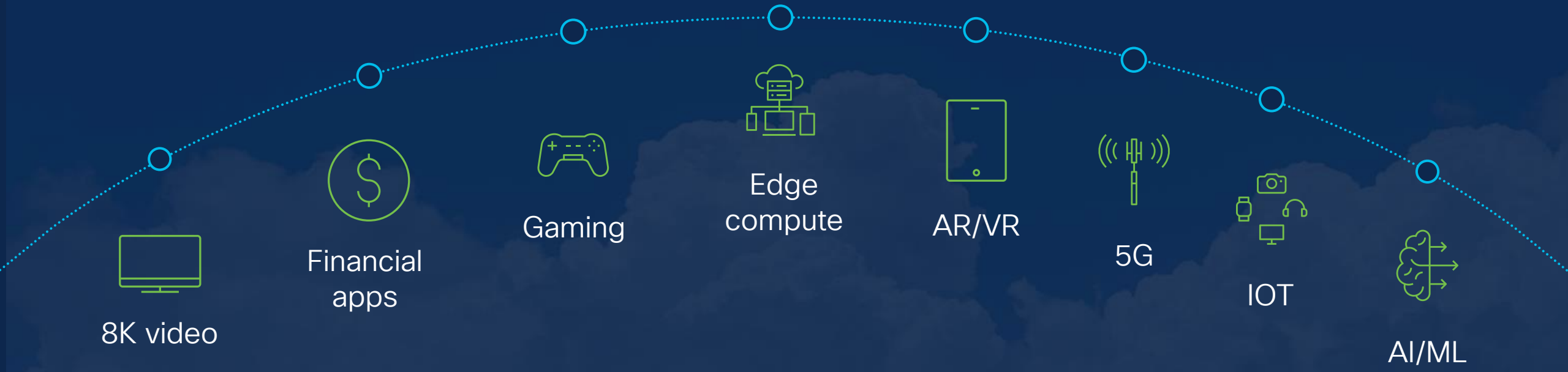
more applications will run in the data center and edge locations

500 million

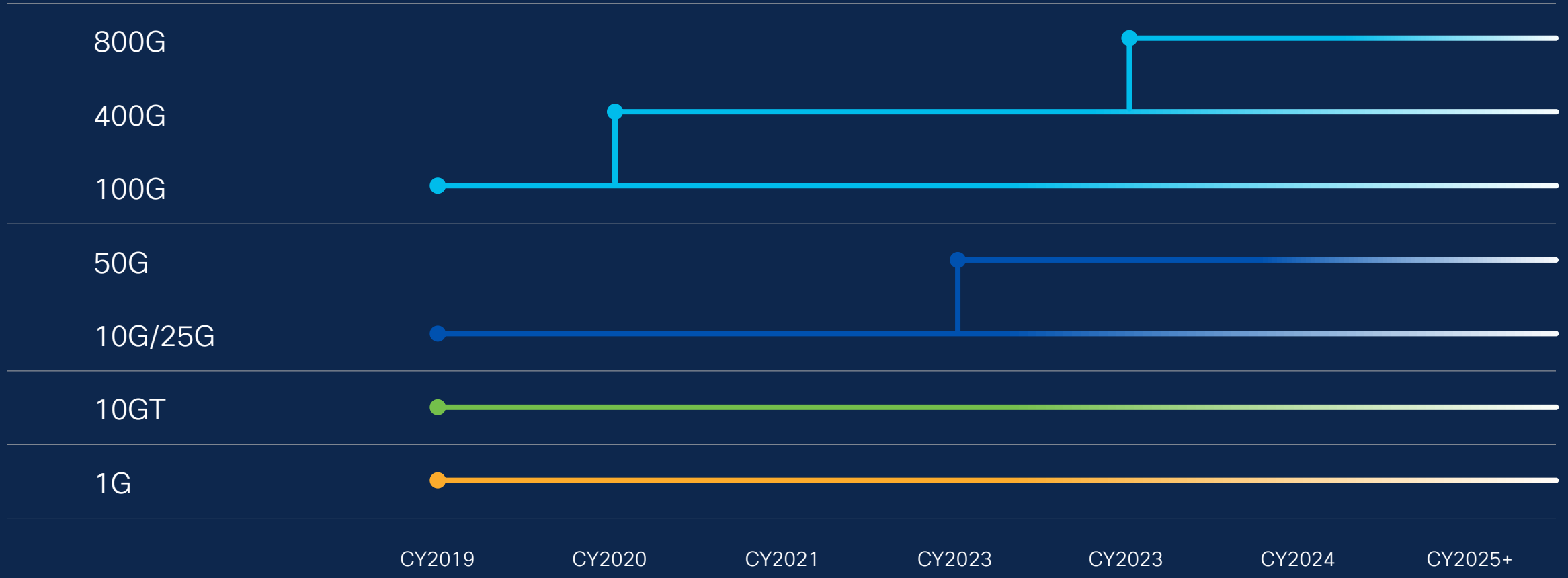
cloud-native applications and services will be developed

40%

of new enterprise IT infrastructure will be deployed at the edge



Data center speed transitions



Importance of optics innovation

Increasing demands for network performance

Growth of data intensive workloads within and between data centers

Strategic importance to delivering connectivity

Growing percent of hardware BOM as speeds increase

Growing sophistication in technology and production

Integrating more capabilities, wafer-scale manufacturing, thermal efficiency and sustainability

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Public

Data center operator top of mind

Increasing data center capacity and sustainability

Preserve investments in existing optics infrastructure and cabling

Simplify operations and management of optical links

Migrating to 400G capacity and preparing for 800G

Technology enablers for more efficient optical modules

Silicon/SerDes advancements

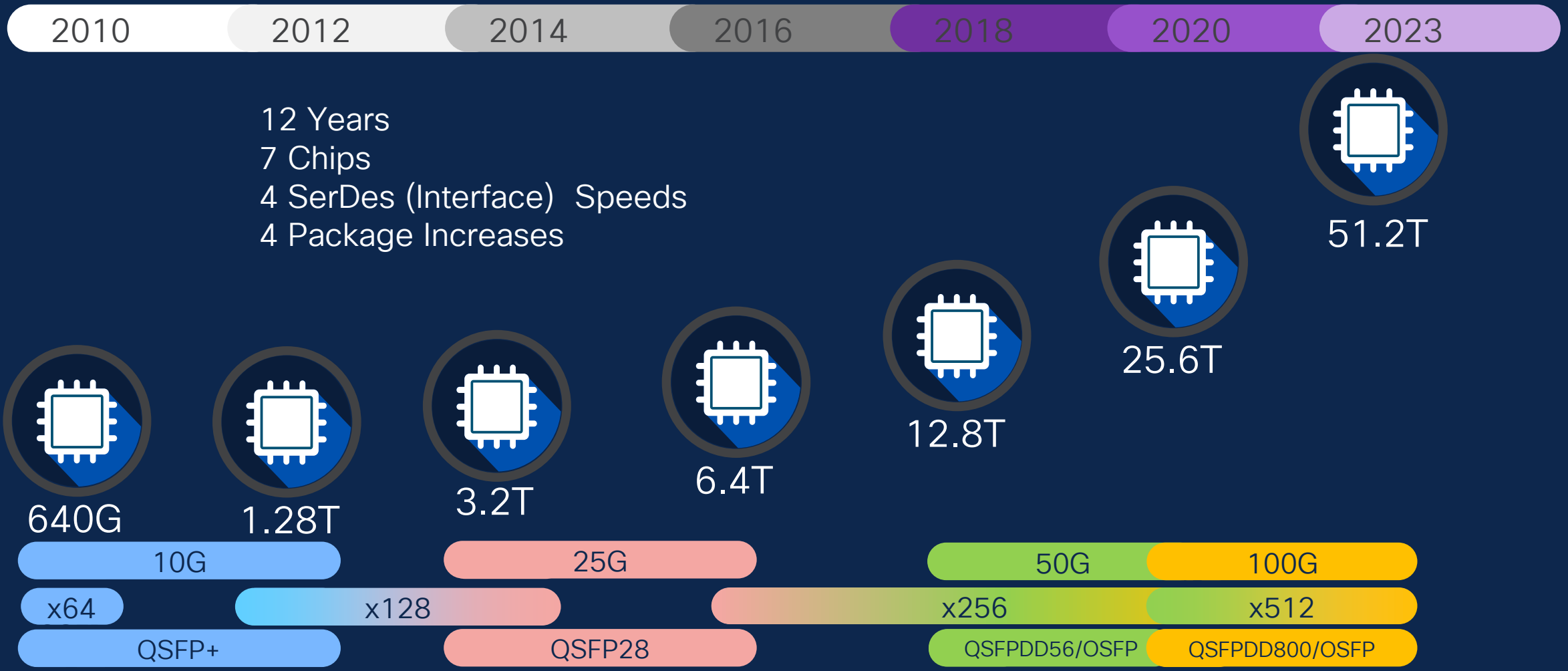
PCIe bandwidth expansion

QSFP-DD form factor

Silicon photonics integration

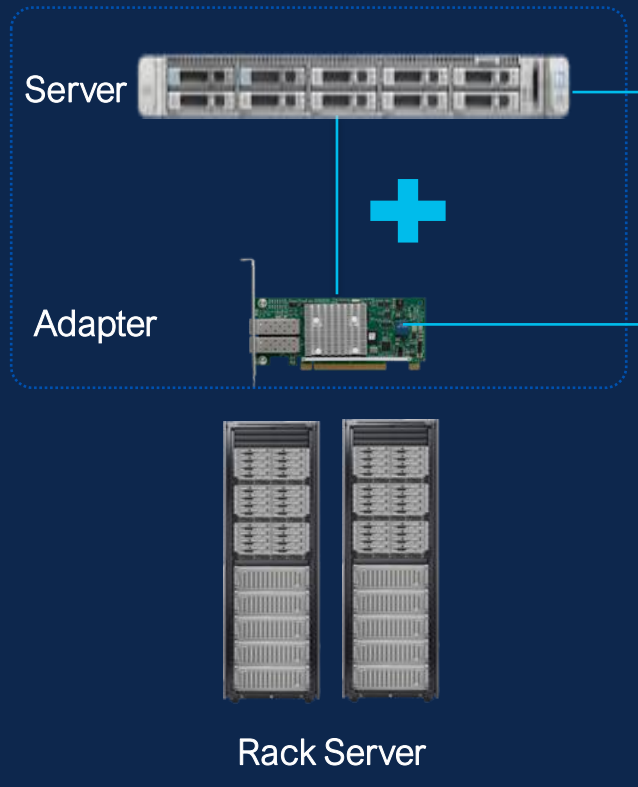
PAM4 and coherent modulation

ASIC capacity drives importance of optics efficiency



PCIe bandwidth expansion driving higher server access speeds

Evolution in NIC and server performance



- Faster CPU
- More and faster memory speeds
- Higher Power

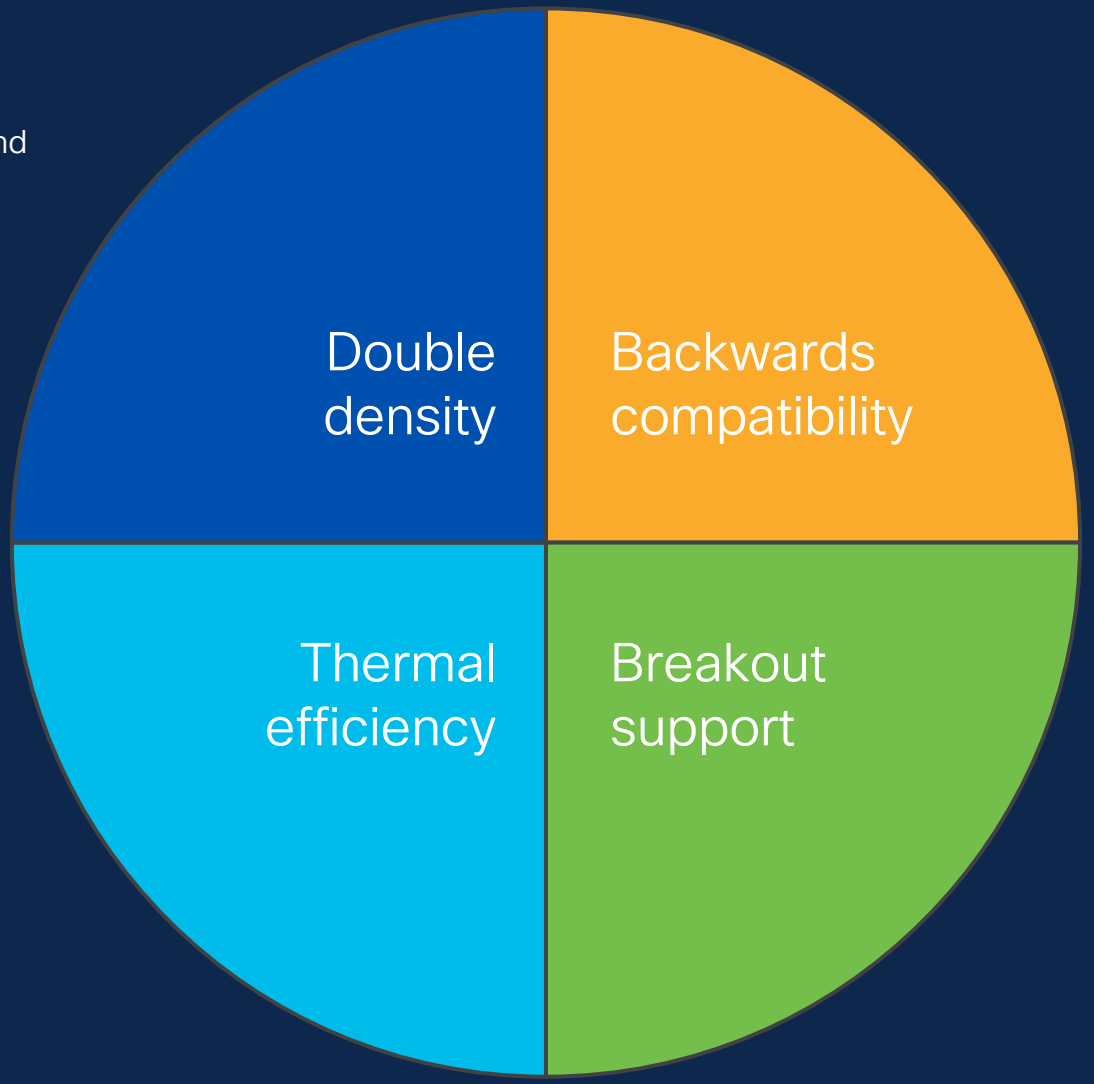
- Increase PCIe Lane performance
- Increase PCIe Lanes from x8 to x16 and x32
- Increase uplink interface speeds from 10G to 25/100G/200G and 400G

Ports	Modulation	BW per Lane (Single direction)	x8	x16	x32
PCIe Gen2	NRZ	5 Gbps	40 Gbps	80 Gbps	160 Gbps
PCIe Gen3	NRZ	8 Gbps	64 Gbps	128 Gbps	256 Gbps
PCIe Gen4	NRZ	16 Gbps	128 Gbps	256 Gbps	512 Gbps
PCIe Gen5	NRZ	32 Gbps	256 Gbps	512 Gbps	1,024 Gbps
PCIe Gen6	PAM-4	64 Gbps	512 Gbps	1,024Gbps	2,048 Gbps

QSFP-DD innovations for 400G

Same faceplate with 2nd row of pins. Maintains 36 ports per RU

Plug existing QSFP+, QSFP28, QSFP56 modules into QSSFP-DD ports

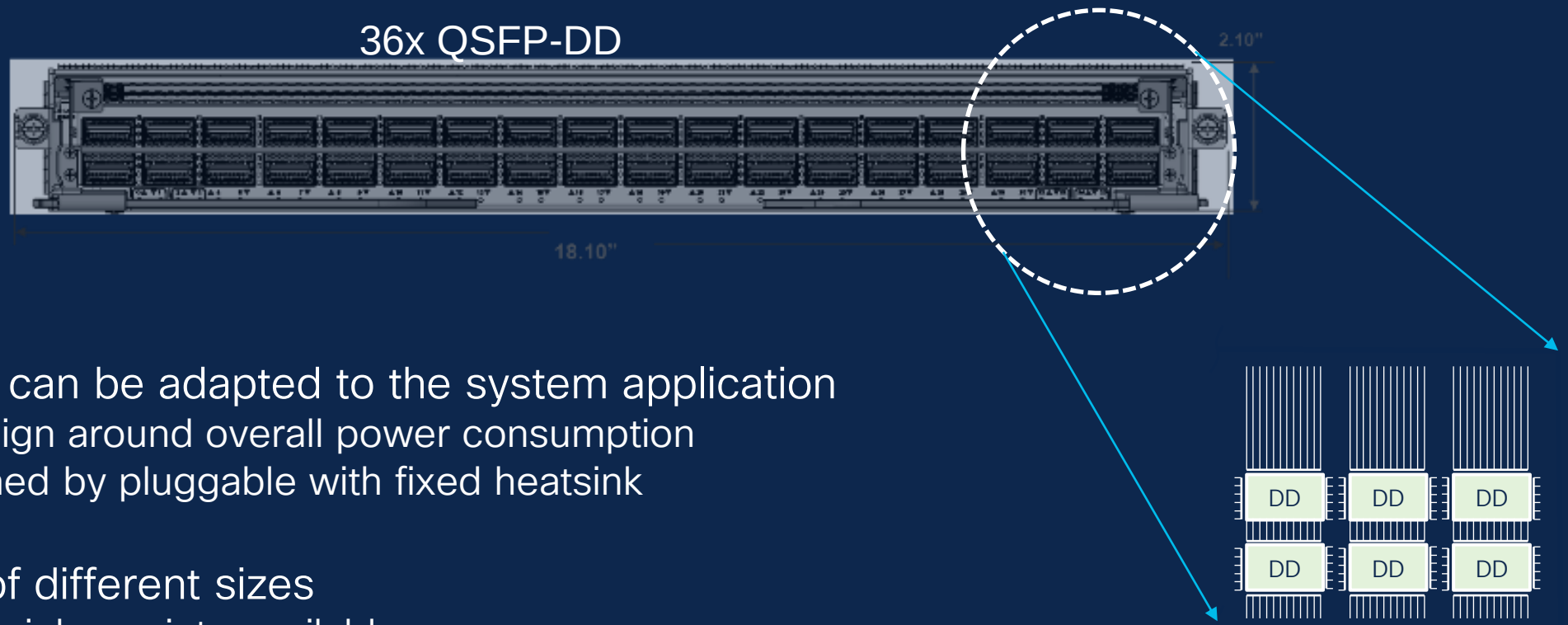


Support for coherent modules with over 25w of power dissipation in riding heatsink on platform

Connect to multiple lower speed interfaces

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Public

Design flexibility with QSFP-DD riding heatsink



- Thermal design can be adapted to the system application
  - Optimize design around overall power consumption
  - Not constrained by pluggable with fixed heatsink
- Add heatsinks of different sizes
  - Extend heatsinks up into available space
  - High-power row, low-power row

Increased integration and efficiency with QSFP-DD modules

Fewer components

- Less lanes
- More optical functions
- Improved reliability
- Economies of scale

Increased density

- Higher speed per lane
- More integration
- Packaging

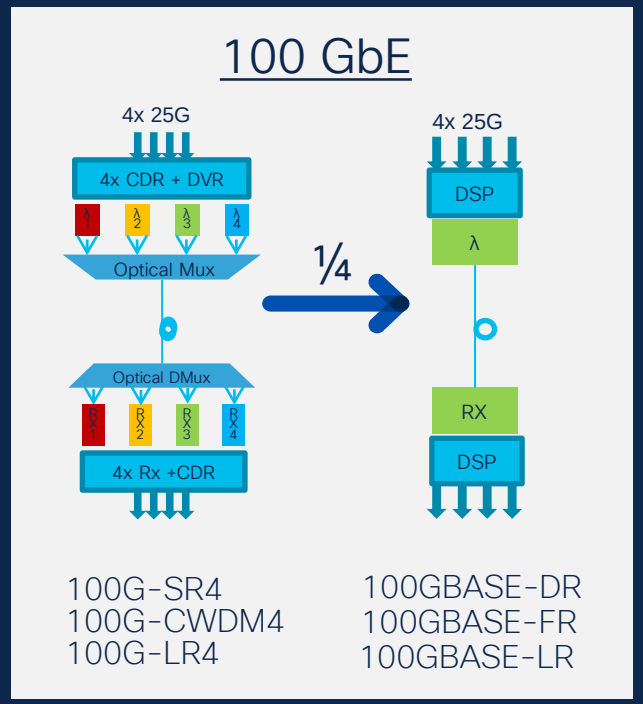
Power efficiency

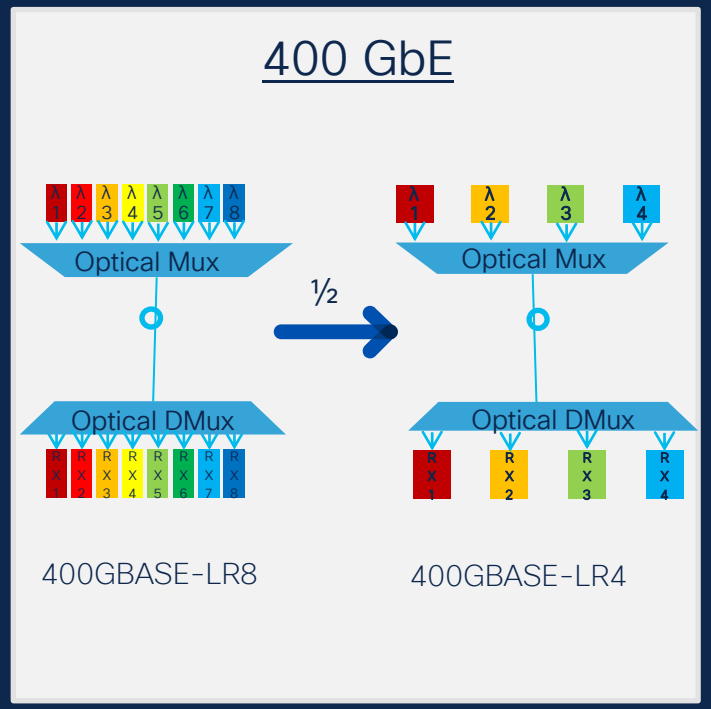
- More integration
- Advanced technology

Ease of use

- Compatibility forward and backward
- Multi-application use case

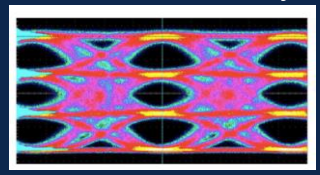
100G modulation enables greater DSP integration and simplification for 400G

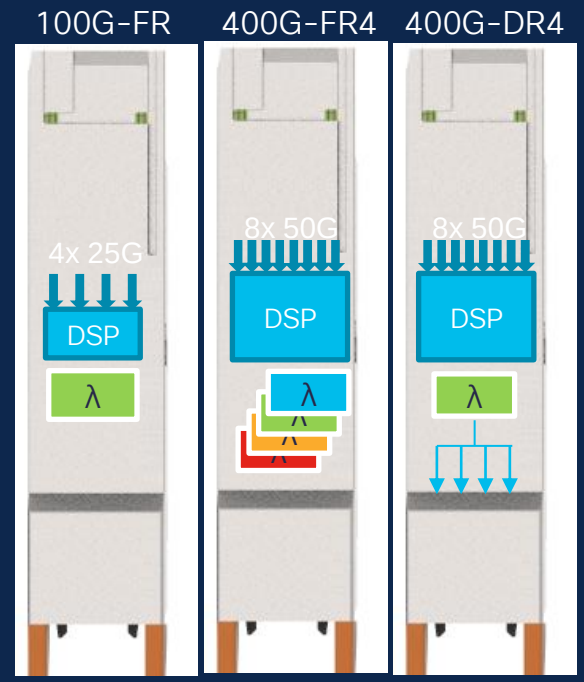




Component reduction to a single DSP

100 Gb/s PAM4 eye

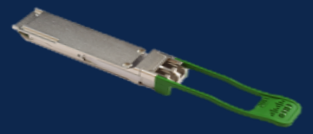




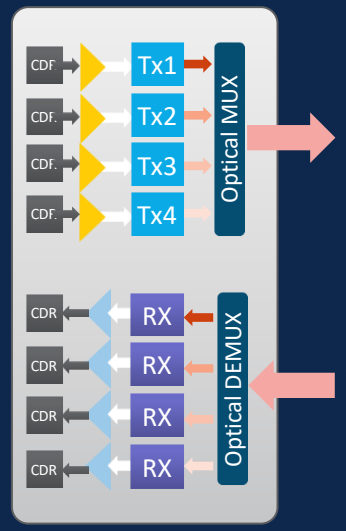
Silicon Photonics enables a single laser to be split across 4 lanes

Single-wavelength 100G optics forward compatibility

Before



4x25Gb/s electrical interface



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

QSFP-100G-SR4 — 100m

QSFP-100G-SM-SR — 2km

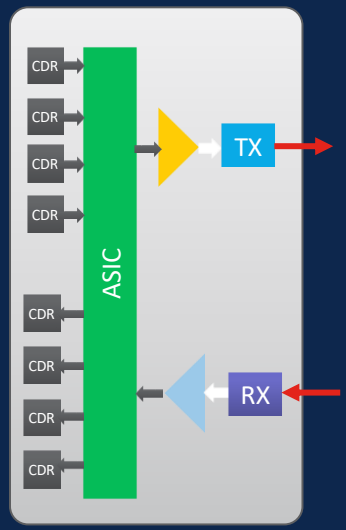
QSFP-100G-LR4 — 10km

QSFP-100G-ER4L — 40km

QSFP-100G-ZR4 — 80km

Now

4x25Gb/s electrical interface



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

500m QSFP-100G-DR

2km QSFP-100G-FR

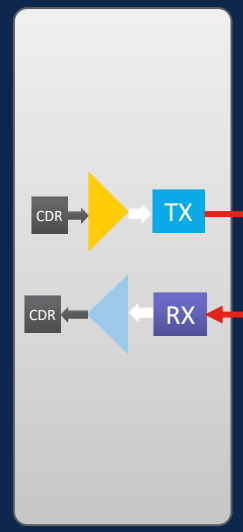
10km QSFP-100G-LR

25km QSFP-100G-ERL

120km QDD-400G-ZR+

Future

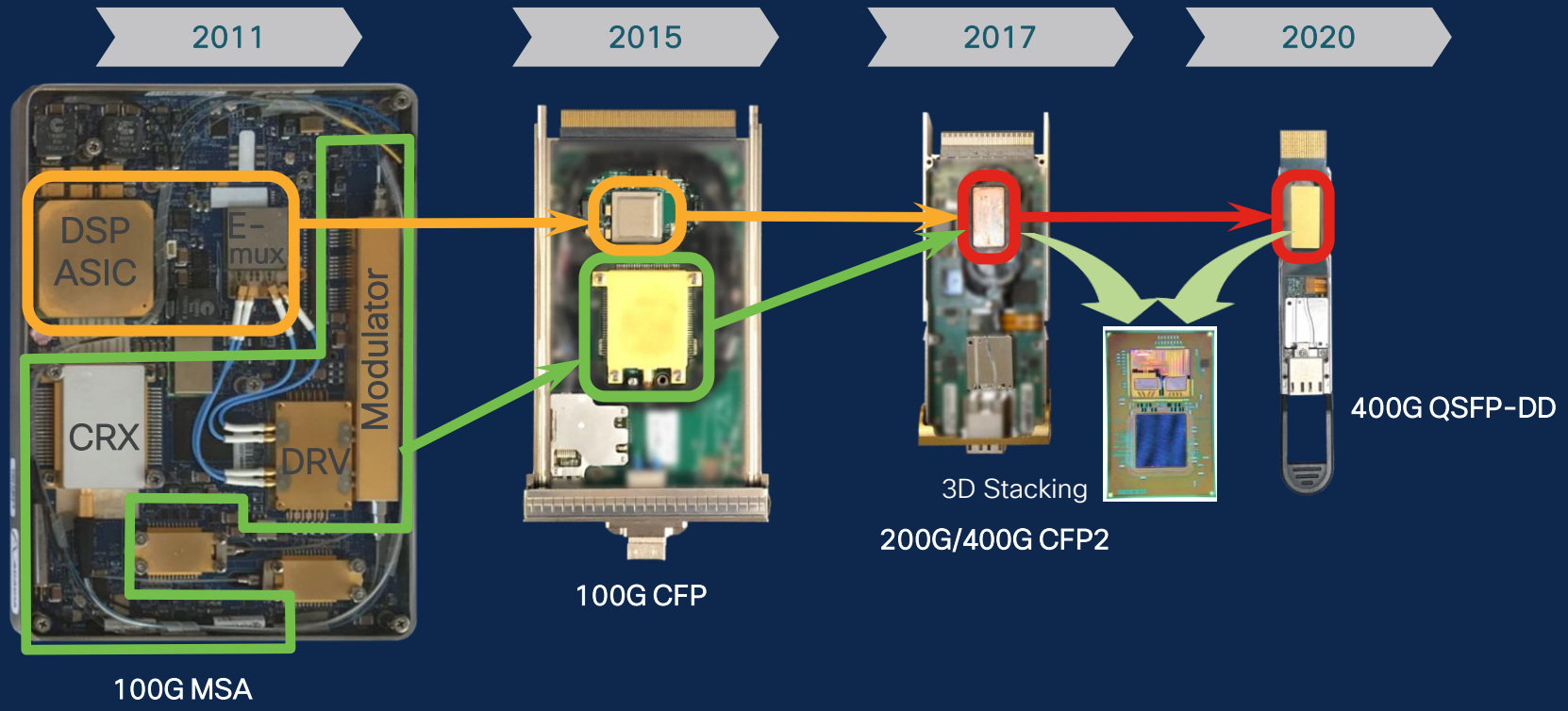
1x100Gb/s electrical interface



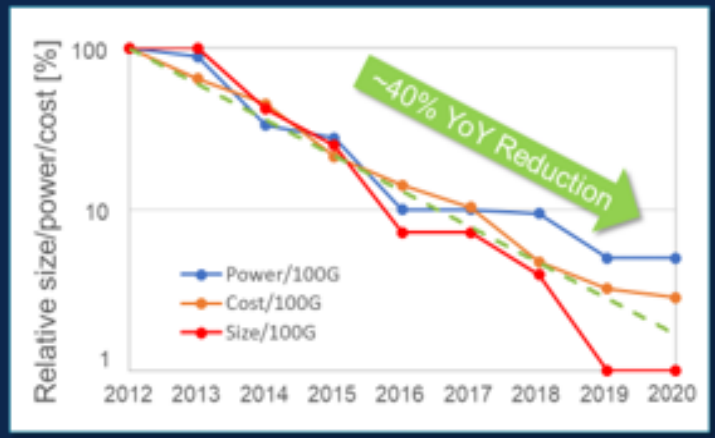
1x100Gb/s optical interface (on a duplex fiber pair)

Forward compatible optical network interface

Increased efficiency with siliconization of 400G coherent optics



Reducing Power, Size & Cost through Siliconization



Strategic decisions for more efficient data center architectures

How do I optimize data center design flexibility to support 400G?

- Reach, operational flexibility, manageability

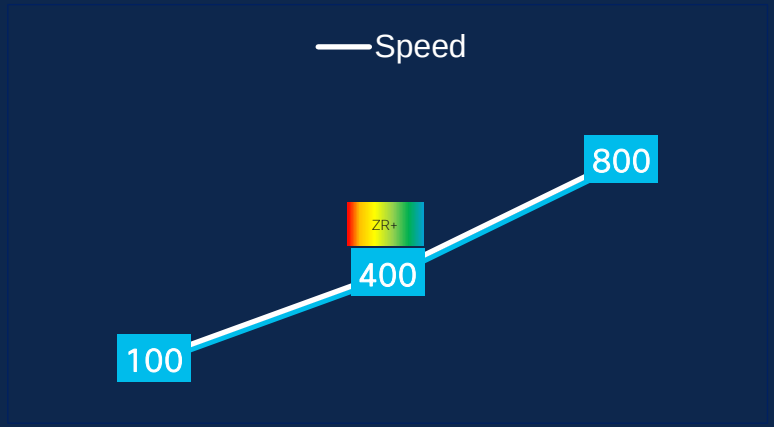
How do I migrate from existing platforms and links to 400G?

- Operationalizing breakout

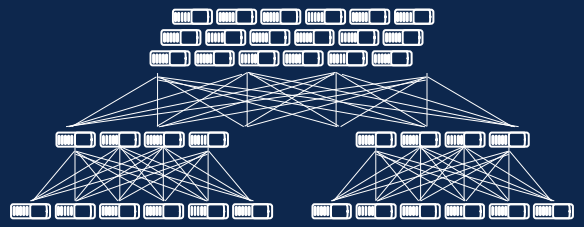
How do I ensure network reliability?

- Component integration improves reliability even at higher speeds

400G impact to data center network fabric



High Scale Leaf/Spine based Designs



- Common network architecture between 100G, 400G and 800G
  - Same physical port densities, same media reaches
  - Coherent and non-coherent deployments
  - Continued investment in fiber plant
- Flexibility to adopt 400G breakout for high radix 100 GbE design
  - Connectivity to 100 GbE equipment
- Design flexibility
  - High bandwidth, high port density platform flexibility w/ fixed, modular
  - Link bandwidth distribution
  - Port flexibility – non / coherent use cases, and mixed data rates
  - Cabling flexibility
- Backward and forward compatibility
  - 100G, 400G, 800G

Operational and design flexibility with 400G

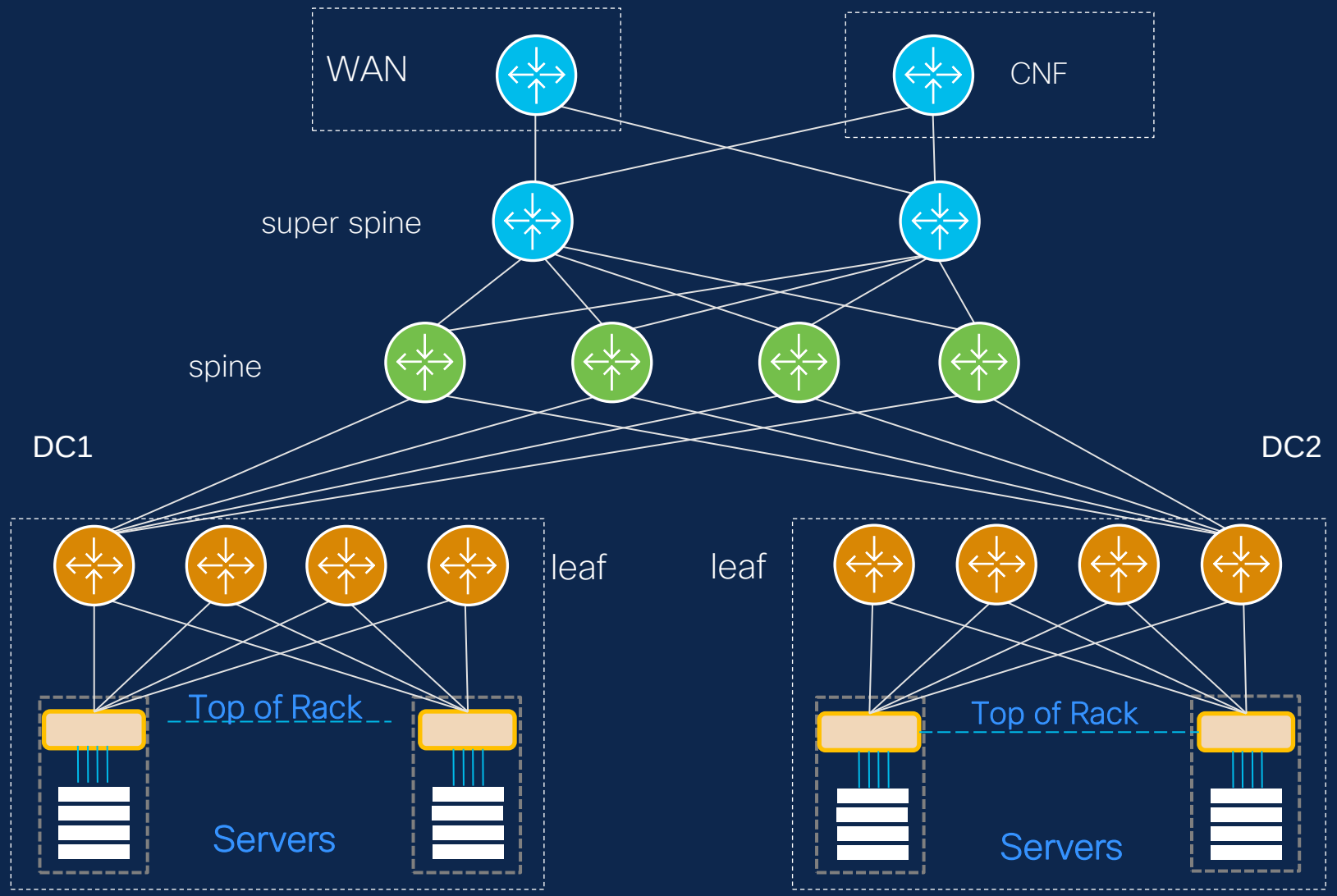
2-Tier Leaf/Spine

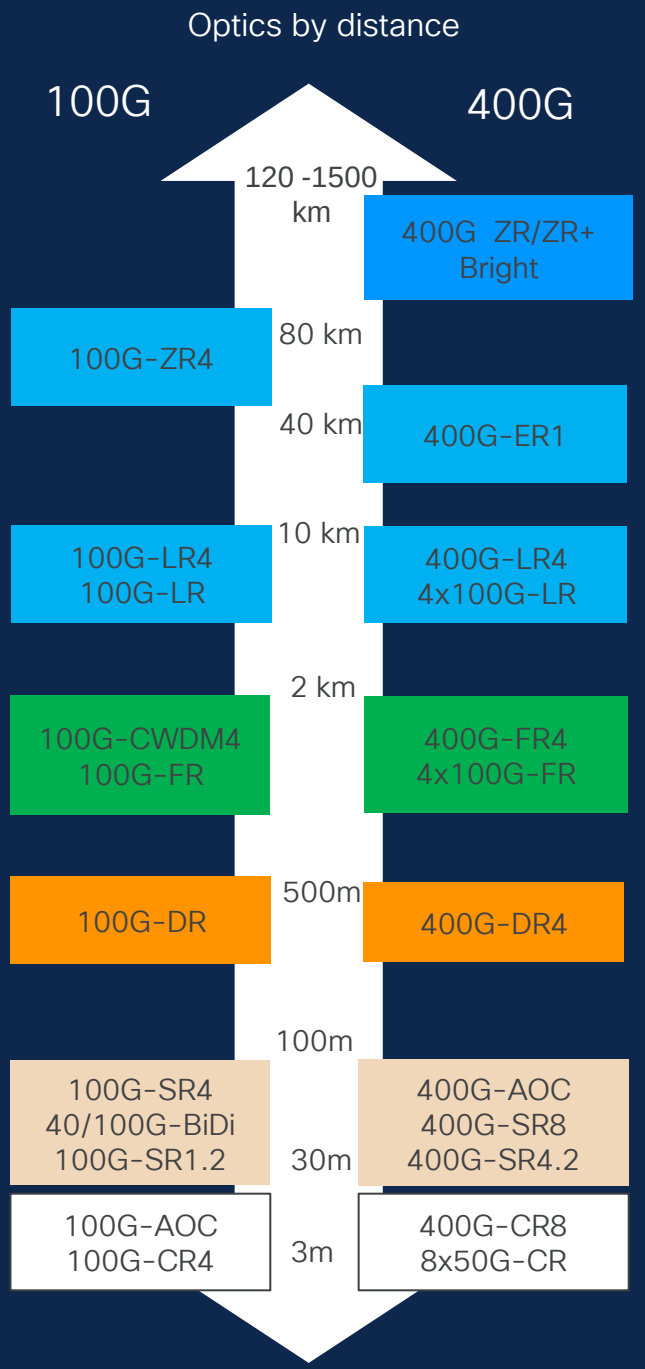
3-Tier Leaf/Spine

Evolved 3-Tier Leaf/Spine

- Increasing scale-out in all tiers
- Improve system capacity with dense 400G/800G platforms
- Cost optimization with lower cost/bit and improved power efficiency
- Latency optimization for network designs
- Improved application performance – high bandwidth 400G fabric
  - Improved ECMP performance – bigger flows, larger flow buckets
  - Intelligent buffering
- Breakout for leaf spine and server access – improved design flexibility

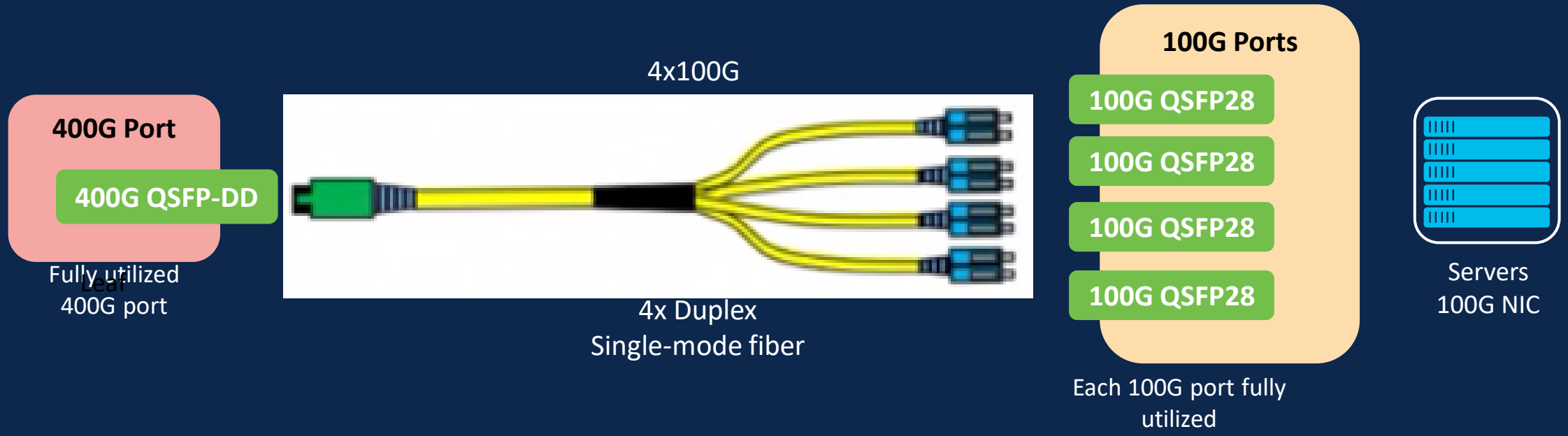
400G optics cover the entire network



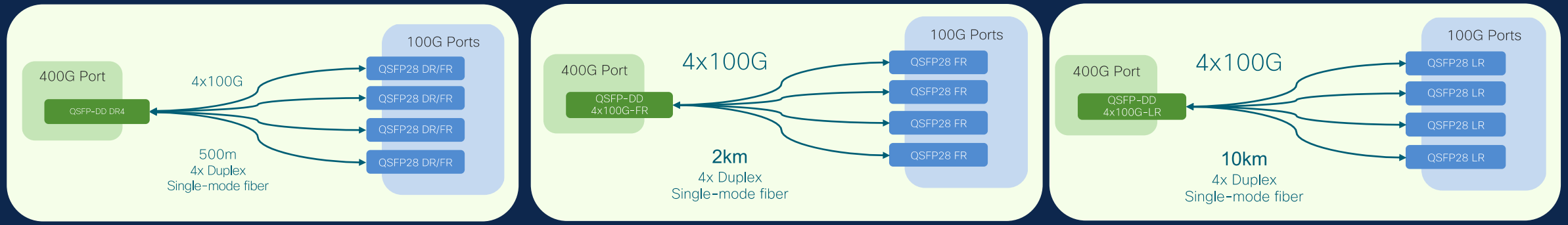


400G to 100G connectivity with breakout

Maximize port efficiency + forward compatibility with 100G single lambda



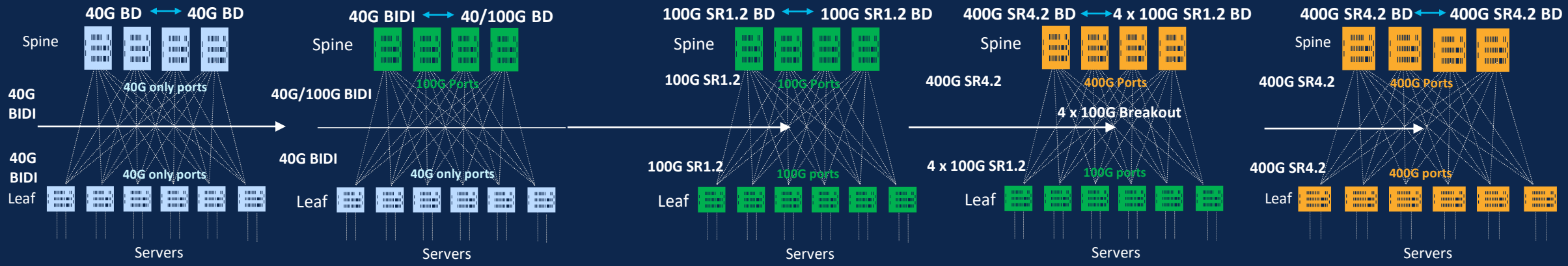
Breakout Options



Migration to 400G over multi-mode fiber

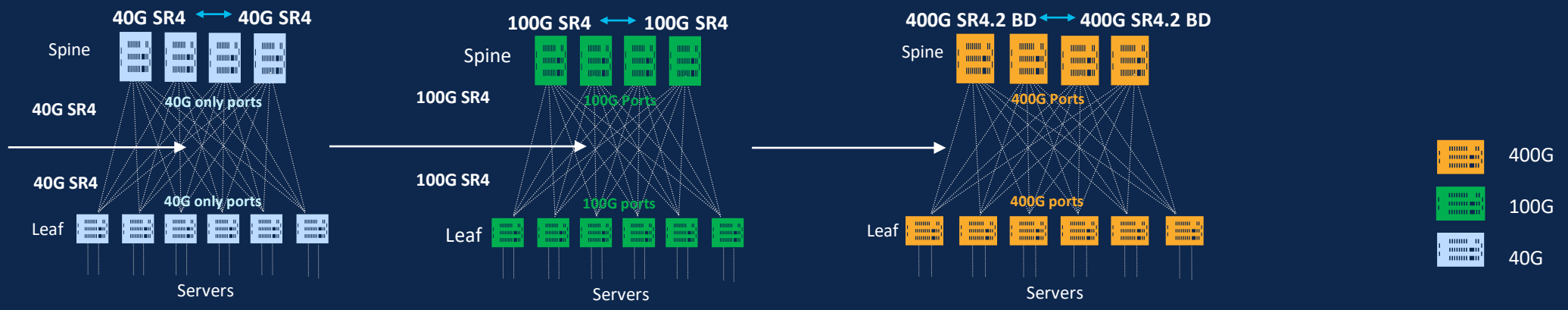
SR1.2 IEEE802.3bm
SR4.2 IEEE802.3cm

Maintain Duplex Fiber Infrastructure



Maintain SR4 Fiber Infrastructure

Transition from Duplex to SR4 Fiber Infrastructure



Architectures and use cases

Multi-tier fabrics

AI/ML clusters

Edge compute

Data center interconnect

Customer requirements

400G/800G

100G/400G

Hyperscalers

100G/400G/800G fabrics

AI/ML compute clusters

Webscalers

Scale-out fabrics

25/50/100G server NICs

Enterprise

High performance IO

AI/ML compute clusters

Media providers

8K uncompressed video driving 100G endpoints

Need for 400G uplinks

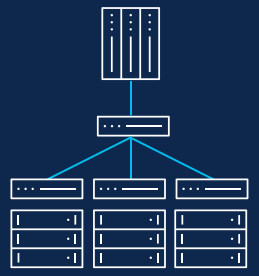
Telco service providers

100G/400G fabrics

Space constrained DC and edge locations

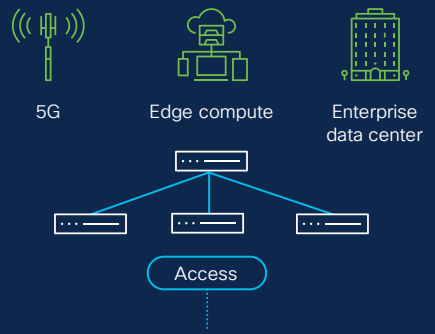
400G data center use cases

Multi-tier fabrics for leaf, spine, super-spine

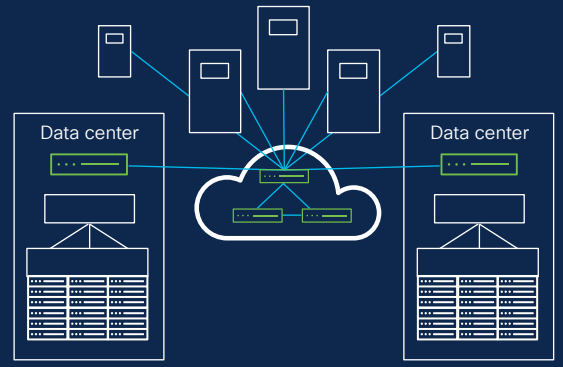


Large scale AI/ML clusters

Edge compute



Data Center Interconnect (DCI)



Hyperscaler and webscaler deployment

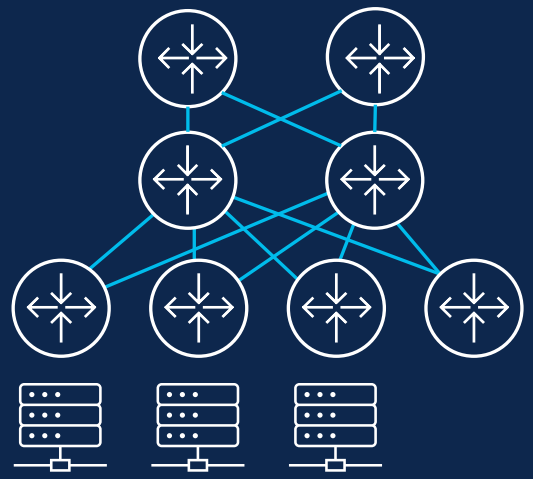
Topology

Requirements

- Cloud-based, large-scale AI/ML clusters and low latency networks
- 100G/200G/400G compute with proportionally scaled-out fabrics and interconnects
- High radix switches for optimized power consumption
- Data protection via MACSec and IPSec

Financial services deployment

Topology

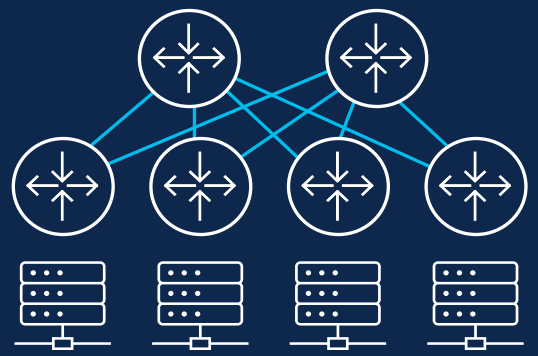


Requirements

- Super Spine connecting multiple data halls
- N+1 Redundancy across all spines
- 400G in different fabric tiers
- Encryption between data halls and sites

Service provider deployment

Topology



Spine

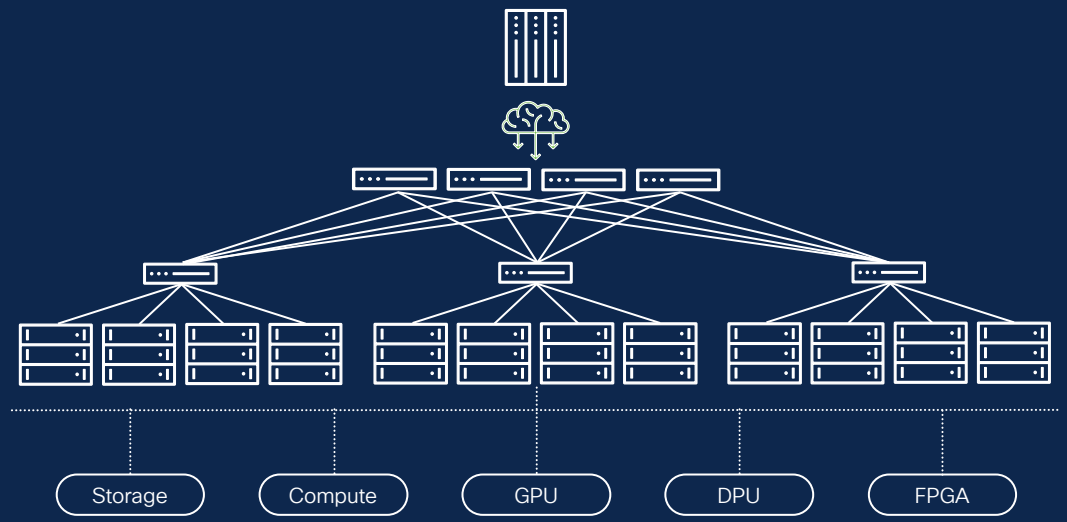
Leaf

Requirements

- 100G/400G fabrics to support edge compute and 5G
- Space-constrained environments in service-provider edge locations
- 100G/400G breakout capability allows low power and cooling
- Ready for NFV/5G adoption cycle

High-performance computing

Topology



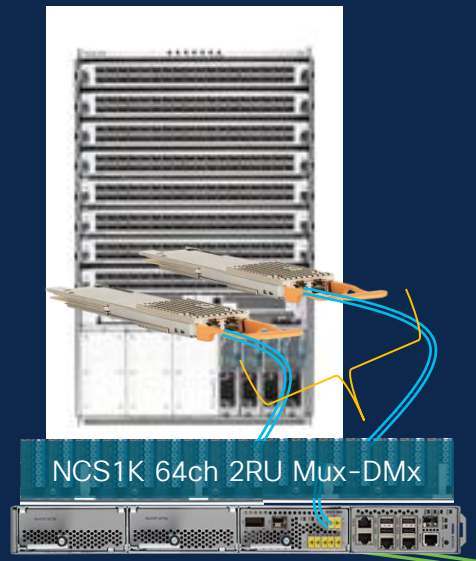
Requirements

- High-throughput, low latency connectivity fabric
- High throughput node-to-node access for compute, storage, GPU
- Network needs to allow parallel access for storage and compute nodes
- Scalable architecture to meet growing HPC demands

400G ZR/ZR+ Enables Simplified DCI

DC A

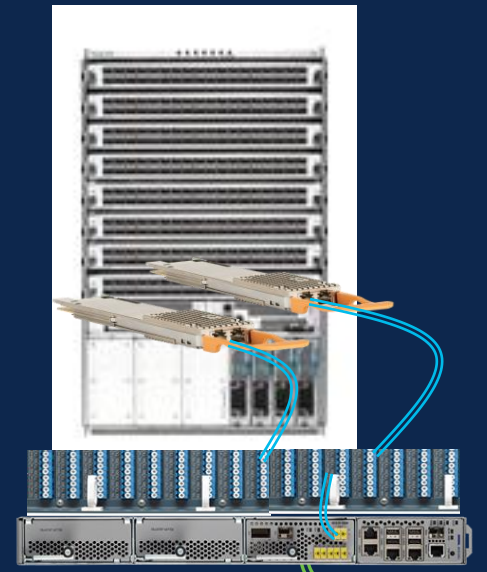
Router / QSFP-DD ZR/ZR+



OLS in a QSFP-DD ZR/ZR+ transport optimization provides optical amplification in a pluggable QSFP-DD form-factor for low and high channel count application

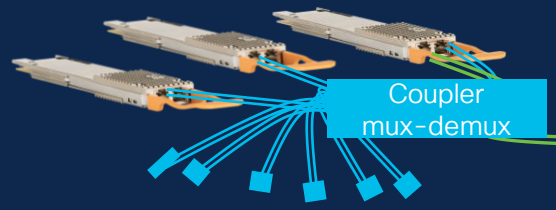
DC B

Router / QSFP-DD ZR/ZR+

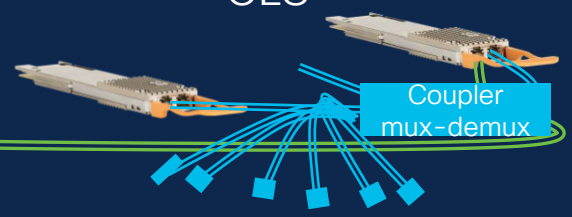


Up to 140km*

OLS



OLS



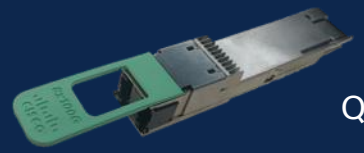
© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Public

Choosing the right optic with a path to 800G and beyond

QSFP-DD800

Roadmap to 1.6 T

High density 400G with QSFP-DD800 modules



QSFP-DD 8x100G FR

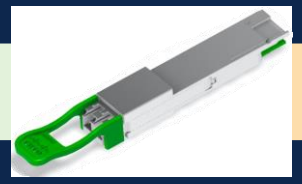
QSFP-DD 2x400G FR4

Increased density	Double port bandwidth for Single Mode Fiber links up to 2km
Investment protection	Reuse existing cabling infrastructure: **Dual** Duplex LC and **Dual** MPO-12 SMF connectors
Backwards compatibility	Connect existing pluggable transceivers: QSFP+, QSFP28, QSFP56, QSFP112, QSFP-DD
Flexible design support	800G port to port | Breakout to 400G or 100G ports
Improved sustainability	Supports over 30W of power dissipation and riding heatsink in host platform
Standards compliant	QSFP-DD 800 MSA, IEEE 400GBASE-FR4, 100GBASE-FR1

Roadmap to 1.6T

QSFP-DD extends thermal efficiency and backwards compatibility to 1.6T for coherent and non-coherent optics

QSFP-DD

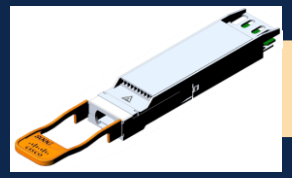


All QSFP

400G (8x 50G)

800G (8x 100G)

1.6T (8x 200G)



OSFP

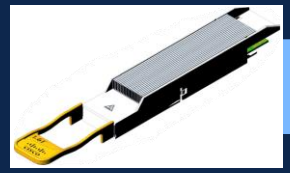
400G (8x 50G)

800G (8x 100G)

1.6T (8x 200G)

Customers can maintain backward compatibility with 400G choices

OSFP-XD



1.6T (16x 100G)

1.6T (8x 200G)

OSFP-XD is **not** backwards compatible with 400G/800G OSFP

Host side interface width

8x

16x

Market demand (time)

Summary

Network efficiencies

High speed optics portfolio

Next steps

Data center network efficiency with high-speed optics

High performance

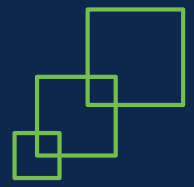
High density pluggable optics to support increasing port bandwidth for 400G, 800G

Backwards compatibly

Investment protection to connect lower speed optics and migrate data center networks seamlessly to higher speeds

Integrated capabilities

Highly integrated capabilities combined with silicon photonics improve optical module efficiency and reliability

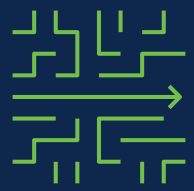


Extensive Portfolio

Connectivity options over copper, single-mode, and multi-mode fiber addressing various reaches and speeds for data center applications

Sustainability

Thermally efficiency designs leverage riding heat sink on host platform



Flexibility

Flexible design support for coherent/non-coherent and breakout applications to preserve existing investment in fiber infrastructure

Portfolio of high-speed optics for the data center

Distance	100 m	500 m	2 km	10 km	40-100+ km
100G GbE	100G-SR4 100G-BiDi 100G-SR1.2	100G-PSM4 100G-DR1	100G-CWDM4 100G-FR1	100G-LR4 100G-LR	100G ZR 100G ZR4
400G GbE	400G-SR8 400G-SR4.2	400G-DR4	400G-FR4 4X100G-FR1 8x100G-FR 2x400G-FR4	400G-LR4 4x100G-LR	400G-ER1 (40km) 400ZR 400ZR+ 400G High Power ZR+

Get started today

3 Speak with an expert

Go to: cs.co/askoptics

2 Evaluate optical transceivers compatibility and interoperability

Visit: tmgmatrix.com.com

1 Learn more about the Cisco optics portfolio

Go to: cisco.com/go/optics

The bridge to possible