

Optical Scale-Up AI Systems: Architectures, Standards, and Resiliency Strategies

Adrian Motamedi Interconnect System Architect

Fotini Karinou Director of Optical Technologies & Architecture

Azure Hardware Systems and Infrastructure (AHSI)

Microsoft

Key messages

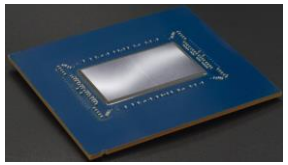
- Optics make multi-rack, multi-row AI scale-up practical.
- System value comes from reach, bandwidth density, efficiency, utilization, and resilience.
- Reliability-by-design can reduce failure blast radius and accelerate adoption.
- Open standards are needed to align a fast-moving optical ecosystem. OCI-MSA provides a scale-up specification, with OCP playing a key ecosystem role.

Microsoft Azure Hardware Innovation – Silicon to Systems

Silicon



Azure Cobalt 100



Azure Cobalt 200



Maia 100



Maia 200



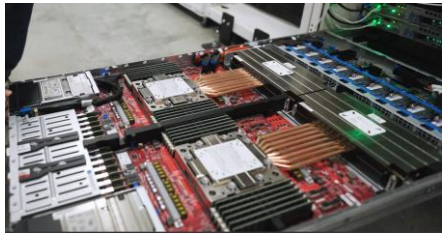
Maia 300



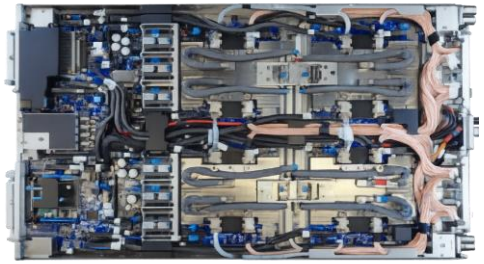
Azure Boost DPU

[Cobalt 200: Azure's next cloud-native CPU]

Platforms

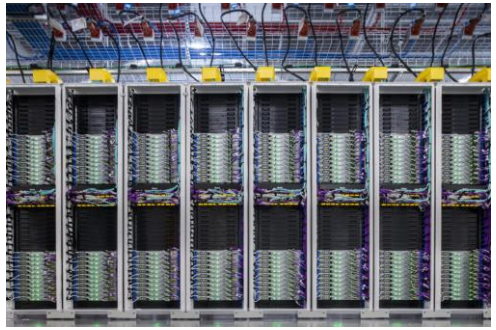


Azure Cobalt 200 Platform



Maia 200 Platform

Systems



Azure Cobalt 100 Racks

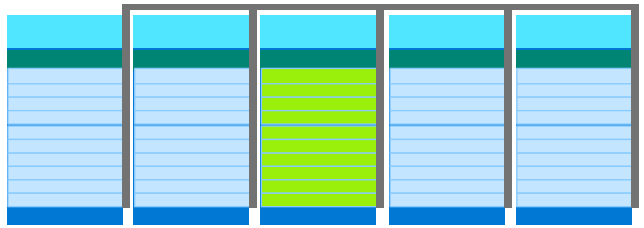


Maia 200 Rack

[Maia 200: The AI accelerator built for inference]

Optics make multi-rack scale-up practical

REACH
Move scale-up beyond a single rack and support multi-row domains.



ENERGY EFFICIENCY
Reduce pJ/bit pressure as domains grow.

BANDWIDTH DENSITY
Pack more bandwidth into accelerator and switch trays.

MULTI-RACK SCALE-UP

UTILIZATION & RESILIENCE
Enable better resource utilization with pooled capacity, better resiliency via spares, and new architectures (switch disaggregation).

The value proposition is system-level: larger scale-up domains can improve utilization, reduce stranding, and relax rack-boundary constraints.

Optical POD deployment challenges

Cost

Optical components (lasers, OEs, fibers, shuffles) materially impact system TCO, with uncertainty from evolving form factors, scale-up economics and manufacturing ramp limits.

Reliability

Millions of links. Each optical link combines many active and passive components (in-tray, in-rack, rack-to-rack), requiring quality, telemetry, operations and failover improvements at scale.

System complexity / form factors

Form factors' trade-offs balance cost, reliability and manufacturability as optical connectivity evolves across accelerator, switch and data center integration points.

DC operations

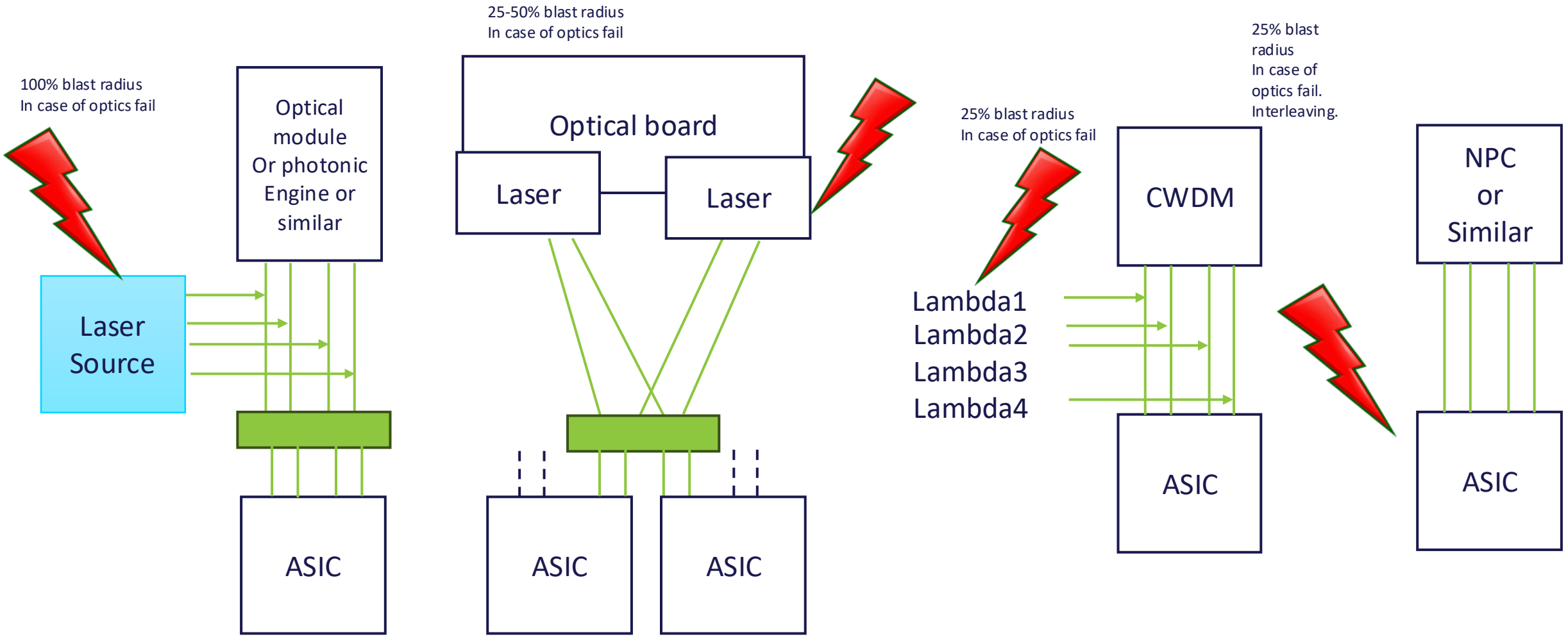
Optical pods add operational complexity: new infrastructure, fiber management, workload scheduling and RAS strategies to ensure performance and failure mitigation.

Example of need for Reliability Standardization

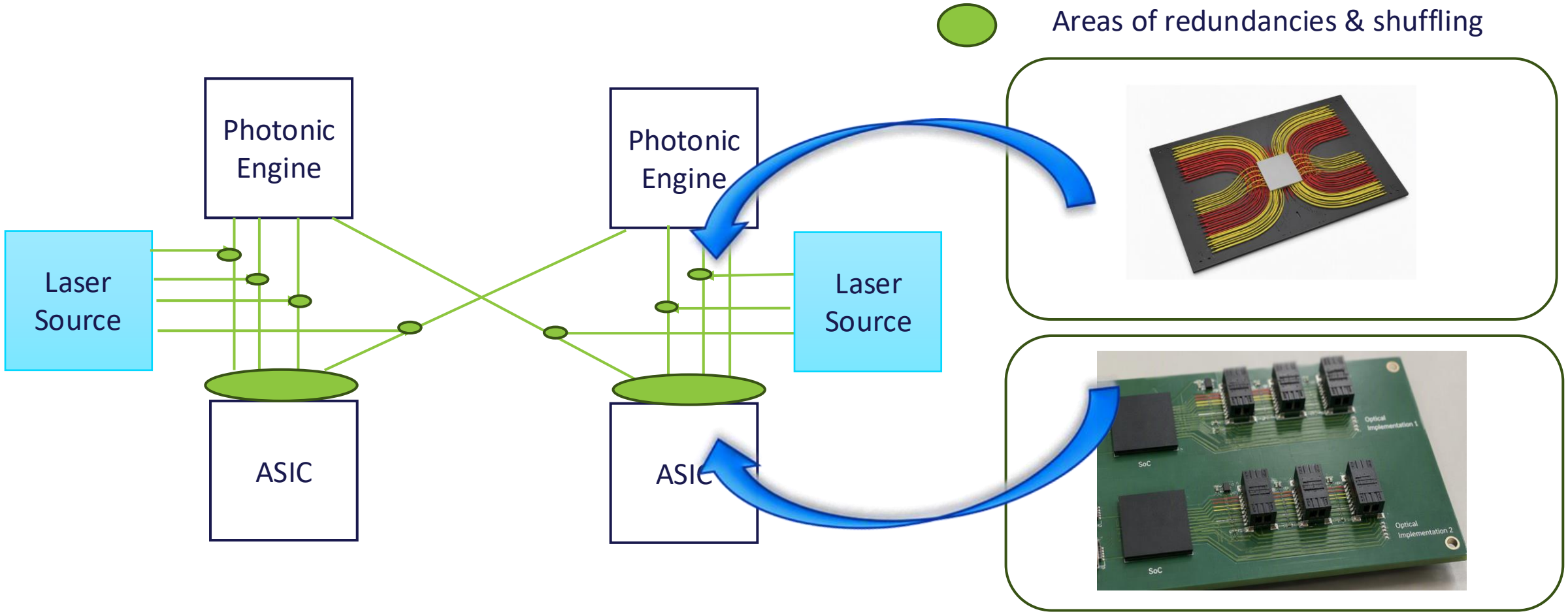
Problem statement (Redundancy and reduction of blast radius)

- Regardless of the protocol, OBO style, NPO and NPX-style connectors or DR4, OCI, etc… enable high-density electrical-to-electrical/optical interfaces but introduce complex coupled failure domains.
  - Optical engines share common laser sources and photonic integration.
  - A single failure (e.g. laser or PIC) can impact multiple channels — many lanes or lambdas can be lost.
  - Blast radius is far larger than traditional copper links, where failures are isolated and reroutable.
- System boundaries often require a single dense connector (NPC/NPO, e.g. CPX MSA, or BOB, etc.), limiting physical separation of redundancy paths.
- Industry needs a unified standardized approach to improve resiliency and drive interoperability.

Redundancy and reduction of blast radius



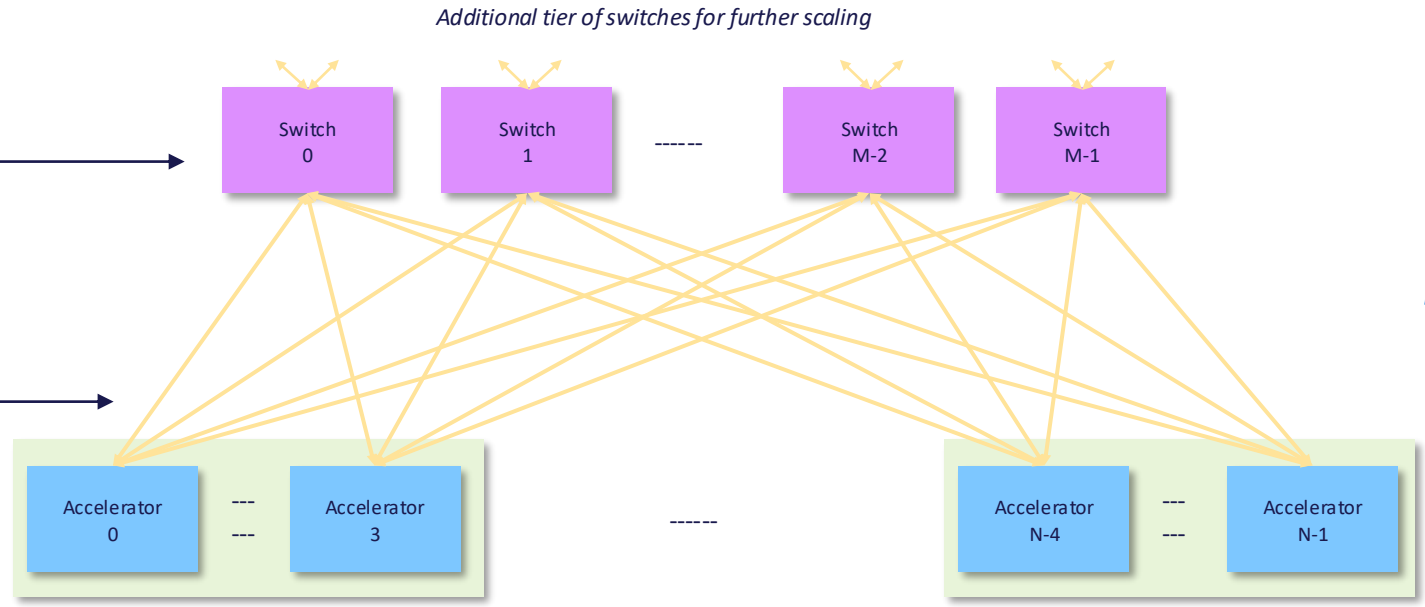
Proposed Solutions, call for action



Scale-up system topologies

High Radix switches in a single tier enable uniform communication domain

Reach, Bandwidth and Connectivity enabled by optical interconnects in Accelerator trays



Example Topology

Optics provides the reach and bandwidth density required to scale AI systems beyond the rack.

Richly connected, multi-Rack and multi row scale-up domains.

Challenges in AI scale-up interconnect specification

SCALING PACE

AI systems are evolving faster than “traditional” Ethernet interfaces cycles. OCI MSA gives scale-up optics a multi-generation path at the required pace AI systems need and the required power efficiency.

TOO MANY CHOICES

The industry needs alignment across optical technologies, rates, wavelengths, signaling, packaging, and link behavior without freezing innovation.

ECOSYSTEM INTEROP

Scale-up networks need interoperable endpoints and components so suppliers can innovate around a common PHY.

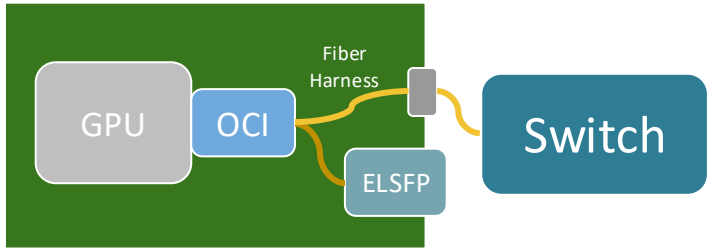
BANDWIDTH ROADMAP

Bandwidth	Spec Timeline
200 Gbps	Today
400 Gbps	2027 (?)
800 Gbps	TBD, >2030 (?)
OCI GEN1	Today

WHAT THE SPEC MUST ALIGN

- Modulation – rate
- Modulation - type
- System Architecture
- Technology - component
- Package Integration

OCI LINK: ENDPOINT TO ENDPOINT



OCI-MSA brings industry focus and accelerates AI scale-up innovation through a common multi-generation PHY that enables interoperability and a scalable supplier ecosystem.

Optical Compute Interconnect (OCI) - MSA

Founding Members

Industry alignment around an optical scale-up specification.

March 2026

Future AI infrastructure built on open standards.



Optical Scale-Up

Architecture that delivers the efficiency required to scale AI infrastructure.

Move Scale-up from In-Rack to Multi-Rack & Multi-Row

For more information and the OCI-MSA GEN1 Specification please see *www.oci-msa.org*

OCI Optical PHY Specification for Scale-up

INDUSTRY ALIGNMENT

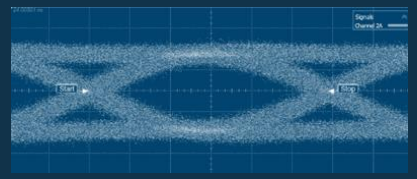
Multi-rack scale-up



OCI-MSA · March 2026

SIGNALING

NRZ



Scalable, lowest power

OPTICAL SOURCE

WDM

λ1 λ2 λ3 λ4 λ5 λ6 λ7 λ8

8 wavelengths, one standard laser set

FIBER EFFICIENCY

Bi-Di, one fiber



Reduce cost w/o radix impact

Defines optical link layer for AI scale-up

GPU / SWITCH endpoints

Physical endpoints

COMMON PHY

Endpoint interoperability

OCI MSA SPEC

Focus on tech and KPIs

SCALING

Path to multi-generation

SUPPLIER ECOSYSTEM

Multi-vendor, faster adoption

A common OCI PHY enables lower power, cost and latency, efficient fiber utilization and multi-generation scaling.

Enabling an Ecosystem for Scale-up Optics

Optically Interconnected System Stack
Performance, Reliability and Interop Testing
Rack and Pod Integration
Fiber Shuffles
Connectors
Server and Switch Platform Integration
Packaging
Form Factors
Optical Engines and Lasers
Electrical Interfaces
Optical Interfaces
Modulators, Photodetectors, Lasers
Materials

Industry Standardization Mechanisms

MSAs
OIF
OCP
APC
IEEE
ITU-T
....

Optical stack complexity is driving urgency for some level of standardization/converge.

OIF Compute Optical Interfaces (COI) Whitepaper (requirements for scale-up), OCI MSA (NRZ, WDM, BiDi)

Other efforts: Short Reach Optical Interconnects (SROI), Open SiPhotonics for AI Systems **(NEW)**, others

Remaining stack still needs ecosystem alignment!

Coordinated scope enables scale-up standards

Example focus areas

OCI MSA momentum

OIF

ELSFP for OCI

CMIS improvements

Link reliability

Link initialization

Interoperability

OCP

Physical optical trays & rack designs

Form Factors

System reliability

Connectors & interfaces

Ecosystem requirements

APC

Wafer-to-Systems Standards

SiP manufacturing specifications

CPO & SiP test and measurement

OSAT-level work

Building on the OCI specification coordinated standards efforts and complementary ownership are needed to align the optical ecosystem across the full technology stack.

Call to Action: Make OCP the home for scale-up standards

From hyperscaler requirements to open implementation guidance

JOIN
Build on OCI MSA and join adjacent standards work.

DEFINE
Create complementary OCP workstreams for trays, form factors, reliability.

COORDINATE
Align OIF, OCP, APC, others, with OCI MSA on complementary scope.

DELIVER
Turn requirements into open, testable specs.

OCP should actively participate in scale-up standards so the ecosystem can move faster together.

Thank You!

