OCP SPECIAL FOCUS: PHOTONICS

The Interconnect Inflection Point: Evolving Intra- and Multi-Rack Scale-Up AI Networks

Akarsh Bhatnagar

Product Line Manager, Celestica

OCP PLATINUM Member

Agenda

1. System-Level Convergence
2. XPO: The Interconnect “Swiss Army Knife”
3. Closing The Critical Reach Gaps
4. Promising Technologies
5. Interconnect Blueprint
6. CALL TO ACTION: Get involved with the open standards

Interconnect Bottlenecks

- As AI/ML clusters expand from single-rack to multi-rack architectures, the scale-up interconnect has become the primary design bottleneck.
- Power and cooling
  - Front-panel thermal limits and excessive power consumption caps module packaging density.
- Physical reach limits
  - Sub-2m copper reach; can no longer reliably span a full rack
- Interconnect latency
  - Reconditioning degraded signals with DSPs

System-Level Convergence MSAs

- Industry is converging on complementary open specifications spanning the full path from front panel to silicon packaging, preserving multi-vendor ecosystems

Pluggable Density Pivot

XPO (eXtra-dense Pluggable Optics)

Next-generation form factor redefines front-panel interconnect density:

12.8T Module Capacity: Enables up to 204.8T total bandwidth density within a single 1RU switch chassis.

Liquid Cooling Integration: Direct blind-mate cold-plate cooling engineered for modules requiring up to 400W+ per slot.

On-Board & Packaging

Open CPX & OCI

System-level standardization targeting near-die and silicon-centric connectivity

Open CPX MSA: Standardizes socketed, pluggable optical & copper engines for Near-Packaged Optics (NPO), creating an open multi-vendor ecosystem.

OCI MSA (Optical Compute Interconnect): Silicon-centric scale-up paradigm taking slow-wide approach using DWDM supporting bidi links.

XPO: The Interconnect “Swiss Army Knife”

- One footprint, many engines: Swap engines as per requirements, without a redesign
- Retains field-serviceability and hot-swappability while achieving CPO-class density

DAC
Direct Attach Copper

Shortest reach, lowest power, lowest cost - in-rack only

AEC
Active Electrical Cable

DSP/retimer-based; extends copper reach for intra-rack links

ACC
Active Copper Cable

Linear-equalizer copper; ~2.5W, short reach, minimal power

XPO

Shared mechanical & thermal envelope, matched to reach and cluster tier, eliminating the redesign cost of switching interconnect types

LPO / LRO
Linear (Re)drive Optics

DSP-free or lightly retimed optics; lower power, shorter reach

FRO
Fully Retimed Optics

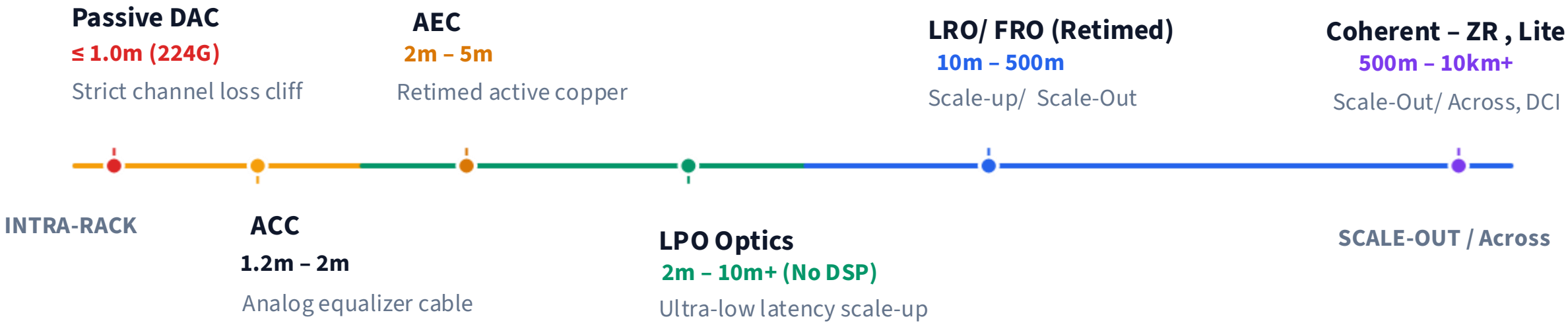
Full signal reconditioning for longer reach, higher robustness

ZR / Coherent-lite
Coherent/ Coherent-lite Engines

long-reach coherent optics for scale-across fabrics

Closing the Critical Reach Gaps

- No single electrical or optical engine covers the full reach continuum a rack-scale AI cluster needs.
- **Why it matters:** Prevents architectural & reach gaps as AI clusters expand across racks.



Power Consumption

Promising Technologies

MicroLED Optical Engines

- A “wide-and-slow” architecture utilizing arrayed MicroLED emitters for massively parallel optical channels.
- Can be integrated directly onto silicon, operating at sub-picojoule per bit energy efficiencies.

Active Radio Cables (ARCs)

- Employs RF/microwave guided-wave media as an alternative to copper and lasers.
- Standard pluggable connectors & formfactor.
- Achieves ultra-lightweight, high-frequency intra-rack connectivity with lower loss.
- Expected reach 2m-10m

Interconnect Blueprint: Intra- to Multi-Rack

- No single technology wins across the whole fabric.
- The winning architecture balances density, serviceability, and thermal efficiency differently at each tier.
- Liquid cooling Next-generation switch integrate dripless quick-disconnect liquid manifolds to handle module thermal loads.

Intra-Rack (Scale-Up)	Rack-to-Rack (Scale-Up/Out)	Row / Multi-Rack (Scale-Out)	Row / DC (Scale-Across)
DAC / ACC / AEC Emerging : ARC/ MicroLED	**CPO / NPO / Open CPX / XPO**	**XPO pluggables (LPO/LRO/FRO)**	**Coherent - Lite, ZR/ZR+**
Density & power dominate; serviceability is secondary at short, fixed reach.	Thermal efficiency dominates; silicon-adjacent optics cut electrical loss.	Serviceability & density balance; liquid-cooled pluggables preserve field swap.	Reach dominates; coherent-lite and ZR/ZR+ optics extend the fabric.

Breaking the Validation Deadlock

Production Adoption Requires Proof at Scale ⇒ ⇐ Proof at Scale Requires Real Deployments

Phase 1: Managed Risk

XPO & Socketed NPO

Liquid-cooled. Swap-ability/ serviceability isolates optical failure risk.

Minimum Redesign

Phase 2: Early Deployments

Interop & Qual

Gather multi-vendor qualification data across temperature, BER margin, and liquid cooling manifolds

Multi-Vendor Supply Resilience

Phase 3: Scale

Optical Compute Fabrics

Graduate qualified optical engines interconnects at volume.

Efficient Cluster Architecture

Multi-vendor MSAs eliminate proprietary lock-in, enabling rapid volume scaling and supply chain resilience

Call to Action

- Join the XPO, Open-CPX & OCI MSA
  - **Contribute to the specifications and the roadmap**

- Pilot XPO/ Open-CPX ready platforms
  - **Bring liquid-cooled, pluggable interconnects into your next 102.4T/204.8T design cycle**

- Potential OCP discussion areas:
  - **MSA for emerging technologies**
  - **Qualification and validation methods**
  - **Multi-vendor interoperability**

Thank You!

