Democratizing AI Infrastructure: Accelerating Innovation at Scale

Asad Khamisy

Senior Vice President and General Manager, Core Switching Group, Broadcom





OCP GLOBAL SUMMIT

October 12-15, 2026
San Jose, California

BROADCOM

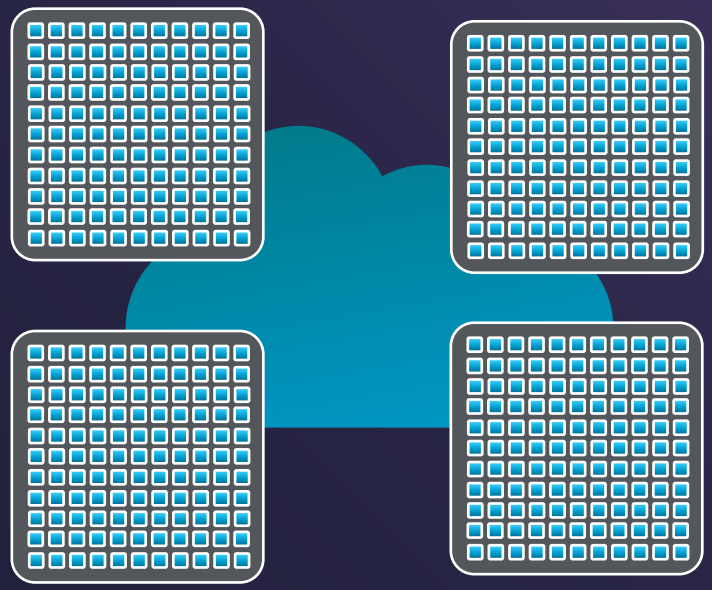
Scaling AI Infrastructure Through Open Innovation

Asad Khamisy, Ph.D.

SVP/GM, Core Switch Group, Broadcom

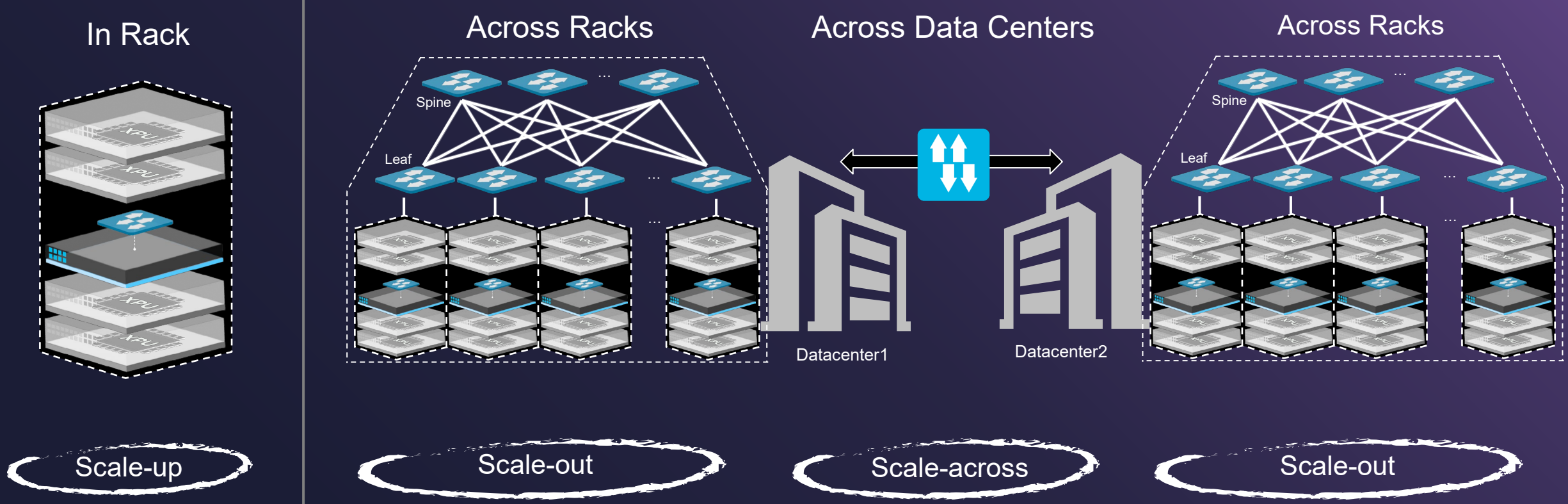
Oct 12th, 2026

A Very Very Large Distributed Computing System



100K+ XPU Clusters

AI Scale-up, Scale-out and Scale-across Networking



Complete Coverage of HPC and AI with *Ethernet*

HPC	Scale-Up	Scale-Out	Scale-Across

SWITCH



Tomahawk Ultra
51.2 Tbps



Tomahawk 6
102.4 Tbps



Jericho 4
51.2 Tbps

SAI / SONiC / SDK

NIC



SUE-T + ESUN (OCP)

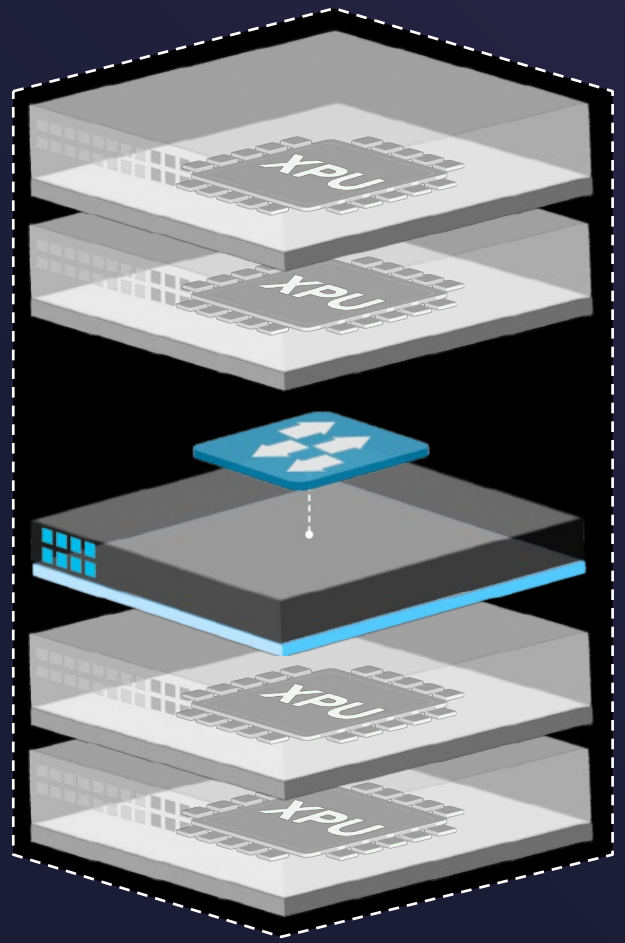


Thor Ultra
800G

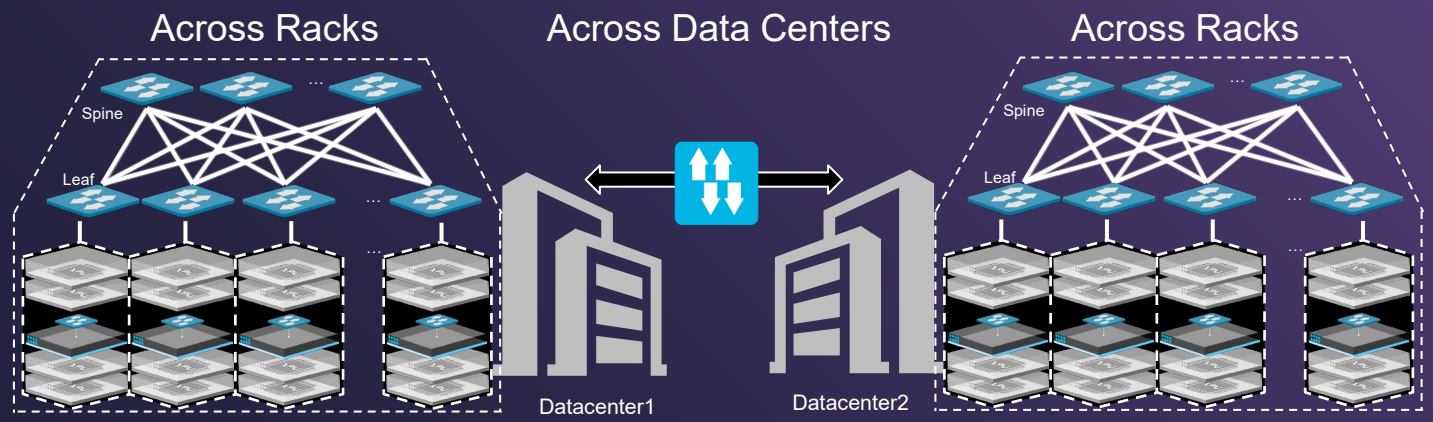
ESUN / Enhanced RoCE Stack

AI Scale-up

In Rack



Scale-up



Scale-out

Scale-across

Scale-out

Scale-up Connects Tightly-Coupled Compute

ARCHITECTURE

Open Standards

PHYSICAL MEDIUM

Copper vs Optics

Scale-up Connects Tightly-Coupled Compute

ARCHITECTURE

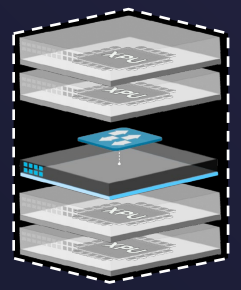
Open Standards

PHYSICAL MEDIUM

Copper vs Optics

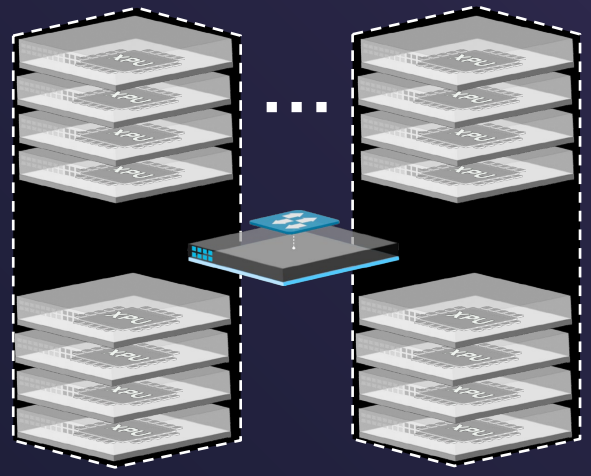
Ethernet for In-Rack and Multi-Rack Scale-Up

Today



64 - 144 XPUs
Single Rack

Tomorrow



1K - 2K XPUs
Multiple Racks

Common Ethernet Operational Foundation

Within a rack → Across racks
Electrical + optical connectivity

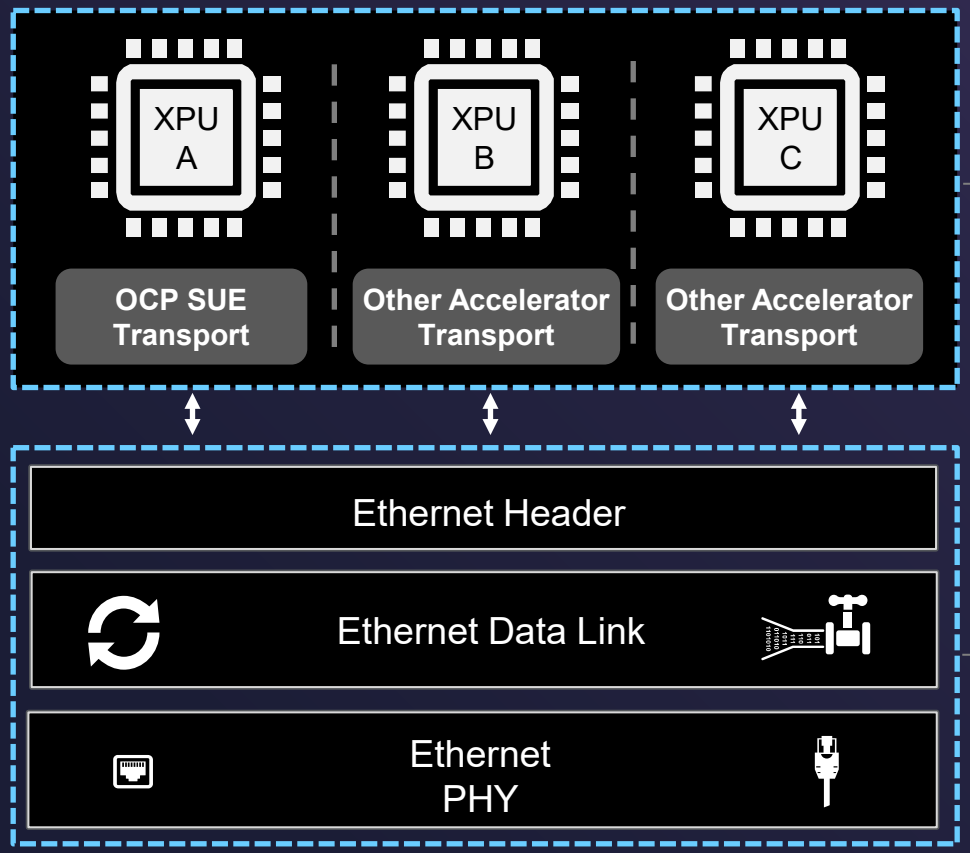
Common operations
Telemetry • Diagnostics • Automation

Common infrastructure
Cabling • Optics • Qualification • Sparing

Across the AI fabric
Scale-up • Scale-out • Scale-across

Ethernet from Scale-up to Scale-out to Scale-across

Scale-up Architecture: SUE-T + ESUN



SUE-T: Co-optimize Transport and Compute

- Push vs pull memory access
- Memory access ordering model
- Transaction packing
- Reliability
- Load balancing
- Scheduling

SUE-T spec publication target: 2H2026

Ethernet for Scale-up Networking (ESUN)

- Highly reliable & lossless Ethernet
- Efficient support for small messages
- Efficient multi-hop scale-up topologies

v1.0 spec published February 2026

ESUN announced at OCP Global Summit 2025

175+ members joined and growing

ESUN racks are now in production

AMD ARISTA arm BROADCOM cisco HPE MARVELL Meta Microsoft NVIDIA

Ethernet Scale-up: From Concept to Production!

5+ Hyperscalers running Ethernet Scale-up in production

10+ partners building Scale-up Ethernet based racks

Multiple custom & merchant XPU Scale-up Ethernet racks to be in production in the next 6 months

Ethernet as the *Scale-up Fabric*

Ethernet *Today*



512 x 100G-PAM4
Single hop Scale-up: 256

Ultra-Low Latency
250 ns

51.2T@64B
77 BPPS

IP or Optimized Ethernet Headers

In-Network Collectives

Link Layer Retry

Credit-Based Flow Control

Performance

Efficiency

Reliability

Ethernet Inherent Efficiency: Variable Packet Sizes

Use case	Link Efficiency
Unpacked 256B posted write	92%
2 packed 256B posted writes	96%
4 packed 256B posted writes	98%
Native 2kB block transfer	99%

Variable-size packets are inherent to Ethernet

- Gives more optimization tools to XPU and system designers
- Switch achieves full throughput for all packet sizes

Optimized UFH header and Reduced IPG

- Next version of OCP ESUN spec

Scale-up Connects Tightly-Coupled Compute

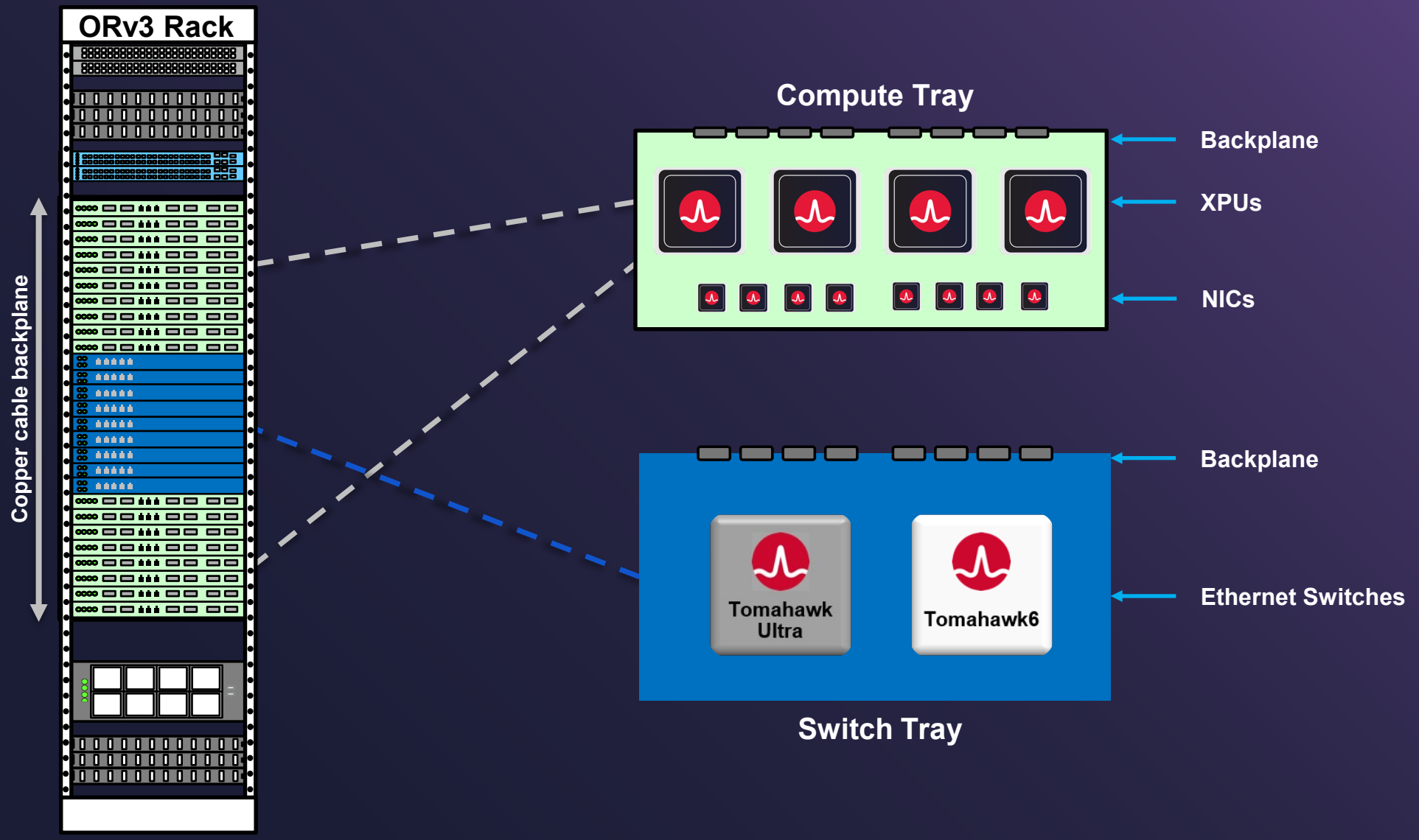
ARCHITECTURE

Open Standards

PHYSICAL MEDIUM

Copper vs Optics

In-rack Scale-up



In-rack Scale-up: Doubling Bandwidth with Copper Backplane

CURRENT GEN

1.04Pbps
TOTAL RACK BANDWIDTH

14.4Tbps per XPU
72 XPUs

512x200G PAM4
9-12 Ethernet switches

Twinax density: **1024 per OU at 26 AWG**
XPU to Switch copper reach: **~2m**
SerDes power: **7 pJ/bit**
Total rack power: **~250 kW**

Bandwidth limited by backplane density
Production volume with Tomahawk6 today

Possible NEXT GEN (2x BANDWIDTH)

2.08Pbps
TOTAL RACK BANDWIDTH (2x)

28.8Tbps per XPU
72 XPUs

576x400G PAM6
9 Ethernet switches

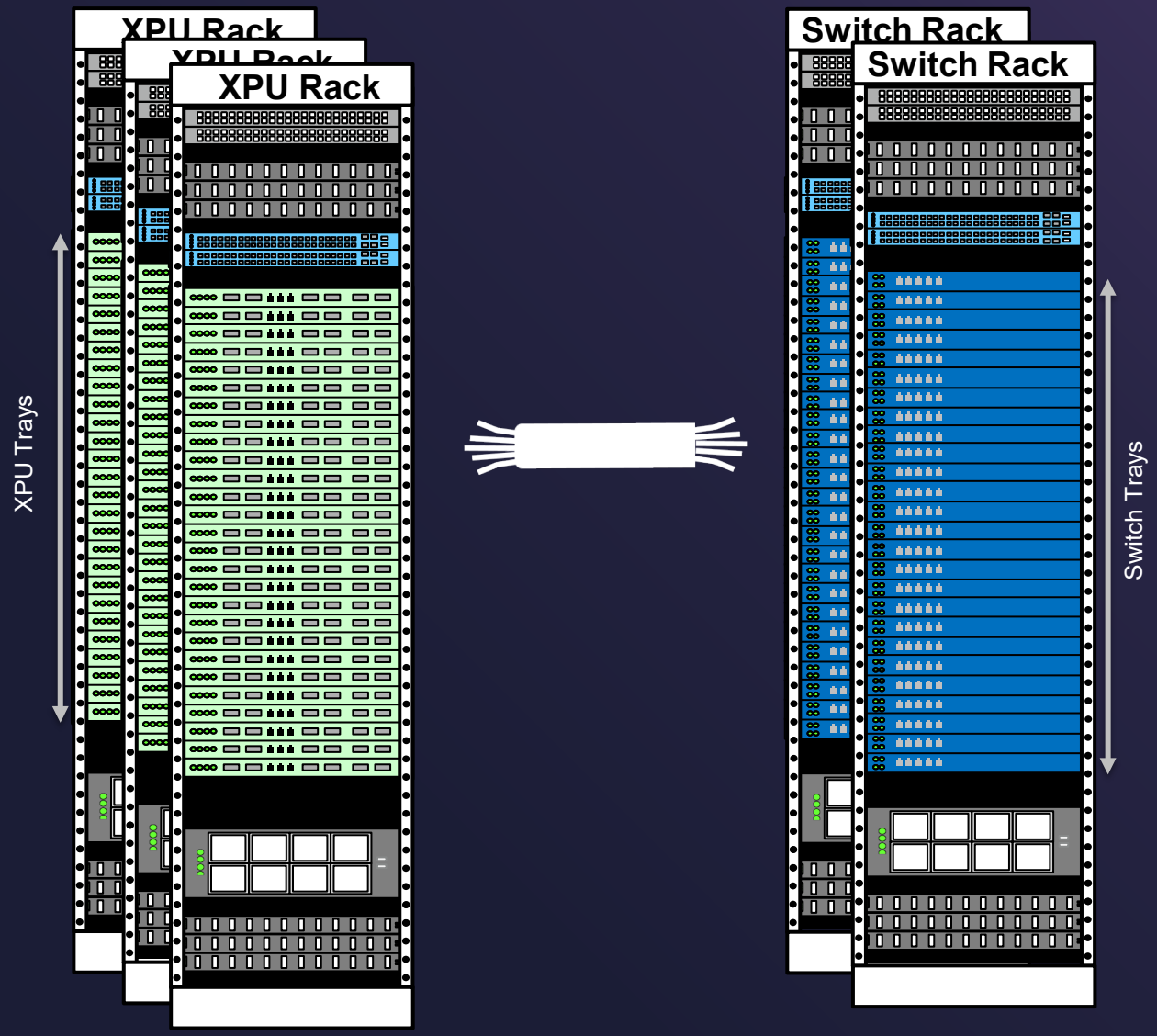
Twinax density: **1024 per OU at 26 AWG**
XPU to Switch copper reach: **~2m**
SerDes power: **7 pJ/bit**
Total rack power: **~500 kW (2x)**

2x Density, 2x Bandwidth

Scaling to 1K and beyond

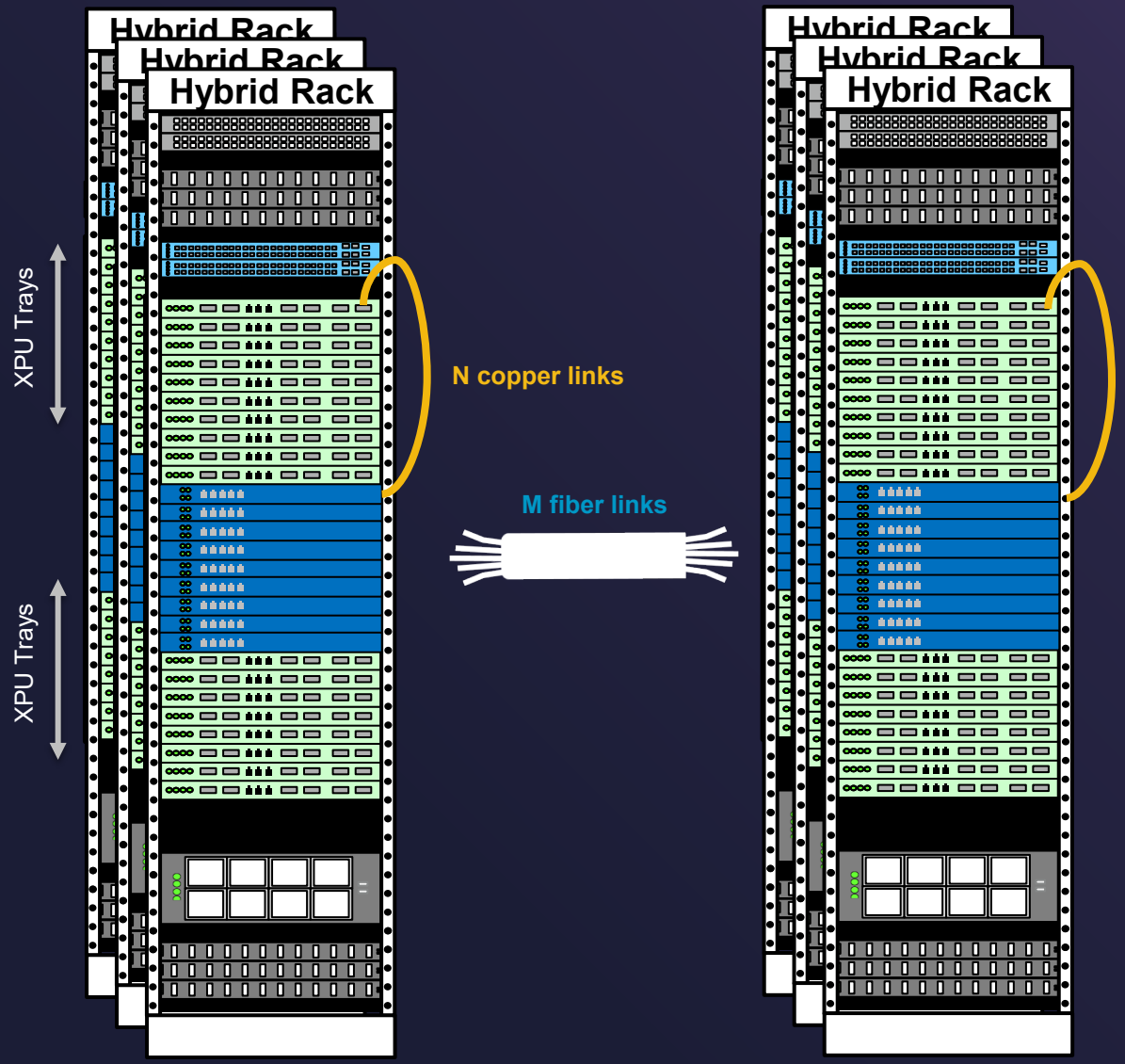
Expanding Scale-up Interconnect Beyond Single-Rack Limits

Disaggregated Rack Scale-up with Optics



- Larger scale-up domain (e.g. 576)
- Lower power per rack
- Reduced weight per rack
- Higher cost optics compared to copper
- Higher latency across racks
  (5ns / fiber meter)

Disaggregated Rack Scale-up with Copper and Optics



- Preserve large scale-up domain
- Breaks scale-up into “low latency domain” and “higher latency domain”
- Reduces optics costs

Introducing Optical Compute Interconnect (OCI)

Founding Members

Industry alignment on optical scale-up

OCI-MSA

Est. March 2026

Future AI infrastructure built on open standards

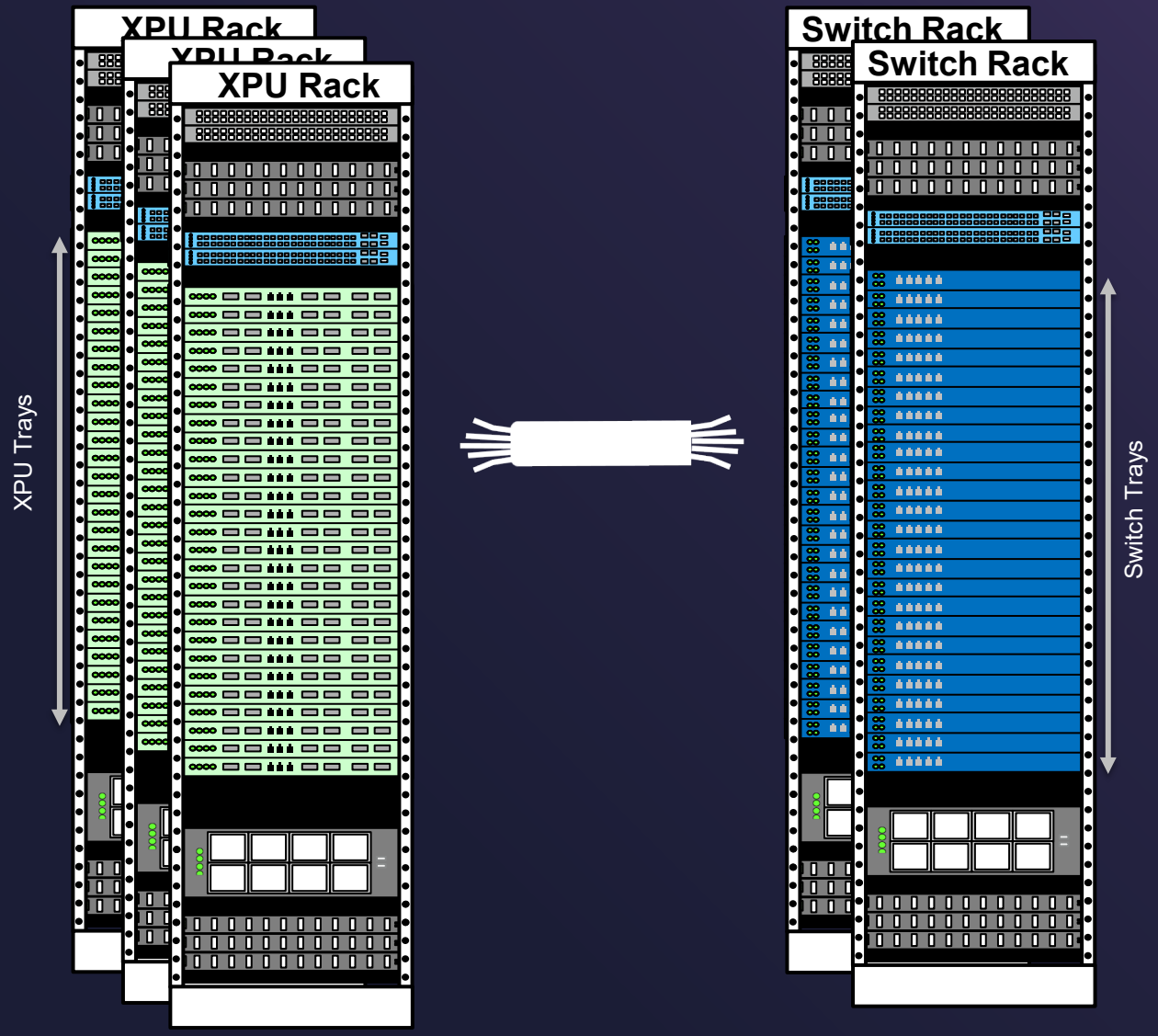


Optical Scale-Up

Architecture that delivers the efficiency required to scale AI infrastructure

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

Disaggregated Rack Scale-up with OCI



- Preserve large scale-up domain
- Same power as copper
- Architecturally defined to enable cost & reliability to approach copper KPI’s

Total Link Power and Reach Comparison - 200G

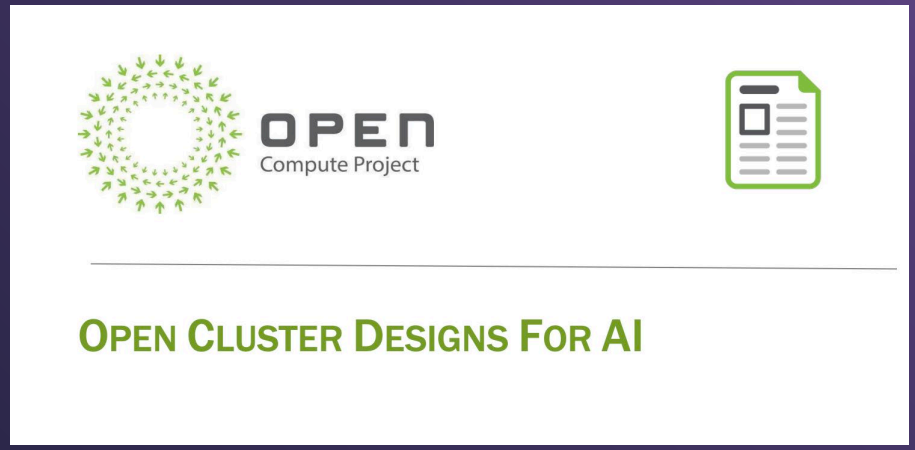


Interconnect	Total Link Power (pJ/bit)	Reach
DAC	7	2m
Integrated OCI	7	500m
DR LRO Front Panel Pluggable	23	500m

OCI – the reach of Front Panel Pluggable, the power of Copper

OCP: Open Cluster Designs for AI

- Delivers open, multi-vendor blueprints to democratize AI infrastructure.
- Modular designs from individual racks (XPU pod) to datacenter fabrics (XOC-N).
- V1.0 blueprints published.
- Scope includes scale-up, scale-out, front-end, storage, and management networks.
- 2nd generation designs under development – please join!



53 Member Companies

BROADCOM
Celestica
hedgehog
SYSTEM STACK



Thank you

Copyright © 2026 Broadcom. All Rights Reserved. The term “Broadcom” refers to Broadcom Inc. and/or its subsidiaries.


BROADCOM®