Oriole

PRISM: The World's First Full-Photonic Network for AI

Compute is Fast, But Your Network Might be Holding You Back

In today's High-Performance Computing (HPC) and Distributed Deep Learning (DDL) systems, it's not the processors slowing things down - it's the memory and the network. Despite huge leaps in hardware computing power, real-world performance often hits a wall, with only a small percentage of peak performance being achieved due to network bottlenecks.

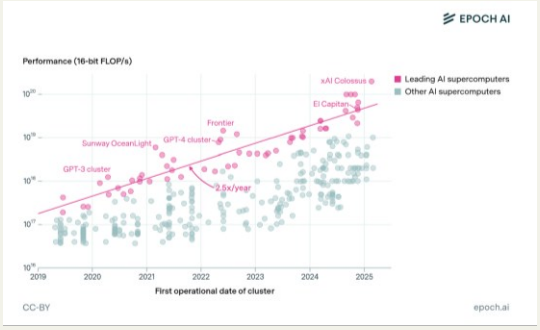


Figure 1: Performance of leading AI supercomputers has doubled every nine months

Network-related delays can account for up to 40–60% of total training time in deep neural networks (DNNs), even with relatively small-scale setups of around 128 nodes. And with the number of neural network parameters doubling every few months, the pressure on interconnects only increases. The extent of this network overhead largely depends on your scaling method.

Weak scaling boosts throughput by adding more workers, typically through data parallelism, which keeps compute and communication times stable as you scale. It's popular because it plays nicely with current network limitations. So long as batch sizes stay large, modest bandwidth (Gbps range) can handle the load. But it hits a wall with very large models since it doesn't always improve training efficiency, nor does it reduce memory demands.

Strong scaling, on the other hand, speeds up each training iteration by slicing up the model itself using model parallelism. This approach helps with memory but demands far more frequent and heavier communication, requiring ultra-fast (multiple Tbps per node), low-latency networks.

Current electronic packet-switched (EPS) systems can efficiently support strong scaling only within a scale-up domain that hosts up to 72 devices (without incurring significant network overhead).

From a Packet-Switched Network Back to Circuit-Switched Network

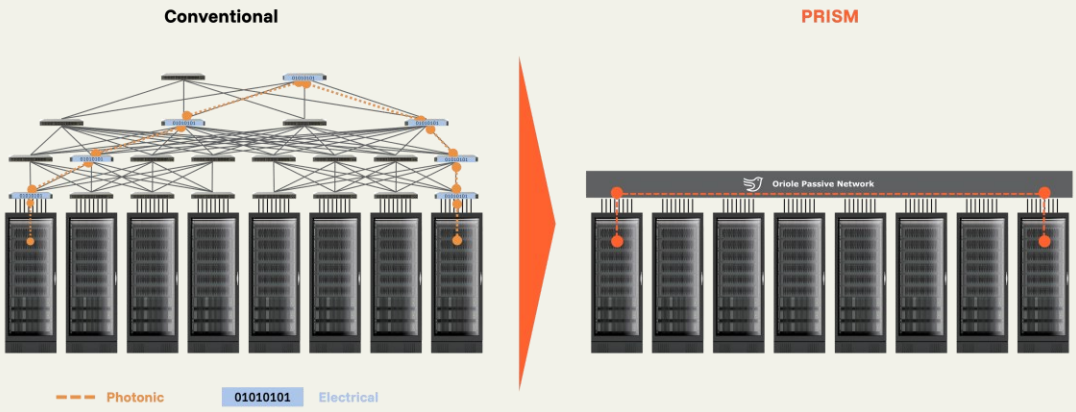
One way to push for strong scaling is to boost the capacity of electronic packet-switched (EPS) systems—but that's easier said than done. As I/O bandwidth and transistor density hit physical limits, power and cost climb fast, making it harder to sustain performance gains. Building bigger switches or duplicating entire EPS networks can add bandwidth, but with serious trade-offs in energy, cost, and complexity.

Limited-connectivity topologies like Dragonfly or Torus offer better scalability on paper but often lead to inefficient bandwidth use and high latency, as data must travel through many intermediate nodes to reach its destination.

A promising alternative is replacing electronic interconnects in packet-switched networks with optical ones. EPS is widely used because it's flexible, scalable, and handles mixed traffic well, but that flexibility comes at a cost. These networks rely on complex control logic like congestion control and forwarding/routing tables, and queueing, which cannot be implemented directly with optics due to high data processing and storage demands. As a result, asynchronous optical packet switching (OPS) doesn't meet the demands of high-performance workloads.

That's where optical circuit switching (OCS) comes in, a simpler, more deterministic approach that plays to the strengths of optical hardware. OCS separates the control and data planes: routing and scheduling are handled by a logical controller (centralized or distributed), while the optics focus on moving data fast. This leads to lower energy use, less complexity, and better scalability. Communications are pre-arranged through "logical circuits," which the controller translates into point-to-point requests. After a brief hardware reconfiguration, data flows seamlessly, and pipelined scheduling keeps things moving without delay.

OCS networks have serious potential, but a few big hurdles still stand in the way of wide adoption in large-scale data center and HPC systems. The technology has often struggled with fast, large-scale traffic scheduling, hardware setup and reconfiguration times (of transceivers and switches), and tight synchronization across thousands of devices. More than anything, deploying an OCS network means rethinking the entire network architecture. Oriole Network's novel approach was developed to overcome these exact challenges.



The Keys of PRISM

Meet the Photonic Routing Infrastructure for Scalable Models – PRISM: the first large-scale, high-capacity, full-bandwidth architecture built from the ground up for AI data center networking (DCN), high-performance computing (HPC), and distributed deep learning (DDL) workloads. It's designed to break through the limits of today's interconnects and keep up with tomorrow's demands. Here's what sets it apart:

1. **Handles Any Traffic:** Nanosecond-level switching using wavelength and space switching. The combination of Time, Wavelength, and Space Domain Switching (TDM, WDM, and SDM) allows for large-scale networking with fast circuit configuration. It supports small bursts of optical data transfer that work efficiently for large and small data transfers, the so-called elephant and mice flows. Because of the switching speed, the system supports both deterministic collective communications and non-deterministic dynamic traffic.
2. **True All-to-All Connectivity:** Offers port-level all-to-all communication where any endpoint can reach any endpoint.
3. **Lower Power and Temperature:** With entirely passive interconnects and switches, the network core stays clean and efficient (it does not dissipate power or require cooling). Complexity is greatly reduced while all the control moves to the edge.
4. **Resilient and Reliable by Design:** No single point of failure. Each node has multiple paths to every other, so even if something breaks, communication continues.
5. **Built to Scale:** Handles systems with up to one million endpoints with only 1-hop diameter, enabling support for increasingly complex and distributed workloads.
6. **Designed With AI Workloads in Mind:** Purpose-built algorithms for collective communication strategies tailored for optical circuit switched networks, enabling schedule-less, contention-free data

These operations are completed in just a few steps, even at maximum scale, dramatically reducing latency and keeping GPUs fed without stalling on communication.

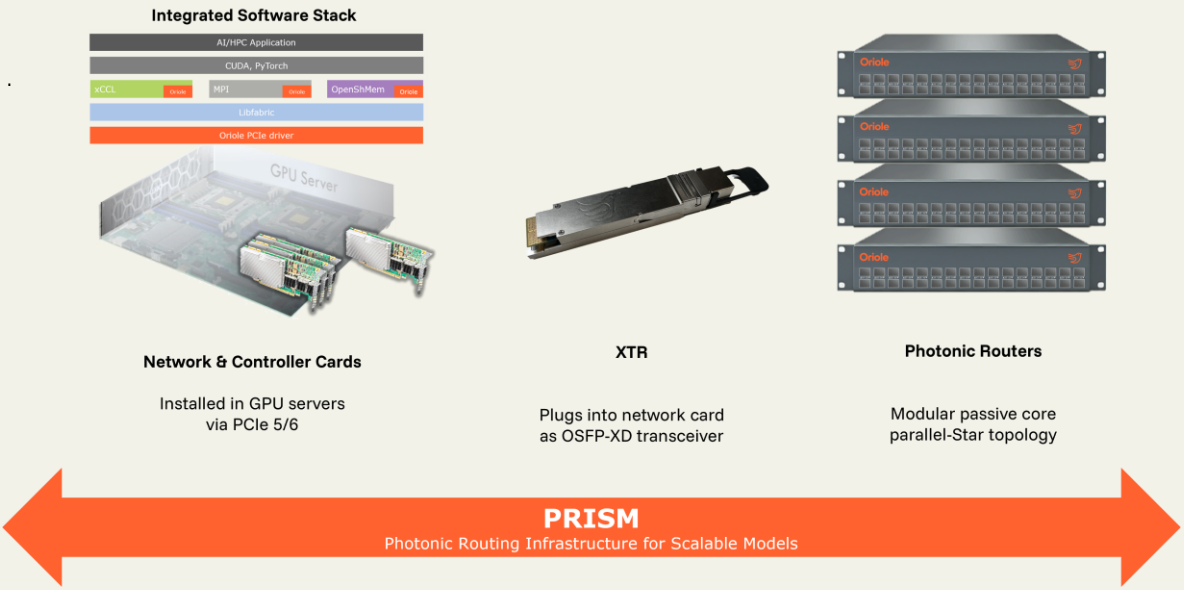
With this combination of performance, scalability, and resilience, PRISM provides a solid foundation for the next generation of high-performance data centers.

What Does the Network and Infrastructure Look Like?

Here's what PRISM looks like under the hood: a full-stack, all-optical fabric purpose-built to connect GPUs at scale with speed and simplicity. While the architecture is a bold change in the way HPC/AI networks are currently run, the components are designed to fit right into AI data centers, offering high performance and reduced complexity.

Software Stack: We are developing a custom xCCL plugin, compatible with NCCL and other collective communication libraries, that integrates directly with our network stack. This plugin replaces the default transport layer used by these libraries, typically Ethernet or InfiniBand, with direct hooks into PRISM's optical circuit scheduler and collective communication logic. By tailoring collective operations to PRISM's architecture, the plugin ensures that data flows are optimized for the photonic topology, unlocking the full performance of distributed training at scale.

High-Performance Network Interface Card: Using the PCIe interface, this card delivers 800 Gbps of bandwidth in a compact form factor. Built to fit most standard servers, it provides full bisection bandwidth within a cluster and manages connectivity through a pluggable interface. A dedicated set of controller cards handles network transfers.



access and coordination, granting transfer requests, configuring circuits with as fast as sub-microsecond update rates, and monitoring network health. All this with full redundancy and pipelined processing for real-time responsiveness.

XTR - Integrated Photonic Switch and Transceiver: This pluggable module will do the work of a transmitter, a switch, and a receiver all in one! It plugs directly into the network card and handles wavelength and route selection on the fly. It combines an optical engine and integrated ASIC, enabling connectivity to as many as one million nodes.

Passive Router: The heart of the network is a power-free modular optical router. It supports both top-of-rack and central deployments and offers dual-path resilience. By separating data and control planes while keeping everything optical and passive, it delivers ultra-low latency with zero energy consumption in the core.

What Does This Mean For You?

Lower Power & Cost in a Power-Limited Industry

As NVIDIA CEO Jensen Huang emphasized during GTC 2025, "AI revenues are power-limited." In today's large-scale AI clusters, the ability to generate value isn't just tied to how much compute you have, it's tied to how much of it you can power efficiently. Every watt matters.

Oriole Networks takes a fresh approach to solving the GPU bottleneck problem with a tightly co-designed system that brings the network architecture, communication model, and scheduling under one roof. Our architecture will significantly reduce power consumption of both the network and compute parts of the cluster and will virtually eradicate (<1%) the communication-dependent idle time of GPUs. It replaces expensive, complex, and power-hungry switches with fewer tunable transceivers and affordable routers, making the system less complex and more reliable.

Low GPU utilization is also one of the sneaky killers of efficiency in HPC and distributed deep learning (DDL) systems. You've got these incredibly powerful (and expensive) GPUs sitting idle, not because they're waiting on data or compute, but because they're stalled by the network. In large-scale training and inference or simulation workloads, GPUs often spend a surprising chunk of time just waiting for data to arrive from other nodes. This becomes a bigger problem as systems scale. Reducing that overhead is now essential not just for performance but for profitability.

In inference, especially with the latest Mixture of Experts (MoE) models, models dynamically route each token to a small subset of expert networks, making inference highly dependent on low-latency communication between compute nodes. Traditional optical interconnects introduce millisecond-scale delays during expert selection and data movement, creating a bottleneck for real-time inference. Our contention-free solution enhances performance by avoiding these issues altogether. While MoE architectures rely heavily on specialized all-to-all

operations, inefficiencies in EPS (congestion) can be addressed through complex pipelining, as demonstrated by DeepSeek. However, our hardware inherently avoids these problems. Our network reconfigures at a nanosecond scale, enabling just-in-time expert activation without pre-allocated routes or static topologies. This dramatically improves expert utilization, reduces tail latency, and allows inference to scale efficiently to larger expert counts. By minimizing communication delays at sub-microsecond granularity, our architecture unlocks the full performance potential of sparsely activated MoE models in production environments.

This is where PRISM's fully photonic, nanosecond-level switching becomes so impactful. By reconfiguring optical circuits so quickly, our network can minimize the idle time of your GPUs and help you get the most compute from them. There's no waiting around for the network to catch up, and less need to overprovision links just to cover peak loads. This result is better GPU utilization, lower energy waste, and more performance per watt across the system compared to EPS systems.

Scale Your AI Cluster Without the Performance Hit

One of the biggest challenges in scaling AI infrastructure is maintaining consistent performance as you add more nodes. In conventional networks, congestion, packet loss, and variable average and tail latency can all degrade application throughput as systems grow. Thanks to the large bandwidth, nanosecond-level reconfiguration, a scheduled and synchronously deterministic fully photonic network like PRISM can minimize GPU idle times and mitigate the performance issues that often come with scaling.

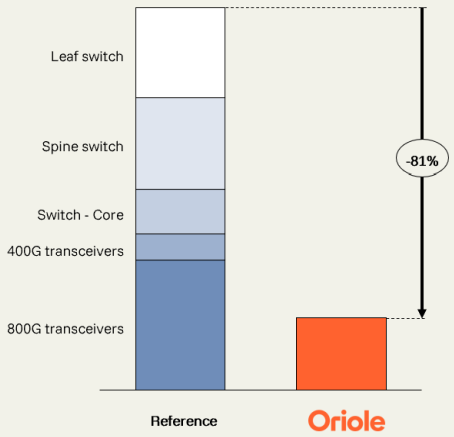
PRISM delivers a **step change in network performance** over conventional solutions and **unlocks AI training and inference performance**, all while using less power and costing less to build. You're not just getting better speeds; you're getting predictable performance at scale, which is exactly what HPC and DDL workloads need.

A comparison between a traditional and an optical AI cluster with 32k nodes

Architecture Overview

	Reference architecture		Oriole architecture
Network layers	3		1
Leaf switches	1024		0
Spine switches	1024		0
Core switches	512		
Total switches	2,560	eliminated!	0
Node transceivers	32,768		32,768
Switch transceivers	163,840		0
Total transceivers	196,608	83% reduction	32,768

Power Consumption [MW]



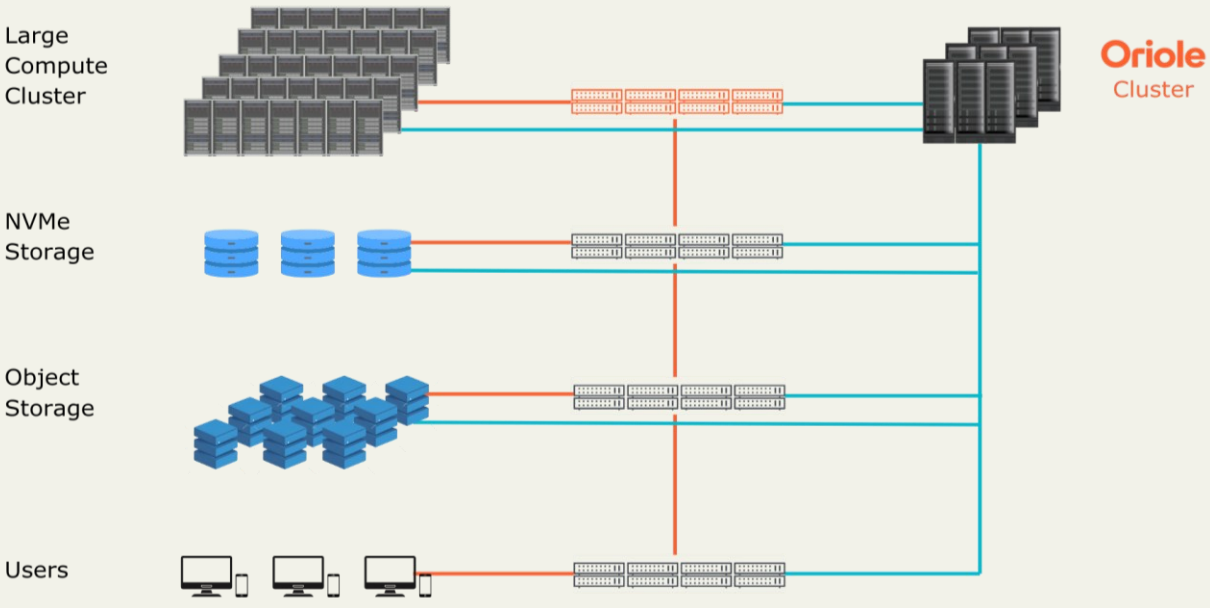
Reduced Complexity & Increased Reliability

By eliminating switches entirely from the network's core, we strip out a massive amount of electronic equipment that traditional EPS networks depend on. That means fewer ASICs, fewer power-hungry chips, and far less heat to manage. In addition, by eliminating the EPS switches and their associated transceivers, system reliability is enhanced significantly by reducing the number of potential points of failure. The result is a simpler, cleaner architecture that's easier to scale, easier to cool, and easier to manage. The figure above compares our architecture with a reference architecture. It shows that eliminating switches (and their associated transceivers) can reduce the power consumption of the core network by 81% compared to a reference architecture, based on EPS switches.

This shift also carries some serious sustainability benefits. With less hardware in the loop, PRISM reduces the volume of electronics that end up as e-waste, an increasingly urgent problem as data center infrastructure grows. In addition, moving away from dense, custom silicon in the network core helps reduce dependency on critical minerals like rare earth metals and high-purity copper, which are becoming harder (and more geopolitically sensitive) to source. Additionally, reducing the cooling requirements means reducing the water supply needs for data centers.

How PRISM Fits Into Existing Ecosystems

PRISM is designed for more than just performance. It's built for practical integration, making it easy to slot into today's data centers without tearing everything down. Most existing setups use a spine-and-leaf topology with traditional EPS switching, and Oriole's network can be introduced as a high-performance cluster right alongside that. A lightweight Ethernet gateway bridges PRISM with existing infrastructure, so legacy compute and storage resources can keep doing their job while high-bandwidth, latency-sensitive workloads move to the optical side. This makes it easy to evolve the network over time: deploy it where it's needed most and expand the optical footprint as demand grows. It's a plug-in upgrade path, not a rip-and-replace job, which means faster adoption and smoother scaling without disrupting what's already working.



Takeaways

In a world where GPUs keep getting faster and models keep getting bigger, your network shouldn't be what's slowing you down. PRISM tackles the real bottlenecks head-on, with higher throughput, lower latency, and a drastic cut in power-hungry hardware. No sprawling switch fabrics. Less of a jungle of cables. Just a clean, photonic core that scales without turning your data center into a furnace.

By cutting complexity and power while boosting performance, PRISM can keep up with modern AI workloads and help scale further. It's a smarter way to wire up your DDL and HPC infrastructure with a network architecture that was designed specifically for large-scale deep-learning networks.

Oriole Networks' PRISM is the key to unleashing next-gen distributed AI training and inference with the world's first fast-switching and energy-efficient pure photonic network.

Oriole Networks Ltd

London
4 City Road,
London, EC1Y 2AA
United Kingdom

Tel: +44 (0)20 814 81981

info@oriolenetworks.com

Paignton
EPIC, White Rock Park, Waddeton Close,
Paignton, TQ4 7RZ
United Kingdom

Tel: +44 (0)18 037 14763

oriolenetworks.com

Palo Alto
380 Portage Ave
Palo Alto, CA 94306
United States