



ONE FIBER, TWO DIRECTIONS

The Case for Bidirectional Optics
in AI Data Center Scale-Up Networks

Authors: Taylor Groves, Lars Johnsson

September 14, 2026



Abstract

The scale-up domain is a tightly coupled cluster of AI accelerators that behaves as a single logical GPU. Its size directly determines training throughput, inference latency, and token generation rates. At 200 Gb/s per lane, copper's reach of roughly one meter constrains scale-up clusters to one or two racks — 72 to 144 GPUs. Expanding beyond that boundary has historically meant crossing into the scale-out network, where pluggable optics cap bandwidth and multi-hop switch architectures introduce latency that puts the burden to recover performance on software optimizations. To connect 512 or more accelerators across racks, photonic interconnects are the only viable path forward. However, the increased volume, complexity, and supply shortages of fiber cabling add significant time-to-market, cost, and reliability risks. A 512-GPU scale-up pod built with conventional unidirectional (UniDi) optics requires in excess of 130,000 optical fibers to connect all endpoints. Considering the shuffle infrastructure to route them this requires more than 260,000 fiber strands, totaling ~400 miles.

In this analysis, we outline how bidirectional (BiDi) optics, which carry transmit and receive on a single fiber using two wavelengths, mitigate this bottleneck. Based on a Lightmatter 512-GPU scale-up model, consistent with other independent analysis [1], we show that BiDi (1) cuts fiber and connector counts in half, (2) reduces total scale-up network spend by approximately 15%, and (3) removes roughly half the potential points of failure in the passive optical path, all while preserving switch radix, bandwidth and link reach. Lightmatter has incorporated BiDi across its interconnect portfolio. We illustrate the advantages of BiDi using the Passage L20 optical engine (OE) as an exemplary solution coming to market in 2027.



1. Introduction

The race to build larger AI models has become coupled with the race to build larger scale-up domains. At the chip and package level, GPUs are constrained by silicon area, yield, and power, and the industry's response has been to combine hundreds of accelerators into a single high-bandwidth, low-latency pod that presents itself to AI models as one enormous GPU. Mixture-of-Experts (MoE) models have pushed this domain to its limits, because the all-to-all expert-parallel traffic they generate can account for as much as 47% of forward-pass latency over a 7,200 Gb/s scale-up interconnect [2]. Larger scale-up domains translate directly into more deployable experts, faster training and better model quality.

The pod cannot scale with copper. At current 200G per lane SerDes speeds, the maximum reach of passive electrical interconnects today is approximately one meter, which effectively confines an electrically connected scale-up domain to one or two racks [2]. Fiber optics, which incur a negligible loss penalty to traverse the tens of meters that separate racks in a pod, are the inevitable replacement [2]. But simply replacing copper with optical links introduces new challenges. In a single-tier scale-up fabric, each GPU must connect to every switch. The size of the scale-up network is fundamentally limited by the size of the switch and the radix (number of links) that come out of it. In order to saturate the GPU bandwidth, data is striped across many rails of switches that sit at a single layer in the network. A UniDi optical fabric for a single 512-GPU pod creates a sea of fiber: hundreds of thousands of strands, tens of thousands of connectors, and the shuffle boxes that reorganize them, each fiber and each mated connector an opportunity for contamination, misrouting, or failure.

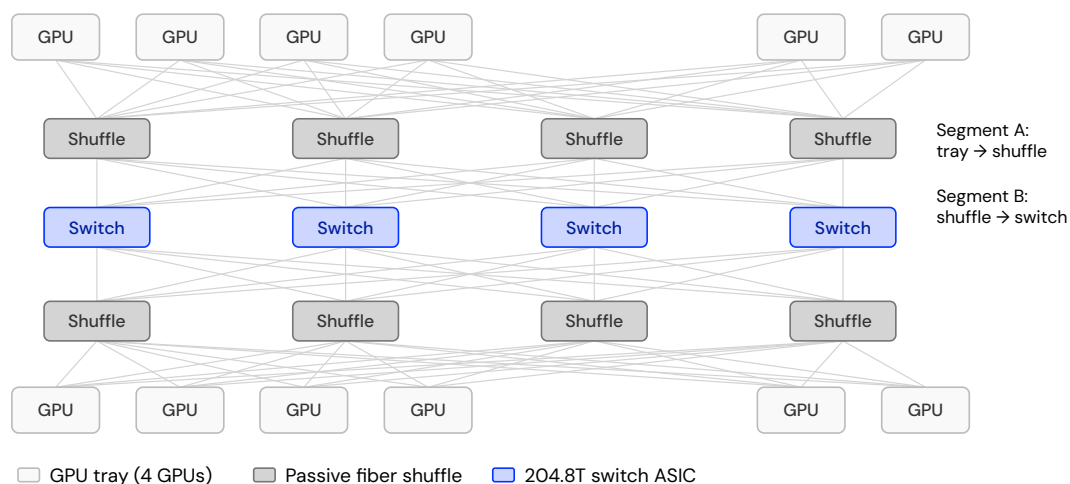


Figure 1: Topology of a 1-layer, 512-GPU scale-up cluster



This assessment poses a consequential question: given that scale-up must go optical, how do we mitigate the escalation of fiber-related costs and failures due to increased cabling complexity? The immediate and most effective answer is bidirectional optics. We first establish why scale-up is the critical domain and why it is going optical (Section 2). We then describe the BiDi link architecture and the physics of carrying two directions on one fiber (Section 3). We then quantify the fiber, connector, cost, and reliability benefits using the Lightmatter Passage L20 BiDi interconnect reference design and address the “cost of BiDi,” the optical insertion-loss adder, and where it lands in the link budget (Section 4). In conclusion we highlight the economic and operational savings of a BiDi scale-up fiber harness while addressing implementation trade-offs and open questions (Section 5).

2. Background: Why Scale-Up, and Why Optical

Scale-up versus scale-out. We need to distinguish between the types of networks because the requirements differ sharply. A traditional scale-out network uses optical transceivers to connect more than 100,000 GPUs at microsecond latencies, around 1.6 Tb/s¹ bandwidth per GPU, consuming roughly 16 pJ/bit in power. A scale-up network would leverage co-packaged optics to connect hundreds or thousands of GPUs at low latency, delivering bandwidth of 25.6 Tb/s or more per GPU, at under 5 pJ/bit [2]. Scale-up is, in essence, the domain where bandwidth density and energy efficiency matter most and where the latency budget is least forgiving.

The faceplate wall. Scale-up traffic in modern pods is routed through a single layer of high-radix switches so that any GPU can reach any other at full bandwidth in one hop. As scale-up bandwidth per GPU climbs, the number of fibers and lasers that must terminate on a switch tray grows faster than the faceplate can accommodate. As shown in figure 2, a 51.2Tb/s CPO scale-up switch requires 16 ELSFP laser modules and 512 UniDi fibers, which maxes out what a 1RU switch faceplate can accommodate. A next-generation 204.8Tb/s CPO scale-up switch utilizing today’s fiber and laser solutions would require 64 ELSFP laser modules and 2,048 UniDi fibers, which would consume 4RU for a single tray that is housing just one switch ASIC. The faceplate real estate, not the switch silicon, has become the binding scale-up constraint, and it is what forces a rethink of optical interconnects.

¹Bandwidth (bits per second) is quoted in each direction (rather than aggregate TX + RX) in this document.

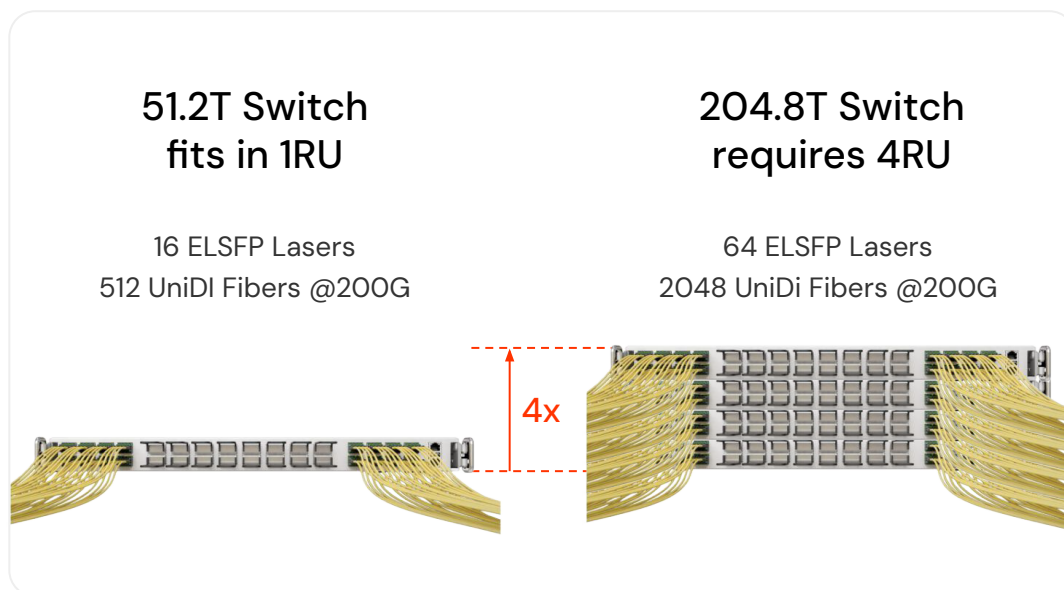


Figure 2: The laser and fiber scaling wall of CPO switches

The fiber plant behind the faceplate. Solving the faceplate problem highlighted in Figure 2 requires solving the two interconnect scaling challenges associated with today's ELSFP laser modules and UniDi fiber plant. Pluggable ELSFP modules trade off optical output power density for easy field serviceability. To efficiently support next generation switches, laser modules that provide much higher optical output power density are needed. Equally vital to addressing the faceplate real-estate crisis is maximizing fiber plant utilization to curb the growth of fibers and connectors. This analysis outlines a solution for the fiber plant challenges, while a future analysis will detail how novel form-factors and silicon photonics laser chips will dramatically increase the optical output power density of pluggable laser modules for scale-up networking.

Consider the UniDi fiber network of the 512-GPU scale-up cluster shown in Figure 1. At 25.6Tb/s bandwidth per GPU and 200G per fiber, each GPU-to-switch connection requires 128 transmit and 128 receive fibers, and in a typical 4-GPU compute tray layout 1,024 UniDi fibers would need to escape the GPU tray's faceplate through 64 SN-MT16 connectors, while a 204.8T switch faceplate would need to accommodate 2,048 fibers in 128 SN-MT16 connectors. To connect the 512 GPUs through the 64 scale-up network switches and the required shuffle boxes, the pod carries 131,072 data fibers. Each fiber traverses a path with multiple connectors from a GPU tray to a switch tray: a front-panel connector at the GPU, two shuffle connectors, and a front-panel connector at the switch. Assuming that the fabric uses MPO-16 or SN-MT16 connectors, the fabric requires



32,768 connector endpoints that must be cleaned and tested during deployment, at an estimated 1000+ hours of labor for a single pod. The fiber plant, in other words, comes with its own scaling wall, and it is this wall that bidirectional optics is designed to breach.

UniDi Fiber Routing in a 1-Layer Scale-up Network:

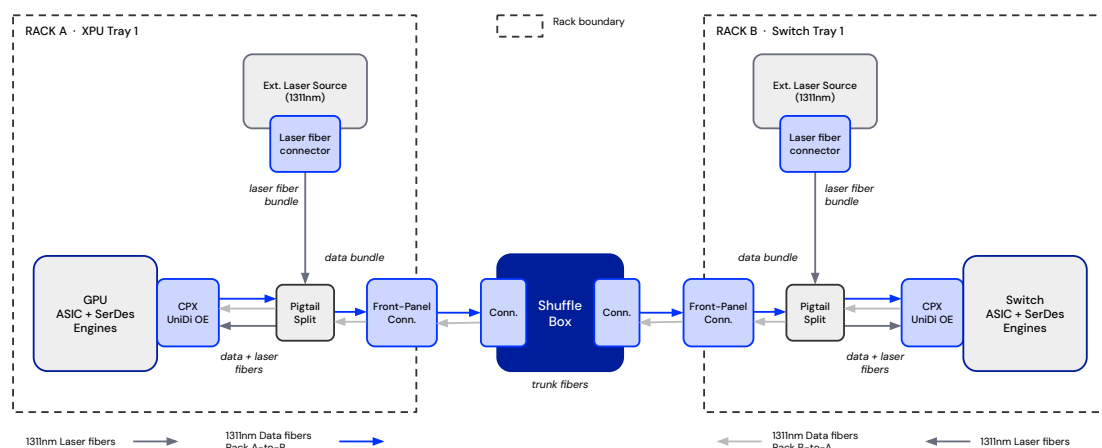


Figure 3: Example of a UniDi fiber path between GPU-and-switch in a scale-up cluster

3. The BiDi Link: Two Directions, One Fiber

The core idea. A conventional UniDi optical link, such as Direct Reach (DR), dedicates one fiber for transmit and a second fiber for receive, each carrying the same wavelength. A BiDi link instead carries transmit and receive signals on different wavelengths in a single fiber, combining them with a wavelength interleaver for transport. A conventional 6.4Tb/s OE that normally needs 64 fibers (32 Tx + 32 Rx) in a DR implementation needs only 32 fibers with BiDi, because each fiber carries two wavelengths. The Lightmatter Passage L20 is the first OE designed for scale-up networks that leverages the BiDi advantage, carrying both transmit and receive signals in each of its 32 fibers, delivering 6.4Tb/s bandwidth in each direction. The advantages are obvious, as shown in Figure 4. The entire fiber and connector fabric between the GPUs and switches is reduced in half without loss of bandwidth, performance or radix.



BiDi Fiber Routing in a 1-Layer Scale-up Network:

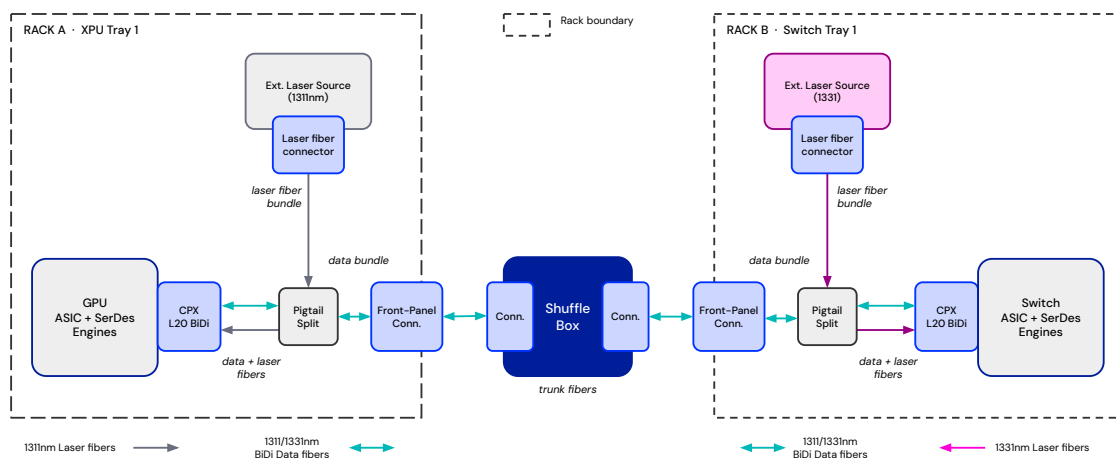


Figure 4: BiDi Fiber savings between GPU-and-Switch in a 1-layer, all-to-all scale-up cluster

Wavelength assignment. BiDi requires that the two link endpoints transmit on complementary wavelengths. The Passage L20 grid uses λA and λB at 1311 nm and 1331 nm at 20 nm spacing. The link is organized into two complementary groups: one set of ports transmits on λA and receives on λB , the other transmits on λB and receives on λA , and the only wiring rule is that a Group-A port must connect to a Group-B port on the far side. Figure 4 illustrates the fiber and connector savings and the wavelength assignment along with the two different External Laser Sources needed to power the BiDi network.

Flexible topologies. A reasonable concern is that splitting ports into two wavelength groups halves the usable radix or forces two distinct SKUs. It does not. Both endpoints carry identical hardware, optics and electronics alike, and the connectivity of the cluster is unchanged relative to DR. Each OE is programmable at the port-level to transmit on either λA or λB . This enables a variety of topologies such as trees, meshes or dragonfly, where an algorithmic mapping ensures a proper two-coloring (each λA port connects to a λB port) for the topology. BiDi is therefore a change to the fiber plant and the OE, not to the switch topology or the radix the architect gets to design with.



4. The BiDi Advantage, Quantified

Fibers and connectors, halved. The headline benefit follows directly from the architecture: carrying two directions on one fiber halves the fiber and connector count without sacrificing radix. For the 512-GPU pod, BiDi reduces the fiber count from 131,072 to approximately 65,536, a saving of more than 200 miles of fiber, and eliminates 16,384 of the fiber connectors along with the shuffle boxes that route them.

Total network spend, down ~15%. Fewer fibers, connectors, and shuffles result in a significant reduction in the capital expenditures for the scale-up network, as illustrated in Table 1 below. A 512-GPU scale-up cluster is modeled to compare key scale-up network metrics for a UniDi baseline with a Passage L20 BiDi solution. Moving from a UniDi based scale-up network to a BiDi based scale-up network cuts fiber mileage and the connector and shuffle count in half. For the 512-GPU scale-up cluster this translates into roughly 15% reduction in total network spend that is directly attributable to BiDi. Independent analysis by Broadcom [1] confirms that BiDi creates “optics cost savings” of 15% for a 512-GPU cluster, with savings scaling alongside fiber length and pod footprint.

512-GPU Scale-up Network CapEx	Conventional UniDi	Passage L20 BiDi
Fiber (miles)	~400	~200
Connectors	32,768	16,384
Racks	6	6
Network Spend	Reference	-15%

Table 1. Scale-up network spend comparison (includes switch, fibers, connectors and shuffle interconnect).

Availability, improved. Every fiber and every mated connector is a component that can fail, and the simplest path to improved reliability is to have fewer of them in the scale-up cluster. Halving the fiber and connector count removes roughly half the potential points of failure in the passive optical path: fewer connectors to contaminate, fewer shuffle boxes, and fewer fibers. Eliminating half the connectors and fibers increases pod-level availability and accelerates the cluster uptime that operators actually pay for.



Deployment, accelerated. The labor of deploying a pod scales directly with the fiber plant. With a cleaning, testing and installation time of 2 minutes per coupler the UniDi path requires an estimated 1000 hours of connector cleaning and testing for a 512-GPU pod. Halving the fiber and connector count removes a large fraction of that labor and the cabling complexity that goes with it, which we estimate can save over a week of integration and test time per pod.

Fiber supply constraints, mitigated. The BiDi benefit of halving the fiber count and mileage also mitigates the impact of any fiber supply constraints. In a regime where operators are signing multi-billion-dollar, multi-year fiber supply agreements to secure capacity, deploying the same compute bandwidth with half the fiber is also a hedge against an increasingly constrained supply chain.

Implementation, solved. BiDi requires OEs that have a wavelength interleaver in the optical path to combine and separate the Tx and Rx signals. The interleaver adds a nominal insertion loss to the optical path that is accounted for in the Passage L20 BiDi link budgets for CPO and NPO applications, which are supported by Lightmatter's Guide DR lasers and conventional ELSFP pluggables.

5. Conclusion

Scale-up cluster growth is one of the primary drivers in frontier AI, and copper's one-meter reach confines it to one or two racks, forcing a transition to optics. The question was never whether to go optical, but whether the fiber plant that comes with it would become a liability of its own: over one hundred thousand fibers, tens of thousands of connectors, added labor costs and heightened failure risk that come with it.

BiDi, benefits delivered. Bidirectional optics is the most direct answer available: by carrying transmit and receive on one fiber using two wavelengths, BiDi (1) halves the fiber and connector count, (2) reduces total scale-up network spend by approximately 15% with the savings growing (more as link lengths grow), and (3) removes roughly half the at-risk components in the passive optical path, improving pod-level reliability and accelerating deployment, all without sacrificing switch radix, bandwidth, or reach.



BiDi, ready to deploy. As scale-up domains keep growing, the complexity of the underlying fiber plant will only become harder to ignore. Lightmatter has already built BiDi into its interconnect portfolio, with the Passage L20 platform, powered by the Guide DR light engine, coming to market in 2027. Lightmatter's [Passage EVK50](#) and [Guide EVK](#) platforms are running in rack-level validation in Lightmatter's data center today, giving hyperscalers a multi-year head start to deploy the high-density optical scale-up networks that frontier AI demands.

Industry Validation. The Optical Compute Interconnect (OCI) Multi-Source Agreement (MSA) specification around a BiDi/DWDM architecture validates what Lightmatter has been building for years. Three generations of OCI-ready Passage optical interconnects already support the BiDi and CWDM/DWDM designs that the OCI-MSA has just put on paper.

References

- [1] Broadcom “CPO BiDi An Efficient Solution for Scale-up of AI/ML Clusters Using Optics”. CPO-BiDi-WP-100, May 17, 2024
- [2] M. Bernadskiy, P. Carson, T. Graham, T. Groves, H. J. Lee, and E. Yeh, “Accelerating Frontier MoE Training with 3D Integrated Optics,” 2025 IEEE Symposium on High-Performance Interconnects (HotI), 2025. arXiv:2510.15893.

© 2026 Lightmatter, Inc. All rights reserved.