SKYLANE OPTICS®

AI and ML Architectures: Fundamentals & Practical Considerations

White Paper

Introduction

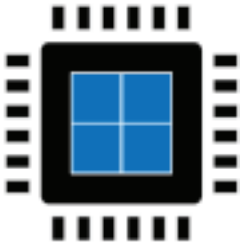
Artificial Intelligence (AI) and Machine Learning (ML) architectures represent a generational shift in how high-performance computing environments are deployed. The process of architecting AI and ML clusters requires specific network infrastructure centered on low-latency and high-bandwidth links to ensure not only optimal network performance but also streamlined deployment timelines.

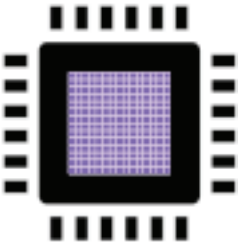
Fundamental to this architectural change is the profound influence of parallel processing capabilities, serving as the driving force behind the learning tasks that form the core of AI's value proposition. Notably, AI clusters are transitioning from traditional CPU-based hardware to GPU-based nodes. This shift not only signifies a change in computing paradigms but also presents a substantial challenge to established data center architectures, challenging current network design paradigms.

Background - Enter the GPU

In the dynamic landscape of AI clusters, data traffic exhibits distinctive characteristics, marked by the presence of substantial data flows often referred to as "elephant flows." These flows play a pivotal role in propelling an intensive workload across the network, setting AI and ML clusters apart from conventional CPU-based architectures. The key differentiator lies in the synchronization of traffic flow across a multitude of parallel jobs facilitated by GPU-based hardware.

Unlike traditional CPU architectures, AI and ML clusters leverage GPU hardware designed explicitly for handling parallel workflows. GPUs boast hundreds, or even thousands, of processing cores, a stark contrast to the four to eight cores typically found in CPU hardware. This distinction enables AI clusters to process and manage large volumes of data threads in parallel—a fundamental requirement for the complex computations involved in Artificial Intelligence and Machine Learning.





CPU – Central Processing Unit	GPU – Graphics Processing Unit
Typical 4-8 Cores	Hundreds or Thousands of Cores
Low Latency	High Throughput
Quickly processes tasks that require interactivity	Breaks jobs into separate tasks to process simultaneously
Good for Serial Processing – Smaller volumes of large transactions. Traditional software is written for sequential CPU execution.	Good for Parallel Processing – Large volumes of smaller transactions. Additional management software is required to bridge traditional CPU functionals to parallel execution.

At the heart of this data traffic is parallel processing, a mechanism wherein the GPUs' numerous cores collaboratively collect, process, and store data in tandem with the learning model deployed by the application. This synchronization across parallel workflows enhances the efficiency and speed of data processing—a critical factor in the performance of AI and ML applications.

Data Flow Characteristics

The data dynamics within AI clusters are characterized not only by their volume but also by their profound purpose. The data shuttled between different components of the cluster is integral to the machine learning process, encapsulating training data, model parameters, and intermediate results—all essential for the iterative learning and refinement of the AI model.

The AI model is a perpetual state of adaptation evolving and gaining insights driven by a distributed compute model, requiring an 'any to any' communication model. This dynamic nature poses challenges for network management, necessitating a high throughput, low latency infrastructure capable of seamlessly accommodating the ever-changing patterns of data traffic within AI clusters.

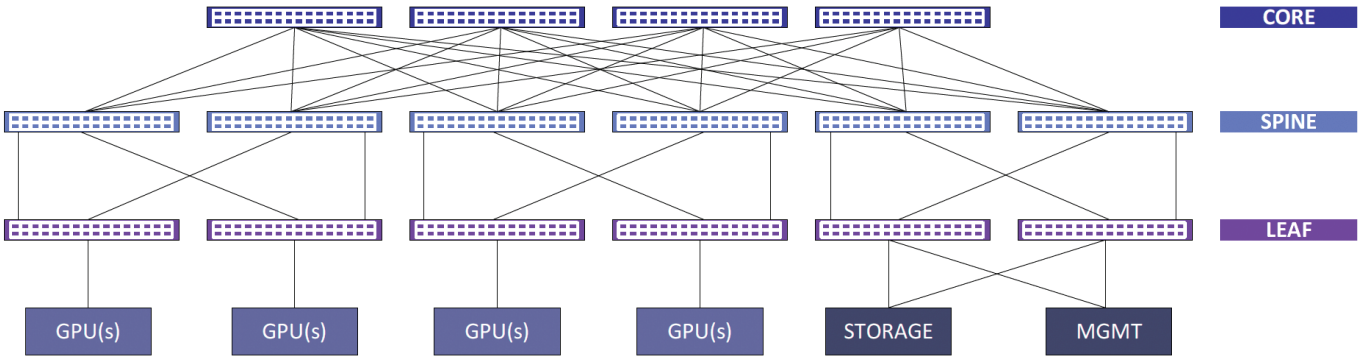
Challenges to the Traditional Data Center

AI and ML networks have undergone significant evolution to cater to the demands of parallel workflows. An AI cluster could be described as one large compute entity, with networking the GPU, storage, and other compute similar to the traffic through a traditional backplane or motherboard may flow. This GPU-based architecture presents data center architects with the challenges of building a contentionless network capable of addressing the density, latency, and bandwidth requirements of unpredictable AI and ML data traffic models.

Density

AL and ML cluster any-to-any models are characterized by GPUs directly connected to the Leaf and Spine architecture, bypassing traditional Top-of-Rack (ToR) switches. GPU nodes within AI and ML clusters will typically consist of two to four GPUs within a single network rack, with each GPU consuming up to 6KW and a dedicated network connection to the leaf. The implications of this power consumption are very high-power density racks, with a relatively low fiber optic cable count per rack. In contrast, a traditional Top-of-Rack server architecture may support ten, fifteen, or more servers in a single 5KW rack. While GPU nodes challenge power density, network switching racks are challenged by cable management density. Network switching racks are configured with high radix switches to accept fiber optic high-speed fiber cables from each downstream GPU and upstream Spine switches. A network rack consisting of eight 64-port Leaf switches must manage 256 downstream and 256 upstream fiber optic cables.

Despite the density challenges, the parallel processing capabilities of GPUs make them indispensable for the intricate computations involved in Artificial Intelligence and Machine Learning.

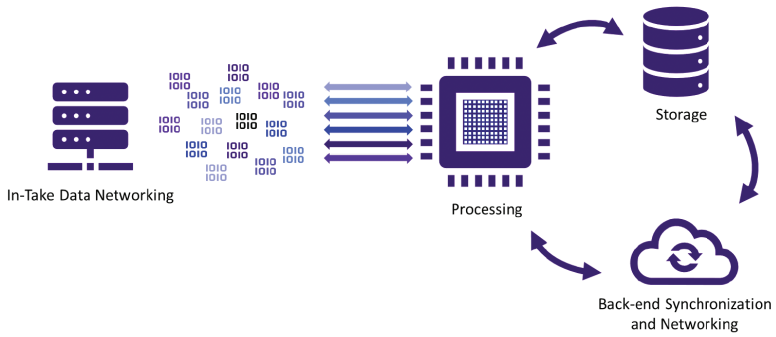


Bandwidth

AI/ML architectures are bandwidth-intensive, requiring a large data pipe to support the myriad of smaller data streams circulating within the cluster. To prevent network bottlenecks, AI clusters typically employ 100G, 200G, and 400G network adapters per GPU. Each network adapter connects to an aggregation port in a leaf switch, typically supporting configurations like 4x100G, 2x200G, or 2x400G. Large capacity network bandwidth is required to ensure the smooth flow of data across the extensive parallel workflows inherent in AI and ML applications.

Latency

AI and ML clusters employ topologies designed to reduce network latency, with architecture decisions bypassing network elements like TOR switches and minimizing fiber distances that may impair the any-to-any compute model. A 100% contentionless network is the objective to optimize latency and jitter to ensure optimal network performance. This imperative on reducing latency places significant importance on optical transceiver selection, supporting network port aggregation configurations such as 4x100G, 2x200G, 4x200G, and 2x400G. Understanding optical transceivers, including components like Digital Signal Processor (DSP) and gearbox technology, is required to mitigate latency in AI/ML clusters.



The other latency considerations are packet latency and jitter, focusing on ensuring in-order packet delivery to minimize delays. Buffering or re-transmission leading to out-of-order results can introduce latency, impacting the performance of AI applications. To mitigate this, Remote Direct Memory Access (RDMA) supported by InfiniBand technology is employed. RDMA enables direct data movement between the network adapter and the application, bypassing the kernel networking stack and avoiding reliance on the processor, cache, or operating system of either network element. This technology, a driving force behind InfiniBand, is also integrated into Ethernet through RDMA over Converged Ethernet and Ethernet iWARP. The integration of Ethernet technologies into AI and ML clusters is critically important for reducing the costs of AI and ML deployments.

Ethernet and InfiniBand

Characteristic	InfiniBand	Ethernet
Bandwidth	100G EDR / 100G HDR100 200G HDR / 400G NDR 800G XDR	100G / 200G / 400G / 800G
Load Balancing	Destination-Based Routing	BGP and ECMP
Packet Stability	Lossless Network	Lossy Network
Latency	~ 1μs	3us ~ 100μs

Impact on Infrastructure

AI and ML infrastructure topologies typically aggregate multiple GPUs and storage nodes into switch ports using parallel optics in a 'break out' configuration. Important variables must be considered in selecting optical transceivers for each topology: interoperability, technologies supported by the NIC, latency objectives, and fiber optic infrastructure.

Interoperability

Ensuring interoperability between parallel series transceivers in the leaf switch and GPU network adapters is crucial for successful AI/ML cluster deployments. When selecting optical transceivers, two key factors need to be considered for interoperability:

- NRZ and PAM4 optical signals are not interoperable.
- Modulation, wavelength, and optical lane data rates must all be aligned.

These interoperability considerations ushered in the next generation of PAM4 100G, 200G, and 400G transceivers. Legacy QSFP28 SR4 transceivers have served as the backbone of modern data centers but are not scalable beyond 200G. The four lanes of 25G, modulated with non-return-to-zero (NRZ) formatting, face interoperability challenges with the single lambda technologies integral to 400G transceivers. Likewise, multi-mode bi-directional technologies are also not interoperable with single-lambda PAM4 communications. Navigating these nuances is pivotal in constructing AI and ML infrastructures that not only optimize performance but also align seamlessly with evolving technological landscapes.

Parallel Series Module Interoperability Reference

Topology	Wavelength	Parallel Series Module Type	Optical Lanes/ Modulation	Constituent Transceiver	Optical Lanes/ Modulation
2x100G	850nm	QSFP28-DD 2xSR4	8x25G NRZ	QSFP28 SR4	4x25G NRZ
2x100G	850nm	QSFP56 SR4	4x50G PAM4	QSFP56 SR2	2x50G PAM4
4x100G	850nm/ 908nm	QSFP-DD SR4.2	8x50G PAM4	QSFP28 SR1.2	5x50G PAM4
4x100G	850nm	QSFP-DD SR8	8x50G PAM4	QSFP56 SR2	2x50G PAM4
4x100G	1310nm	QSFP-DD DR4	4x100G PAM4	QSFP28 DR1	1x100G PAM4
4x100G	1310nm	QSFP112 DR4	4x100G PAM4	QSFP28 DR1	1x100G PAM4
2x200G	850nm	QSFP-DD SR8	8x50G PAM4	QSFP56 SR4	4x50G PAM4
1x400G	850nm	QSFP-DD SR4	4x100G PAM4	OSFP-RHS SR4	4x100G PAM4
1x400G	850nm	QSFP-DD SR4	4x100G PAM4	QSFP112 SR4	4x100G PAM4
2x400G	850nm	OSFP 800G SR8	8x100G PAM4	OSFP-RHS SR4	4x100G PAM4
2x400G	850nm	OSFP 800G SR8	8x100G PAM4	QSFP112 SR4	4x100G PAM4
2x400G	1310nm	OSFP 800G DR8	8x100G PAM4	OSFP-RHS DR4	4x100G PAM4
2x400G	1310nm	OSFP 800G DR8	8x100G PAM4	QSFP112 DR4	4x100G PAM4

IT IS IMPORTANT TO NOTE THAT A TRANSCEIVER SUPPORTING INFINIBAND ALONE DOES NOT MAKE AN ETHERNET SWITCH COMPATIBLE WITH INFINIBAND.

Compatibility

Host compatibility, defined as the seamless operation of an optical transceiver within the intended platform, is fundamental for the success of AI/ML cluster deployments. However, reading the fine print of technologies supported by AI/ML network adapters and switches is crucial to avoid deployment delays. Paying attention to the functions supported by the network element is essential, as network adapters may or may not support functions such as QSFP56 SR2 (100G PAM4) and cannot be 'broken out' to aggregate small constituent optics. Additionally, network adapters may not fully support all of the functionalities offered by a transceiver, and a switch model may exclusively support either Ethernet or InfiniBand. It is important to note that a transceiver supporting InfiniBand alone does not make an Ethernet switch compatible with InfiniBand.

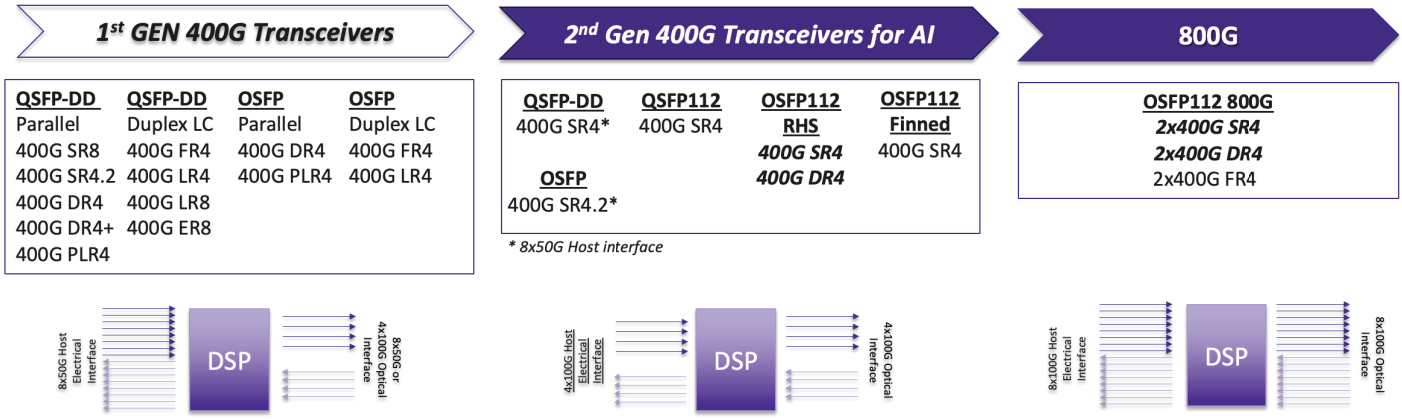
Network switches need the same attention to detail regarding function support but generally facilitate breakout configurations.

The "form" and "fit" of transceivers logically impacts their ability to perform as expected in a host device. The new generation of adapters and switches typically adheres to form factor guidelines from industry Multi-Source Agreements (MSA). QSFP MSAs offer the most generous backward compatible options for hosts:

- QSFP56-DD Ports – Supports 400G QSFP-DD, 200G QSFP28-DD, 200G QSFP56, 100G QSFP28.
- QSFP56 Ports – Supports 200G QSFP56 and 100G QSFP28.
- QSFP112 Ports – Supports 400G QSFP112, 200G QSFP56, and 100G QSFP28.

The OSFP form remains relevant in the industry today due to its larger size, offering superior power handling abilities through the built-in heat sink in the connector shell. However, integrating the OSFP form into a GPU node network adapter poses challenges. The 2019 revision to the OSFP MSA added the "RHS" option to the form factor. "RHS" – short for "Riding Heat Sink" – removes the heat sink from the transceiver body, relying upon a heat sink in the network element to 'ride' on the transceiver. The practical application was to ensure that the OSFP transceiver fits within the footprint of the server PCIE slot. Another distinguishing feature is the connector of the OSFP-RHS being unable to interface with a standard OSFP backplane. First-generation AI and ML network adapters have adapted this form factor for 400G communication.

The OSFP-RHS form factor introduced another twist to the market – the introduction of 112G at the host, even for 400G transceivers. OSFP-RHS transceivers follow the lead of the 800G OSFP transceivers, with only one MPO-12 connector and four of the eight lanes on the host loaded for communication.



Latency

The first generation of 400G transceivers operates on an 8x50G host interface architecture. In the case of four-lane transceivers, a Digital Signal Processor (DSP) and gearbox are employed to retime signals, condensing them from 8 lanes to 4 lanes. Notably, the 8-lane 400G SR8 DSP does not engage in data retiming. The PAM4 DSP without a gearbox in the 200G/400G transceiver ensures a swift 1:1 conversion with a latency of less than 90ns, while the 400G transceiver incorporating a PAM4 DSP with an 8:4 gearbox experiences a slightly extended latency of under 110ns, marking a 20ns difference.

In contrast, second-generation 400G transceivers capitalize on 800G technology, adopting a 4x100G host interface. This strategic shift results in a reduction in latency by up to 20ns per connection, showcasing advancements in the pursuit of more efficient and responsive data transmission. Parallel series modules not only aggregate seamlessly but also eliminate the need for a mux/demux in the module, removing components that may add latency.

Quick Reference Table

Data Rate	Transceiver	Modulation	Technology	Impact on Latency
100G	QSFP28 SR4	NRZ	CDR/No Gearbox	Low
100G	QSFP28 SR1.2	NRZ: Host, PAM4: Optical	DSP/2:1 Gearbox/Mux	High
100G	QSFP56 SR2	PAM4	DSP/1:1 Gearbox	Medium
100G	QSFP28 DR1	NRZ: Host, PAM4: Optical	DSP/4:1 Gearbox	High
200G	QSFP28-DD 2xSR4	NRZ	CDR/No Gearbox	Low
200G	QSFP56 SR4	PAM4	DSP/1:1 Gearbox	Medium
200G	QSFP56 FR4	PAM4	DSP/1:1 Gearbox/Mux	Medium
400G	QSFP-DD SR8	PAM4	DSP/1:1 Gearbox	Medium
400G	QSFP-DD SR4.2	PAM4	DSP/4:2 Gearbox/Mux	High
400G	QSFP-DD SR4	PAM4	DSP/8:4 Gearbox	High
400G	OSFP-RHS SR4	PAM4	DSP/1:1 Gearbox	Medium
400G	QSFP112 SR4	PAM4	DSP/1:1 Gearbox	Medium
400G	OSFP-RHS DR4	PAM4	DSP/1:1 Gearbox	Medium
400G	QSFP112 DR4	PAM4	DSP/1:1 Gearbox	Medium
400G	QSFP-DD DR4	PAM4	DSP/8:4 Gearbox	High
800G	OSFP 800G SR8	PAM4	DSP/1:1 Gearbox	Medium
800G	OSFP 800G DR8	PAM4	DSP/1:1 Gearbox	Medium

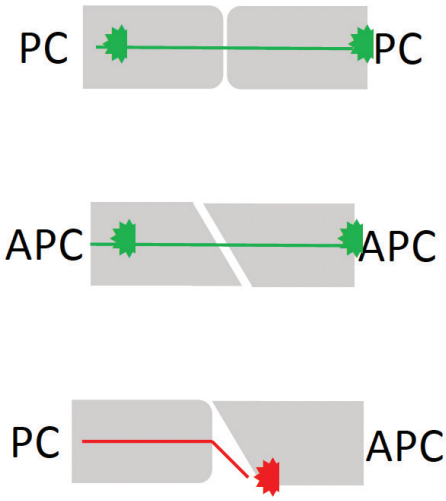
Fiber Infrastructure Requirements

A well-designed and implemented structured cabling infrastructure is essential not only for AI and ML performance but also to protect capital investment. Bypassing the traditional ToR architecture is a harbinger that existing data center fiber infrastructure may not be adequate for the requirements of AI and ML clusters. The planning of structured cabling begins at the fiber optic transceiver interface and accounts for the relative fiber densities of each segment within the network topology.

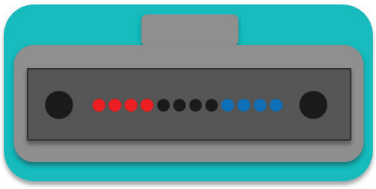
Connector Types

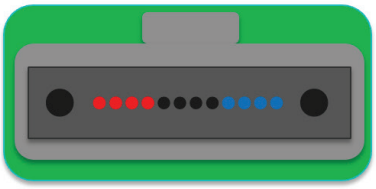
Parallel series fiber optic transceivers are characterized by aligning multiple fibers in parallel across a multi-fiber push on the (MPO) connector interface. The first-generation 400G SR8 transceiver introduced a new connector type – the MPO-16, prompting data centers to evaluate the cost/benefits of maintaining multi-mode short-reach architecture or transitioning to single-mode fiber. AI and ML clusters are shifting towards 100G vertical-cavity surface-emitting lasers (VCSEL) for multi-mode, short-reach connections, realizing the four lanes of 100G “SR4” architectures. In contrast, 400G SR4 transceivers require angle-polish (APC) MPO-12 connections, rather than the physical contact (PC) MPO-12 connectors common to 40G and 100G legacy connections.

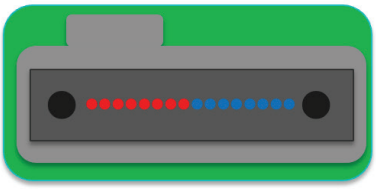
An APC fiber end face is polished to an 8-degree angle to perfectly with another APC end face. A PC-to-APC interface can result in intermittent service interruptions (flapping) and, at worst, cause damage to the transceiver.



Thus, the operational impact of MPO-12/APC in a retrofit or brownfield data center cannot be overstated. Industry standards have been developed to provide visual guides to correctly align fiber infrastructure installation. For example, the offset keyway of the MPO-16 connector ensures that an MPO-12 connector cannot be connected to a 400G SR8 transceiver. APC connectors are now specified by structured cabling standards with green sliding connector housings to avoid confusion with PC MPO-12 interfaces.







MPO-12 PC 4 Transmit, 4 Receive Fibers Multi-Mode Fiber (OM4, OM3) Aqua or Magenta Connector Housing	MPO-12 APC 4 Transmit, 4 Receive Fibers Multi-Mode Fiber (OM4, OM3) Single-Mode Fiber Green Connector Housing	MPO-16 APC 8 Transmit, 8 Receive Fibers Multi-Mode Fiber (OM4, OM3) Green Connector Housing Offset Keyway

Single-mode MPO-12 interfaces are, by default, angle polished and thus are not subject to the same operational complexities.

Connector Type by Transceiver

Transceiver	Fiber Type Supported/Reach (meters)	Connector Type
QSFP28 SR4	OM4/100m	MPO-12 PC
QSFP56 SR2	OM4/100m	MPO-12 APC or MPO-12 PC
QSFP28 SR1.2	OM4/100m, OM5 150m	Duplex LC PC
QSFP-DD SR4.2	OM4/100m, OM5 150m	MPO-12 PC
QSFP28 DR1	SMF/500m	Duplex LC UPC
QSFP28-DD 2xSR4	OM4/100m	MPO-24 PC
QSFP56 SR4	OM4/100m	MPO-12 APC or MPO-12 PC
OSFP-RHS SR4	OM4/100m	MPO-12 APC
QSFP112 SR4	OM4/100m	MPO-12 APC
OSFP-RHS DR4	SMF/500m	MPO-12 APC
QSFP112 DR4	SMF/500m	MPO-12 APC
QSFP-DD DR4	SMF/500m	MPO-12 APC
QSFP-DD SR8	OM4/100m	MPO-16 APC
QSFP-DD SR4	OM4/100m	MPO-12 APC
OSFP 800G SR8	OM4/100m	2x MPO-12 APC or 1x MPO-16 APC
OSFP 800G DR8	SMF/500m	2x MPO-12 APC or 1x MPO-16

Fiber Type

Once again, the decision point between single-mode and multi-mode fiber comes to the forefront in discussions about AI/ML clusters. Retrofitting data centers with existing multi-mode cabling may encounter a reach challenge unique to AI and ML clusters. Typically, short-reach applications over multi-mode fiber extend up to 100 meters over OM4 fiber. The projected massive demand for 100G VCSELs to meet AI/ML demands has introduced the VR4 (or VSR4) marketed under the “SR4” interface type. This 400G short reach variant with a reach of up to 50 meters on OM4, creates another decision point for data center retrofits, large data center rooms, or multiple data center rooms. Without a fiber infrastructure upgrade, these deployments may be limited by their fiber infrastructure.

Structured Cabling Considerations of AI/ML Architectures

The architecture of GPU-based AI/ML topologies fundamentally influences the design of structured cabling. In traditional data center setups, ten, twenty, or even forty CPU-based servers were often connected to a Top-of-Rack switch, with considerations ranging from physical size to available power, cooling, and data center layout. In contrast, GPU nodes from network equipment manufacturers may support up to eight 400G transceiver connections, using over 10KW of power. AI cluster reference architectures exhibit varying GPU node densities, with configurations of up to four nodes in a single network rack. Each node is connected within a Leaf and Spine architecture that incorporates management for storage fabrics. Data center constraints on power and cooling may reduce the number of GPU nodes per rack or increase the overall data center footprint.

Contrasting the challenges in data center infrastructure at the GPU node level, the substantial increase in fiber density within the Leaf and Spine segment of the network poses a challenge for standard data center racks and cable management systems. The reference architecture recommends 1,024 MPO cables from 128 GPU nodes converging onto 32 leaf switches, with an additional 1,120 cables building the fabric between the spine and core levels. While high fiber density is not uncommon in many data centers, the MPO cable density of this magnitude raises operational concerns.

Striking a balance between managing density, operational concerns, and cost is crucial for AI and ML cluster structured cabling infrastructures.

Backbone/Trunk Cabling

The planning and engineering of the structured cabling infrastructure supporting AI/ML clusters usually begins with the decision points on connecting nodes at the leaf level. This analysis considers two design options with divergent approaches – “backbone” and “home-run” cabling. Structured cabling designs with vertical backbone cables interfacing with horizontal cables at interim end-of-row locations offer test access and network reconfiguration points. While this approach provides superior operational and troubleshooting capabilities, it adds significant management and record-keeping complexity. The “home-run” approach utilizes pre-connectorized trunk cables connecting GPU racks directly to the leaf. This approach reduces the complexity of installation, maintenance, and record-keeping, but does limit troubleshooting and flexibility.

In-Rack Considerations

In the absence of Top-of-Rack switches, AI and ML cluster designs connect GPUs directly to the leaf. Best structured cabling practices involve terminating trunk cables to Top-of-Rack patch panels, providing flexibility and access for troubleshooting and polarity management in the data center. The low-cable density nature of GPU racks also opens options to use modular patch panels with options for additional media types, such as LC adapters and RJ-45, are advisable. This approach ensures equal flexibility for the physical layer of the management network.

In contrast to the relatively low density of GPUs, the significant increase in fiber density within the Leaf and Spine segment of the network presents a challenge for standard data center racks and cable management systems. The MPO connectivity originating from the GPU nodes converges to high-radix AI cluster Leaf and Spine switches. In a scenario with 140 GPU nodes, the outcome results in 1,120 MPO cables to the leaf and an additional 1,120 to the spine and core, respectively. While high fiber density is not uncommon in many data centers, the MPO cable density of this magnitude raises operational concerns about airflow, management, record-keeping, and hand access.

MPO Cabling Considerations

The integration of MPO cabling adds complexity to the structured cabling environment, requiring careful considerations of polarity and fiber density to avoid deployment delays and unexpected costs.

Structured cabling designs by credentialed Registered Communications Distribution Designers (RCDD) offer the safest approach to avoid unexpected MPO cabling issues. Even professional designers are limited by the information at the time of installation and may already be off the project by the time late changes are made to the network. Notably, MPO connector polarity often emerges as a last-minute concern. One change in the end-to-end cabling design can throw off polarity and delay deployment until new cabling is procured. Utilizing field configurable MPO patch cables offer a hedge against polarity induced delays. These cables can re-configure not only the polarity but connector pinning. The cost delta between field configurable MPO cables and fixed MPO cables can be mitigated by designating field configurable cables for a Top-of-Rack connection at one end or the other of the span.

MPO Connectors from GPU Nodes to Leaf	1,024
MPO Cables from Leaf to Spine	1,120
MPO Cables From Spine to Super Spine	1,120
Total MPO Cables in Network Core	3,244

Moreover, cable density within the leaf/spine portion of AI/ML clusters is exacerbated by the larger physical size of both the connector and the cable itself. Unlike LC patch cables that have decreased in size with technological advances, MPO cables typically use a 3.0mm outside diameter cable to protect the multiple fibers inside from bends. Increased fiber density will block visual access to connectors and restrict safe hand access in highly populated panels and switches. These restrictions can be mitigated by MPO patch cables with pull tabs. The pull tabs on the connector enable safe hand access to individual MPO connectors, minimizing the potential for unintended consequences when touching adjacent connections. Similarly, MPO patch panels used in the Leaf and Spine environment must offer a balance between high density and reduced complexity. Considerations for patch panels and high-radix switches in the Leaf and Spines are based on 64, 400G port architectures. Each switch will likely have 32 ports to the downstream level and 32 upstream. High-density patch panels configurable to match port count, upstream, and downstream configurations will assist in reducing complexity in this portion of the network.

Conclusion

The design and implementation of robust AI and clusters require a holistic approach to address key network infrastructure challenges of data center density, network latency, and bandwidth.

AI and ML clusters often involve interconnected network elements processing vast amounts of data simultaneously in an any-to-any manner. The diversity of power-hungry GPU nodes, storage, and core switching present practical challenges requiring a holistic view of transceiver selection and the structured cabling environment. Medium to large clusters will tilt the deployment model toward transceivers and structured cabling as GPU nodes and Leaf/Spine switching environments are separated by the power density of GPUs. The sheer volume of MPO cables converging at the data center Leaf presents challenges that require cable management and polarity solutions to avoid project delays.

Latency is another critical factor that demands a holistic perspective in AI and ML cluster design. Reduced latency is essential for real-time AI applications. A comprehensive strategy involves not only selecting low-latency networking technologies but also minimizing communication contentions through an understanding of how transceiver technologies and cabling can seamlessly meet the demands of applications where split-second decisions are crucial. New 112G SerDes at the host high-speed transceivers with 1:1 DSP gearboxes provide lower latency than 56G 4:8 DSP gearboxes. Bypassing the ToR switch reduces a point of contention and latency in the network but presents structured cabling challenges to efficiently cable nodes to the network Leaf.

Bandwidth management is the third pillar in the holistic approach to AI and ML cluster design. In essence, a holistic approach to bandwidth management involving 400G and 800G data rates is essential for future-proofing AI clusters against the escalating requirements of data-intensive AI applications. As AI models and datasets continue to grow in complexity and size, the demand for high bandwidth becomes paramount. A comprehensive strategy involves not only investing in high-capacity network infrastructure but also understanding the impact of transceiver selection on the structured cabling environment. 100G VCSELs have introduced APC multi-mode fiber connectors to the data center market, presenting architects with yet another data point in the tipping point between multi-mode and single-mode cabling.

By addressing density, latency, and bandwidth collectively, designers can create AI clusters that are not only powerful but also scalable, resilient, and capable of handling the evolving landscape of AI technologies.

About Skylane Optics

Skylane, an Amphenol company, is a leading provider of transceivers for optical communication.

We offer an extensive portfolio for the enterprise, access, datacenter and metropolitan fiber optical market as well as for smart home applications and home networks.

We cover the European, South American and North American market with a strong partner network and have offices in Belgium, Brazil, Sweden and USA.

Our offerings are characterized by high quality and performance. In combination with our strong technical support, we enable our customers to build cost optimized network solutions.

We offer an extensive range of high-quality products including transceivers (Optical and copper), Active Optical Cable (AOC), Direct Attach Cable (DAC), Mux/Demux, Coding Box (TCS).

Skylane Optics ®
Rue du Moulin, 18
5650 Fraire - Belgium

T: +32 71 610 640
info@skylaneoptics.com

www.skylaneoptics.com