WHITE PAPER

AI Infusion

The Next Wave of Data Center Networking

Mike Connaughton, RCDD, CDCD
Senior Manager, Product Management,
Leviton Network Solutions

Brooke Ford
Director, Product Management,
Leviton Network Solutions

TABLE OF CONTENTS

INTRODUCTION

The latest advancements in Artificial Intelligence (AI) are introducing revolutionary new ways to automate our workflows, solve problems, and develop insights on large datasets. However, intense processing, higher data rate transceivers, and large data storage requirements for AI computing creates new power and latency implications in data centers that can affect network infrastructure design.

Applications using AI are expanding every day, driving up demand for compute power in financial, industrial, government, manufacturing, and so many other sectors. Generative AI platforms like OpenAI's ChatGPT and Anthropic's Claude offer virtual assistance, including research, task automation, coding, and more. Microsoft Copilot, introduced in 2023, is an AI companion that helps enterprises across many sectors to save time by summarizing meetings, identifying trends in spreadsheets, and more. In the biomedical sector, AlphaFold is an AI program recently developed by the DeepMind research lab to make highly accurate predictions about how amino acid sequences form three-dimensionally — advancing drug and vaccine development and accelerating biological research.

As we continually discover new ways to harness emerging AI technology, AI computing clusters are at the front end of a huge ramp in growth. Industry analyst LightCounting expects a nearly 30% Compound Annual Growth Rate (CAGR) in fiber transceiver sales for AI clusters through 2028. Transceivers for non-AI applications in the data center will see a 9% CAGR — also strong growth — but it pales in comparison to what the AI expansion will bring.

AI and Machine Learning Growth in the Data Center

	% Of Transceiver Sales			CAGR
	2020		2028	2020–2028
AI Clusters	**15%**	▲	**41%**	**29%**
Rest Of Cloud Network	85%	▼	59%	9%

Source: LightCounting

Given AI computer clusters function on large amounts of processing, data storage, and power, the networking architecture and requirements applied to large data centers can also apply to networks running AI algorithms. This paper provides an overview of the challenges data center managers may face when deploying AI clusters, including power, cooling, geography, latency and deployment speeds. These factors will ultimately drive the network architecture, cabling, and connectivity required in these data centers.

POWER CONSUMPTION AND COOLING

Artificial Intelligence relies on high-performance computing (HPC), and this type of computing requires access to substantial power. For this reason, AI clusters can't easily be deployed in existing data centers due to available power limitations.

Take, for example, the NVIDIA DGX B200, the current gold standard graphics processing unit (GPU) for AI applications. The single unit consumes up to 14.3 kilowatts, an amount of power consumption that not long ago would be considered enough power for an entire rack. In an AI cluster installation, up to four of these units, would be mounted in a single rack, as shown in **Figure 1**. Simple math equates to a power consumption of more than 57 kilowatts per rack. Roll this out into several rows of a data center, and several megawatts of power quickly become necessary. For many data centers, and many regions, there is not enough available power to simply drop an AI cluster into an existing building. As data rates and capabilities increase, it should be expected that the power demand will grow.

Due to these increased power demands, new power infrastructures and dedicated facilities are already coming into force. In markets around the world, AI is pushing data centers toward greater power capacity. According to market research from data center real estate analysts datacenterHawk, 36 megawatt leases and even 72 megawatt leases are becoming increasingly commonplace, driven by power requirements from AI's compute-dense deployments.

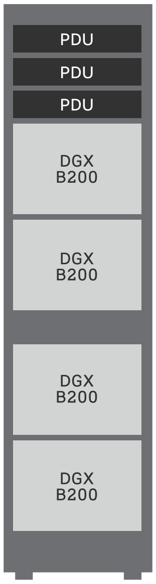


Figure 1:
NVIDIA DGX B200
Server Rack
4 Servers = 57+ kW

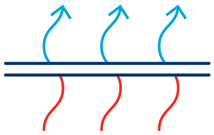
Transceiver Selection

One way to minimize AI's power usage is through transceiver selection. Most high-speed transceivers incorporate digital signal processing (DSP) to maintain signal integrity. However, for many links in an AI cluster, the distance between ports is very short — in the tens of meters. Multimode transceivers use vertical-cavity surface-emitting laser (VCSEL) technologies, which use less power than EML lasers used for single-mode transmission and can be a consideration for shorter links. For example comparing a 800G-SR8 multimode transceiver which consumes 4.5 W of power to a 800G-DR8 singlemode transceiver which consumes 16 W of power there is the potential for over 10W of power saving.

Additionally, Linear-drive Pluggable Optics (LPO) is an emerging transceiver technology that replaces the DSP with simpler components. This option reduces power consumption (and latency) and is a great candidate for short reach connections, whether they are multimode or single-mode fiber. Using LPO technologies offers the potential for 20-25% power savings. When used across thousands of ports in an AI cluster, that adds up to significant power savings, making them increasingly popular in AI deployments.

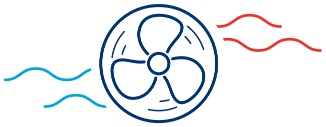
Cooling

When individual devices are using 14 kilowatts, it is crucial that effective and efficient cooling is installed to prevent these devices from overheating. This is among the most pressing topics for data center architects when faced with AI deployments. Cooling is being implemented today in many scenarios, such as air-to-air cooling, liquid assisted cooling, direct to chip cooling, and fluid immersion.



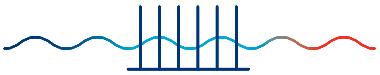
AIR-TO-AIR

Most common today, heavy reliance on HVAC systems.



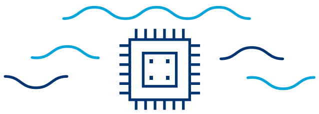
REAR DOOR HEAT EXCHANGE (RDX)

Bring water to the cabinet using a rear door heat exchanger.



DIRECT LIQUID COOLING (DLC)

Fluid is piped directly to a heat sink on or near the chip to draw heat away.



FLUID IMMERSION

The entire board operates from inside a fluid bath.

These examples provide escalating cooling efficiency improvements. RDX cooling and DLC cooling are already growing in popularity. While space needs to be allocated for the equipment associated with these options, there is minimal impact on the cabling. However, with fluid immersion, when the cable is exposed to the fluid, there are many parameters that can be impacted, such as cable flexibility, termination retention, and jacket integrity. With a wide variety of immersion fluids available, and the wide variety of cabling materials, there are a considerable number of combinations that need to be understood.

LATENCY AND DATA RATE IMPLICATIONS

AI calculations are done in parallel and are limited by the slowest calculation. That means addressing latency is vital to creating a fast, operational AI cluster. Ethernet and InfiniBand are the dominant options for connecting AI clusters, but each provide different advantages.

Ethernet is the dominant networking protocol, providing advantages as a diverse, multi-vendor ecosystem offering interoperable Ethernet switches, network interface controllers (NICs), cables, transceivers, optics, management tools, and software from various parties. Additionally, the proven scalability and cost-effectiveness of Ethernet networks provide distinct advantages. However, Ethernet needs to be improved to achieve low tail latency AI needs.

Introduced in 1999, InfiniBand is an industry standard architecture designed to interconnect servers, storage, and communications infrastructure equipment, providing high bandwidth and low latency. Used in both high performance computing and AI clusters, InfiniBand is used in nearly 200 of the top 500 supercomputers, and in more than half of the top 100 supercomputers, according to the InfiniBand Trade Association. While InfiniBand connections still need to adapt to allow simpler connectivity in larger systems, the technology is evolving to meet Ethernet standards for seamless integration.

The protocol RoCE (RDMA over Converged Ethernet) is a solution for sending InfiniBand packets over an Ethernet channel. When optimized, RoCE can offer the same latency as Infiniband. In addition, a group called the Ultra Ethernet Consortium is developing a transport protocol with multipath, packet-spraying delivery and other improvements.

Data Rates

As you can imagine, AI is driving higher and higher data rates in data centers. Today, 200 Gb/s is most likely the slowest data rate in a large-scale AI cluster. In fact, 400 Gb/s and 800 Gb/s are more typical, and 1.6 Tb/s is expected to gain adoption in the near future.

Adoption of 400G took off in 2021, as shown in **Figure 2**, and is forecasted to maintain steady growth, with parallel optics being the most popular (DR4, SR4). 800G grew rapidly in 2023 with multiple hyperscale deployments and was aided by the availability of the NVIDIA DGX H100 servers that use 800G for compute fabric.

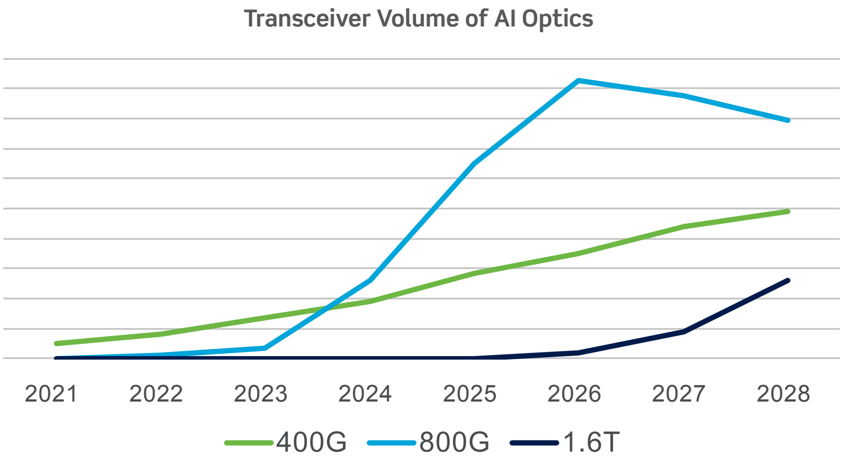


Figure 2: 800G will be the most popular data rate over the coming years with 1.6T starting to emerge in 2026 and 2027. Source: LightCounting

Data center managers are attaining these higher data rates through several approaches, often used in combination with each other:

1 **LANE SPEED IMPROVEMENTS**
Encoding schemes have been developed that squeeze more bits of data into each on-and-off cycle of the light emissions in the fiber. Today, the lane speed of 100 Gb/s is standardized by IEEE, and they are already working on a 200 Gb/s lane rate.

2 **WAVELENGTH DIVISION MULTIPLEXING**
WDM sends several wavelengths simultaneously down an individual fiber, multiplying the available bandwidth by the number of wavelengths. This adds cost to the transceiver, as it incorporates multiple lasers into each transceiver. The plus side: it reduces the number of fibers required.

3 **USE MULTIPLE FIBERS** (parallel transmission)
Splits the data rate into separate lanes. For example, 400 Gb/s can be split into four streams of 100 Gb/s and then recombined at the other end at the 400 Gb/s interface. From a technical standpoint, this is the simplest approach to increase data rates and usually the least expensive option.

NETWORK ARCHITECTURE, CONNECTIVITY AND CABLING

Naturally, the physical interfaces of these data rates lead to different physical connectors and termination requirements. Connection choices at the higher data rates are largely driven by multi-source agreements (MSAs) between vendors, the most recent of which are the OSFP and QSFP-DD800 MSAs.

Connection at the 400G data rate is dominated by traditional 12-fiber MPO connectors and LC connectors, with volume on the LC duplex FR4, and MPO-based DR4 and SR4 parallel optics. One difference is that, at this rate, multimode interfaces will drive the adoption of angled physical contact (APC) to reduce reflectance or return loss. The angled endface geometry improves return loss and performance, supporting higher data rates like 800G and beyond.

These same connectors for 400G are still the most popular choice at the 800G rate, but they are deployed "belly-to-belly" within the transceiver's form factor to increase the fiber count, as shown in **Figure 3**. The MPO-12 for example, still operates at 100G per lane, but instead of using four pairs of fibers for 400G, MPO-12 now uses eight pairs to double that bandwidth and achieve 800G. Similarly, at the 800G rate, MPO-16 will emerge with additional volume. With an offset key, shifted to the side, MPO-16 has its own unique connector interface while still using eight pairs of fibers for communication.

Figure 3: Transceiver interface with "belly-to-belly" 8-fiber MPO connections
Source: NVIDIA

Other interface options have been introduced for these higher data rates. These are commonly referred to as Very Small Form Factor, or VSFF, connectors. The CS®, SN®, and MDC are VSFF connector options that enable higher density duplex connections, but these will see lower levels of adoption initially. As of this writing, the MMC connector, a VSFF connector for parallel optics, is beginning to gain traction in the hyperscale space.

Network Architecture

It is important to note that there are a wide variety of options when it comes to AI network hardware and configurations — it would be impossible to cover all the possibilities. For the purposes of this discussion, this paper uses NVIDIA's reference architecture for its popular DGX B200 system. In this example, there are four different functions, each requiring their own network connectivity:

COMPUTE FABRIC

Consists of the highest bandwidth connections to enable communication between GPUs across nodes, acting as a large supercomputer for intense AI training and learning. The compute fabric operates over the lower latency InfiniBand protocol.

STORAGE FABRIC

Provides ready access to shared data across the nodes in support of those training and learning functions. This fabric would also operate over InfiniBand.

IN-BAND MANAGEMENT NETWORK

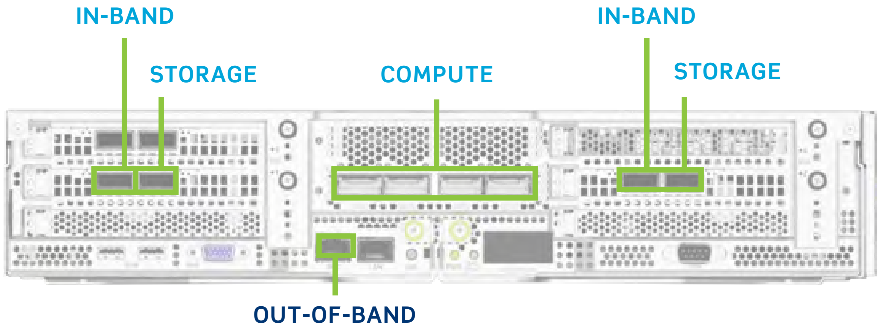
Typically consists of high-speed 100G and 200G links for connecting all services that manage the AI cluster and software communications. It operates over Ethernet.

OUT-OF-BAND MANAGEMENT NETWORK

Uses low-speed copper connections for other basic management functions, connecting to servers, switches, PDUs, and others, and uses the Ethernet protocol.

All four of these network fabrics have interfaces on the rear of the system, as shown in **Figure 4.** Most notable are the compute ports, each operating at 800G, while utilizing twin MPO port connections each operating at 400G. This is an example of the "belly-to-belly" interfaces mentioned earlier. Each server has four twin ports for a total of 8 MPOs and up to 64 fibers for the compute function. Depending on the device configuration, there are 1 or 2 ports used for the storage and In-Band network for 10 to 12 MPO interfaces and as many as 96 fibers per system. There is also a copper interface for the out-of-band management, which must be considered when planning the cabling infrastructure.

DGX B200 Network Ports



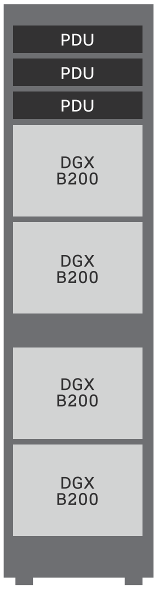
Per Server Requirements

Function	MPO 8 Ports	Fibers
Compute	8	64
Storage	1 or 2	up to 16
In-Band	1 or 2	up to 16
Total Fiber	**10 to 12**	**up to 96**

Function	Copper Ports
Out-of-Band Cat 6	**1**

Figure 4: NVIDIA DGX B200 Network Sled Rear View
Source: NVIDIA

Up to four DGX B200 servers can be placed in a rack. Multiplying the per server requirements by four, the total cabling requirements per rack yields up to 384 fibers, as shown in **Figure 5**. This is a dramatic increase over typical front-end network server architectures that might run duplex 50G or 100G duplex connections to the servers, with four uplink MPO ports to the next layer of switching.



Per Rack Requirements

Function	MPO 8 Ports	Fibers
Compute	32	256
Storage	4 or 8	up to 64
In-Band	4 or 8	up to 64
Total Fiber	**40 to 48**	**up to 384**

Function	Copper Ports
Out-of-Band Cat 6	**4**

Figure 5: Populated AI Server Rack
and Cabling Requirements

Taking this design a step further, NVIDIA clusters up to 8 DGX B200 racks into a row called a “Scalable Unit”, architected to expand to up to four of these rows for a massive supercomputer pod. This multiplies the per rack requirements by 8 for as many as 384 MPO ports, correlating to over 3,000 fibers.

All of these connections are aggregated to the switching racks to create the Compute, Storage, In-Band, and Out-of-Band fabrics previously mentioned. These switch racks could be located in a middle-of-row architecture as shown in **Figure 6**, at the end of the row, or in a more centralized location. The overall rack spacing may vary depending on what is optimized for the particular data center. Regardless, with this number of ports and fibers in such a small number of racks, it becomes clear why the percentage of optics for AI applications will be increasing so rapidly over the next five years.

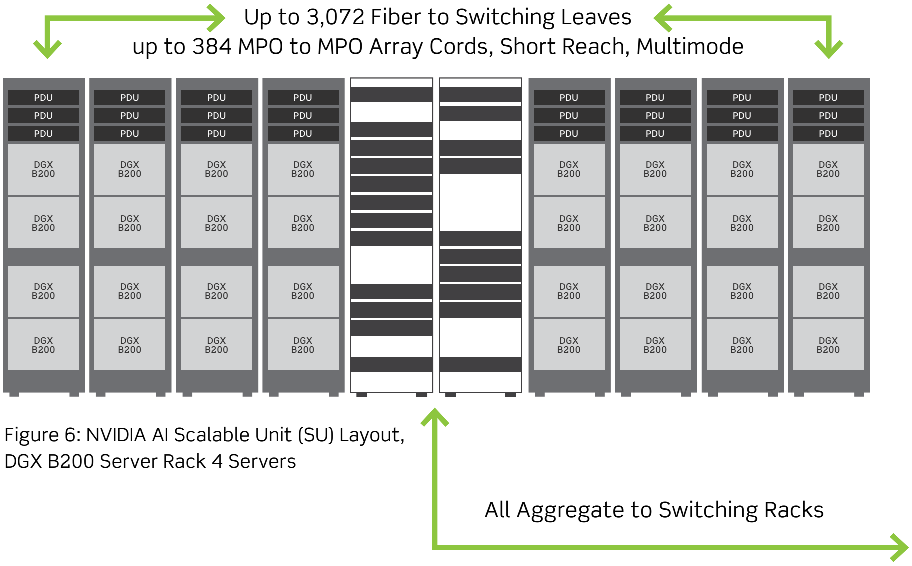


Figure 6: NVIDIA AI Scalable Unit (SU) Layout, DGX B200 Server Rack 4 Servers

Per Rack Requirements

Function	MPO 8 Ports	Fibers
Compute	32	256
Storage	4 or 8	up to 64
In-Band	4 or 8	up to 64
Total Fiber	**40 to 48**	**up to 384**

Function	Copper Ports
Out-of-Band Cat 6	4

Per SU Requirements
(Rack Requirements x8)

Total Fiber MPO 8 Ports	**up to 384**
Total Fibers	**up to 3,072**
Total Copper Ports	**32**

Direct Cabling Connections

What are the cabling options to make the high volume of connections in a SU layout? One option is to install direct connections from the AI systems to the switching fabric, using either active optical cables (AOCs) or discrete MPO array cords, as shown in **Figure 7**. This is the most straightforward approach, but it does create a significant volume of cable in the cable tray and within the racks themselves. Additionally, the QSFP (or OSFP) connector attached to the end of the AOC needs to be routed through the cable pathway. Pulling the transceivers through the pathways can be cumbersome, and the installer must be careful to avoid damage. Also, as more connections are added, routing can become more difficult.

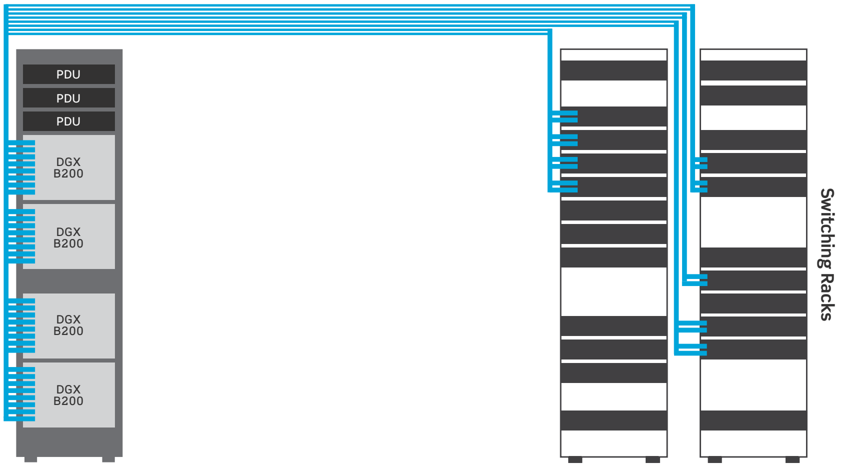


Figure 7: Port-to-port direct connection cabling

If using a direct connect design, it is highly recommended to take steps to ensure the source and destination devices are known to help the installer. Very specific cable labeling is critical, and cable bundling or grouping is recommended.

Structured Cabling

As an alternative to direct connections, structured cabling offers significant benefits in AI clusters. This design replaces the vast number of point-to-point connections in the overhead tray with patch panels on either end, and higher-fiber-count MPO trunk cabling between the racks. .

Structured cabling reduces congestion in the overhead trays and cuts down on cabling runs — creating up to 85% fewer cables to manage. Also, the trunk infrastructure can be pre-installed before active equipment is in place, so only the patch cords need to be installed once active equipment is in place.

Structured cabling also allows for smaller in-rack cables on the front side of patch panels, as shown in **Figure 8**. This helps reduce congestion and improve cable density within the rack itself. AOCs and traditional MPO array cords typically have a 3.0 mm or 3.6 mm diameter, whereas shorter array cords or breakout legs of a harness or trunk can use a 2.0 mm diameter cable, offering a 33% size reduction. These smaller diameter options in a structured cabling design can help improve airflow, contributing to greater overall efficiency in the rack.

Structured cabling also adds flexibility to help installers with cable management, offering panel labeling, easy cable grouping, and color coding of ports and connectors to make it easier to identify cables and reduce troubleshooting.

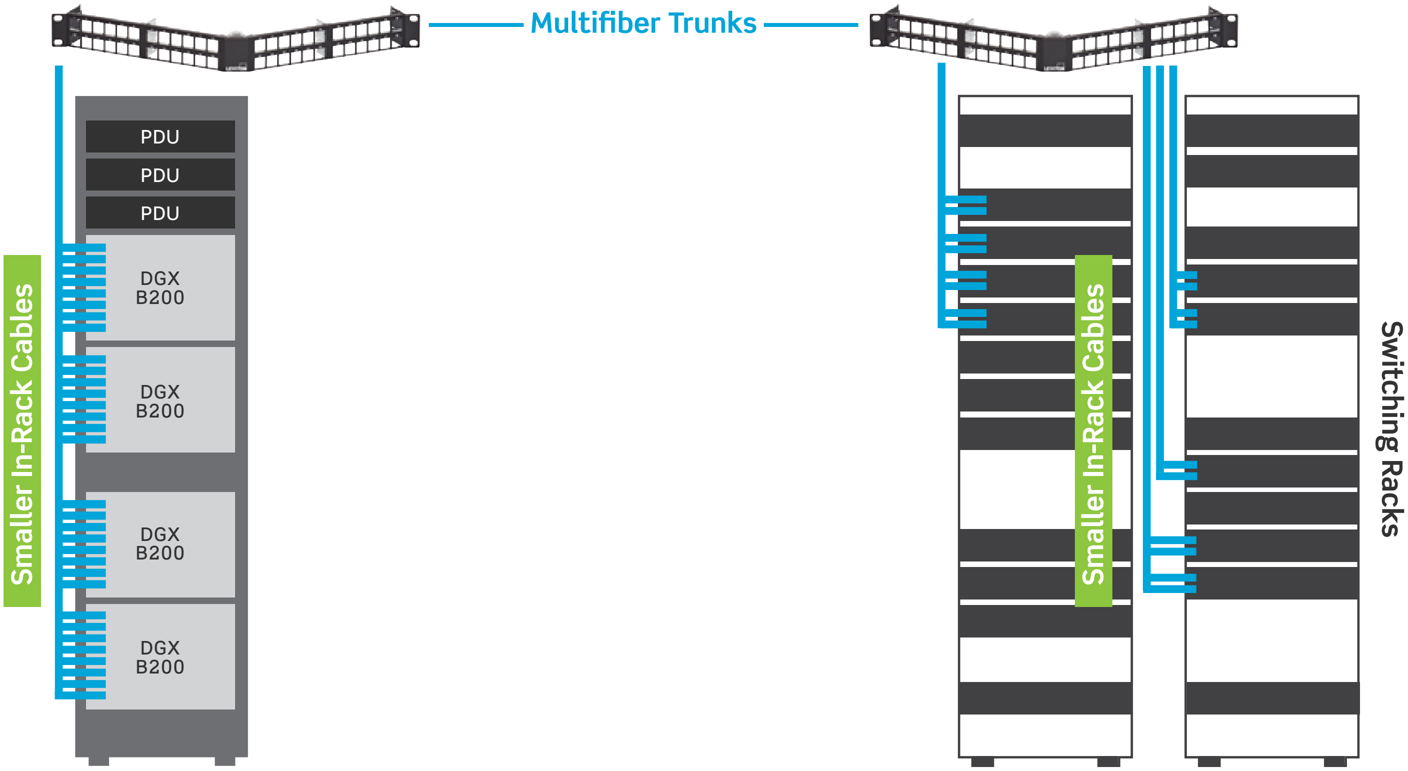


Figure 8: Structured Cabling Design

In addition, with a structured cabling design, pre-terminated trunk cables can be pre-installed so that only patch cords need to be run to the active equipment on the day of turn-up. This can significantly reduce the turn-up time and the time to revenue (TTR).

To quantify this time savings, Leviton partnered with installers to document the step-by-step process used when making direct connections with individual array cords versus using a structured cabling solution where the multifiber trunks were pre-installed.

On the direct attach side, it was estimated that it would take about 60 seconds per port to make the physical connection, including including sorting, opening and cleaning the cords along with other standard practices, and about 5 seconds per foot to pull the cable. But as the pathway becomes more and more congested, it was estimated that it takes approximately 0.1% more time with each subsequent connection, as each cord or groups of cords need to be placed in an increasingly crowded pathway.

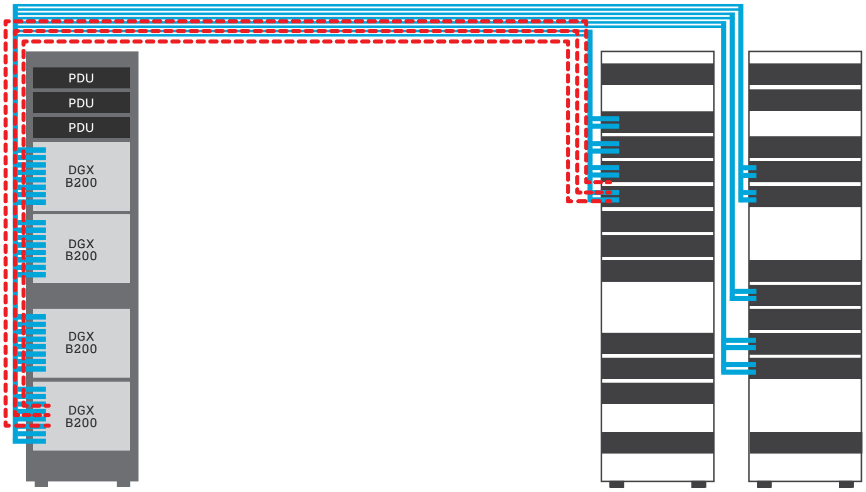


Figure 9: Direct Attach: Server Port to Switch Port Installation

In the structured cabling installation with trunk cables preinstalled, the physical connection time and routing of cables is the same: 60 seconds per port to make the physical connection and about 5 seconds per foot to pull the cable. The difference is that the pathway in this case is short and all array cord connections from the top of the panel down to the physical port in the rack are essentially identical.

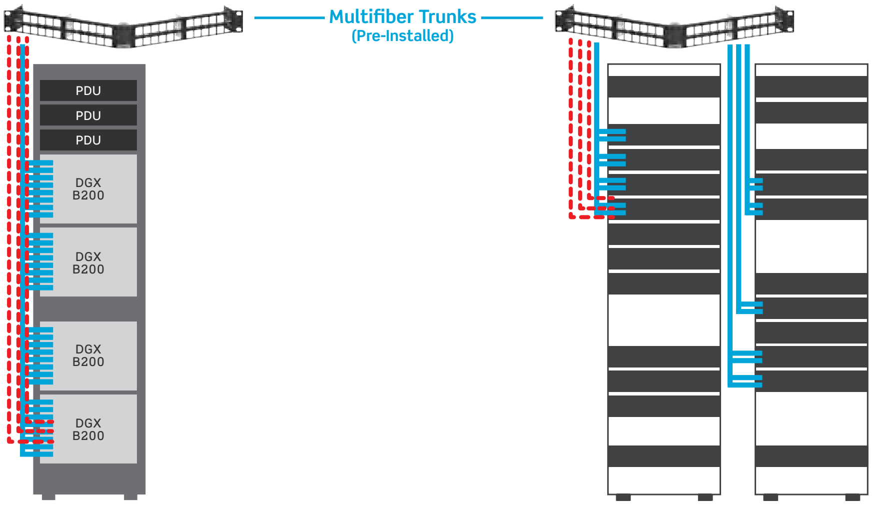


Figure 10: Structured Cabling: Installation with trunk cables pre-installed

As the number of connections required for installation grows into the thousands and beyond — which can easily happen in an AI network deployment — the time savings associated with pre-installing the multifiber trunks grows exponentially. **Figure 11** shows this time savings, comparing the number of hours required to install the AI network with a given number of ports.

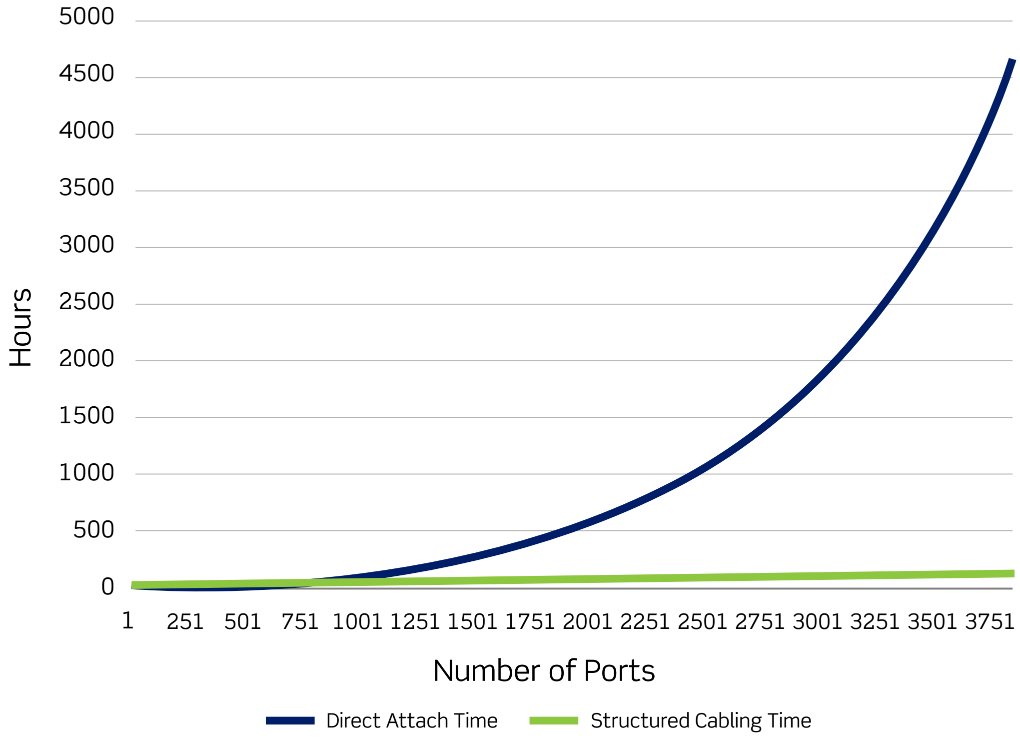
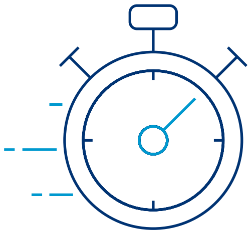


Figure 11: Installation Time: Direct Attach vs Structured Cabling

Addressing Latency



As latency is an important consideration in AI clusters, the question often arises as to which cabling method is best to address latency. With cabling, the key parameter for latency is not the physical optical connection but the overall length of the optical channel. As light passes through fiber, the latency is approximately 5 nanoseconds per meter, and properly designed structured cabling will not add additional latency compared to AOC or direct cabling connections.

In a port-to-port connection scenario using AOCs or longer MPO array cords, on average there will be slightly extra cable lengths as the cords are routed in the overhead tray and down within the rack. Lengths are usually measured at the granularity of 0.5 meter, 1 meter, or multi-meter lengths, depending on the overall length of the connection, so there will be some level of excess cable in that path.

However, in a structured cabling design with patch panels either in-rack or mounted overhead, they are already within the desired point-to-point cable path to make the port-to-port connections. Using shorter array cords and multifiber trunks, network designers can be more precise with cable lengths to help minimize excess cable in that optical path. In addition, the use of fixed panels that do not have sliding trays or other moving components that would need additional slack to operate also help with cable length. Ultimately, in aggregate across all of the port connections, the latency would be about the same whether using direct connection cabling or structured cabling.

Leviton Lab testing has confirmed that adding physical connections does not impact latency of an optical channel, as shown in Figure 12. Our engineers inserted as many as 7 physical connections in the path of a 100-meter channel, and tested the same setup across multiple fiber types, including OM4, OM5, and OS2. For every fiber type, the measured latency in the channel was virtually the same, even as connections were added.

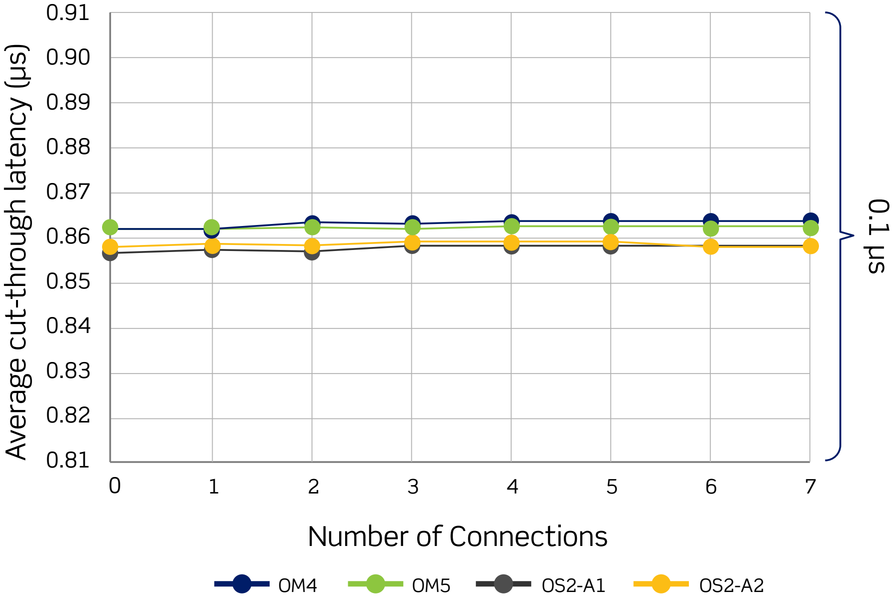


Figure 12: Latency Test with 7 Physical Connections in a 100 Meter Channel

CABLING SYSTEMS FOR SUCCESSFUL AI NETWORKS

As outlined above, network managers need to consider the cabling and connectivity for successful AI applications alongside data rates, scalability, latency, and power consumption. Leviton can help, with a full suite of products available to support AI-enabled networks. With a wide range of pre-terminated fiber cabling for speed of deployment and global availability, Leviton can provide solutions tailored to specific network deployments. Our product portfolio includes multimode fiber APC assemblies, 16-fiber MPO-based connectivity, and end-to-end systems designed for ultra-low loss performance.

In addition, Leviton provides expertise in data center design, rack elevation, and layout optimization to support AI applications effectively for today and tomorrow.

Learn more at **leviton.com/AInetworks**

CN® and SN® are registered trademarks of Senko Advanced Components, Inc.

Today's networks must be fast and reliable, with the flexibility to handle ever-increasing data demands. Leviton can help expand your network possibilities and prepare you for the future. Our end-to-end cabling systems feature robust construction that reduces downtime, and performance that exceeds standards. We offer quick-ship make-to-order solutions from our US and UK factories. We even invent new products for customers when the product they need is not available. All of this adds up to the **highest return** on **infrastructure investment**.

USA — NETWORK SOLUTIONS HEADQUARTERS

2222 - 222nd Street S.E., Bothell, WA, 98021, USA
+1 (800) 722 2082 | infoUSA@leviton.com | Leviton.com/NS

Customer Service
+1 (800) 722 2082
insidesales@leviton.com

International Customer Service
+1 (800) 722 2082 (Option 1)
intl@leviton.com

BERK-TEK™ Cable
+1 (800) 237 5835
infoUSA@leviton.com

Technical Support
+1 (800) 722 2082
appeng@leviton.com

APAC

+85 (2) 3620 2602 | infoAPAC@leviton.com | Leviton.com/NS

Customer Service
+1 (631) 812 6228
infoASEAN@leviton.com

China
+85 (2) 2774 9876
infoChina@leviton.com

South Korea
+82 (2) 3273 9963
infoKorea@leviton.com

CANADA

+1 (800) 461 2002 | infoCanada@leviton.com | Leviton.com/NS

Customer Service
+1 (514) 954 1840
pcservice@leviton.com

Network Solutions products are **available worldwide** in **over 100 countries**. Visit us online at leviton.com/ns to learn more.

EUROPE

Viewfield Industrial Estate, Glenrothes, KY6 2RS, UK
+44 (0) 1592 772124 | infoEurope@leviton.com | Leviton.com/NS/EMEA

Customer Service
+44 (0) 1592 772124
customerserviceEU@leviton.com

Technical Support
+44 (0) 1592 778494
appeng.EU@leviton.com

Benelux
+44 (0) 1592 772124
infoBenelux@leviton.com

Central & Eastern Europe (CEE)
+44 (0) 1592 772124
infoCEE@leviton.com

DACH
+49 (0) 173 272 0128
infoDACH@leviton.com

France
+33 (0) 1709 87826
infoFrance@leviton.com

Italy
+39 (02) 3534896 (Milan)
+39 (06) 68584613 (Rome)
infoitaly@leviton.com

Nordics & Baltics
+46 (70) 9675033
infoNordics@leviton.com

Spain & Portugal
+34 (91) 490 59 19
infoSpain@leviton.com

UK & Ireland
+44 (0) 1592 772124
infoUKandI@leviton.com

LATAM

infoLATAM@leviton.com | Leviton.com/NS

Customer Service
+52 (55) 2333 5963
infoLATAM@leviton.com

Caribbean
+1 (954) 593 1896
infoCaribbean@leviton.com

Colombia
+57 (1) 743 6045
infoColombia@leviton.com

Mexico
+52 (55) 2128 6286
lsamarketing@leviton.com

MIDDLE EAST & AFRICA

Bay Square, Building 3, Office 205, Business Bay, Dubai, UAE
+971 (4) 247 9800 | infoMEA@leviton.com | Leviton.com/NS

Customer Service
+971 (4) 247 9800
lmeinfo@leviton.com