

How Connectors Keep Data Centers Fast, Efficient, and Cool

Connectors' Essential Role in Powering AI

By David Pike for Mouser Electronics



(Photo source: Pakawadee/stock.
adobe.com; generated with AI)

Artificial intelligence (AI) workloads in data centers are projected to grow 350 percent before the end of the decade.¹

That level of growth relies on more than just faster chips. It puts pressure on the entire infrastructure, spurring developers to rethink how computing power and data move through the system so it can move faster and handle more, without creating losses at the connection points.

Data centers can support this growth. Whether owned by a single enterprise or the leading hyperscale providers, data centers are home to the extensive computing and storage systems required to power AI. However, their capabilities depend on the physical infrastructure that connects them.

This paper will focus on the critical role of connectors in the latest data center applications. In doing so, it will highlight three key aspects of connector design: high-speed data, efficient power delivery, and effective thermal management.

Data Connections

AI is optimized for analyzing large datasets to uncover patterns and generate insights. Even if this were the limits of its capabilities, AI would still demand some of the most significant concentrations of computing power ever assembled. However, the development of AI systems is accelerating their expansion into autonomous decision-making and automation at a scale and speed beyond human capacity. This evolution is driving the rapid growth of the global data center market today.²

Modern data centers rely on a combination of components. Graphics processing units (GPUs) handle the massive parallel processing required for deep learning workflows, allowing them to conduct the large numbers of computations needed by deep learning models. These are often paired with AI accelerators, such as application-specific integrated circuits (ASICs) and field-programmable gate arrays (FPGAs), that are designed to perform specific tasks with low latency and high efficiency.

These computing elements depend on high-speed access to data. AI training requires fast-response storage (e.g., solid-state drives) and high-capacity arrays for bulk data handling.

Rather than functioning as a monolithic entity, data centers rely on constant coordination among thousands of interconnected systems. The largest hyperscale data centers feature tens of thousands of GPUs³ and depend on the free flow of information throughout the facility. Keeping data flowing efficiently between processors, memory, and storage without loss depends on connectors designed to meet these requirements.

Meeting the Standards

When it comes to defining high-speed data connections between GPUs, storage devices, and the Ethernet hardware that links them to the outside world, PCI Express® (PCIe) is an industry-wide standard. The PCIe specification governs how data is configured and the physical layer, including card form factors and the connectors that link them.

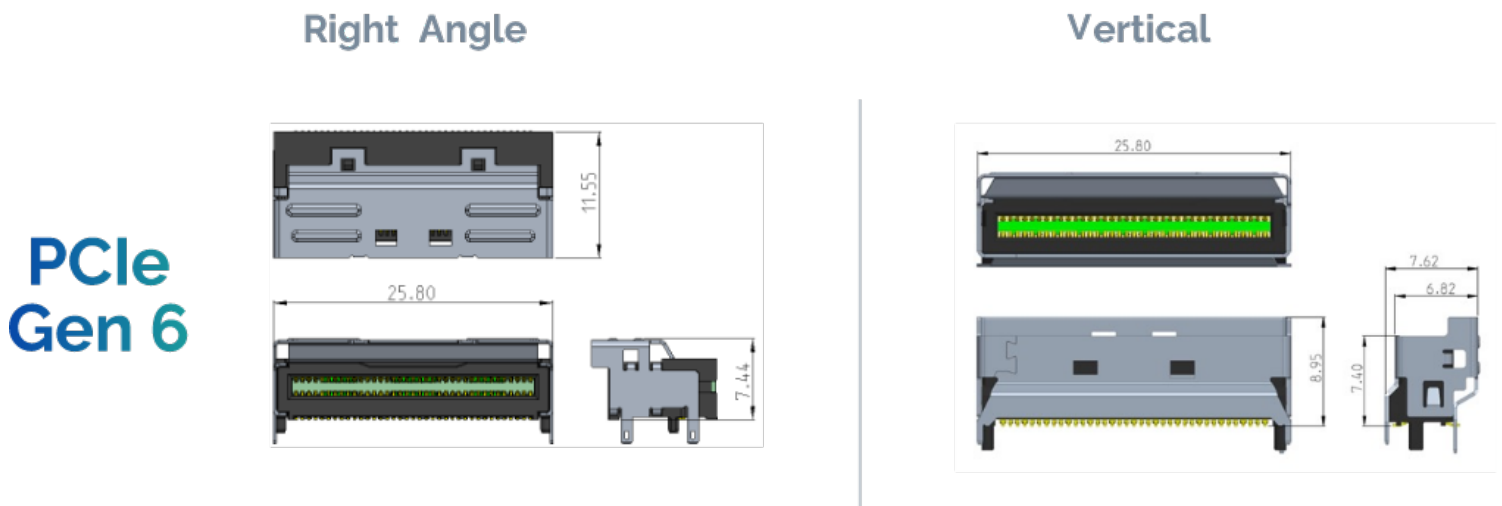


Figure 1: PCIe Gen 6 Card Edge Connectors support up to 64GT/s and are SFF-TA-1016 compliant, with a new inner structure design but the same interface as the PCIe Gen 5 solution. (Source: Amphenol; edited by Mouser Electronics)

The fifth generation of PCIe ([PCIe 5.0](#)) uses non-return-to-zero (NRZ) signal modulation to achieve transfer rates of up to 32 gigatransfers per second (GT/s). NRZ modulation uses two voltage levels to carry one bit of data per clock cycle.

The sixth generation ([PCIe 6.0](#)) doubles this data rate using pulse amplitude modulation 4-level (PAM4). This technique uses four voltage levels instead of the two employed by NRZ signaling, meaning 2 bits of data can be transmitted per clock cycle, thereby doubling the overall data transfer rate.

PCIe connectors are vertically or horizontally mounted card-edge sockets (depending on the form factor) that affix directly to a board (**Figure 1**). They allow for quick installation of any PCIe-compatible device, including GPUs, ASICs, and other high-speed peripherals. These connectors enable high-bandwidth data transfer, helping to minimize latency and prevent issues across the system.

Cross-Generational Compatibility

Regardless of the generation, PCIe connectors conform to a standard physical form factor and PCB footprint, making them scalable and flexible. The standardized interface allows users to upgrade individual modules in response to changing demands. Alternatively, the backward compatibility of PCIe connectors allows existing equipment to be reused where performance allows. Choosing the connector that matches each requirement allows the user to balance performance with cost.

Beyond speed and compatibility, PCIe connectors offer reliable data transmission. At high data rates, signal integrity is important because even minor disruptions can cause data loss. Connectors must



Figure 2: Connectors must be designed for durability and signal integrity in dense data center environments (Source: Matthias/stock.adobe.com)

minimize insertion loss, return loss, and crosstalk across multiple high-speed lanes. Effective electromagnetic interference (EMI) shielding is also essential to protect sensitive signals from external noise, especially in tightly packed server environments. Mechanical durability matters, too. Connectors must handle repeated insertions, thermal shifts, and vibrations without degrading performance (**Figure 2**).

With these various data transmission requirements, the multifaceted characteristics of high-performance connectors make them essential to data center architecture, ensuring that infrastructure can grow with changing demands.

[Amphenol](#) manufactures a comprehensive range of PCIe connectors (**Figure 3**) that conform to the widely accepted standard and exceed the performance required for even the latest applications. Amphenol PCIe connectors support data rates up to 64 GT/s and are rated for 50 mating cycles, offering the signal integrity and durability required for high-performance data center applications.

While reliable data connections are essential for data center operations, they are not designers' only consideration. Without the power to feed the many GPUs, accelerators, and network connections, data centers cannot deliver the full potential of AI.

Power to the Data Center

Connectors do more than just move data in a data center. They also play a critical role in power delivery. Even before the widespread emergence of AI, data centers were considerable power consumers. Estimates from the International Energy Agency (IEA) state that energy consumption in data centers during 2022 represented about 2 percent of global energy production. This number was recorded just before the sharp rise in generative AI workloads.⁴

The full effect of this growth in AI services is reflected in the potential rise of energy consumption. The IEA report also predicts that data center energy use will likely double by 2026, representing approximately 4 percent of the global energy market.

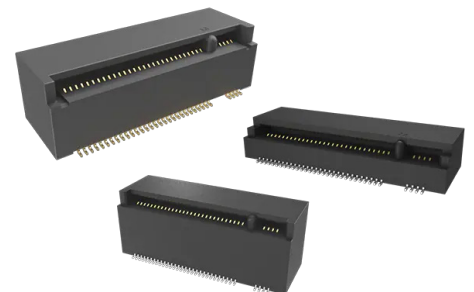


Figure 3: Amphenol Communications Solutions PCIe M.2 Gen 5 High-Density Card Edge Connectors. (Source: Mouser Electronics)

The Scale of Demand

Data centers' power consumption has far more significant local effects. For example, Ireland's location as a bridge between the US and European markets has made it an attractive hub for data services, supported by government incentives. Investment in this technology has directly impacted the energy sector, with data centers consuming 18 percent of Ireland's national energy production in 2022.⁵ The activities of large cloud and internet service providers in the country could increase this figure to 32 percent in the next few years, almost one-third of the country's total energy production.⁶ Not only does this place an enormous burden on energy production, but it also challenges the infrastructure required to distribute it.

This trend is mirrored on a smaller scale. High-performance AI accelerators, such as the NVIDIA H100, consume 700W, while some configurations exceed 1000W. As a result, the infrastructure within the data center must deliver high power densities efficiently and reliably.

Data Center Infrastructure

The path that power takes through the data center is complex. Unlike most consumer power needs, the scale of a data center's needs demand a far larger supply. As a result, power will arrive at much higher voltages directly from the national grid, increasingly supplemented by on-site supplies generated by wind or solar technologies.

Energy is routed to an uninterruptible power supply (UPS), which acts as both a voltage conditioning layer and a backup to cater for momentary fluctuations in supply. The power then flows to power distribution units (PDUs), which divide it into manageable segments, before reaching remote power panels (RPPs) that distribute it further to rows of server racks. This hierarchy distributes the power from transformers and UPS systems to individual racks and devices, delivering power at safe, usable voltages while providing redundancy and scalability.

At each leg of this journey, the power passes through connectors. With their ability to help provide flexibility to any power distribution system, connectors continue to play a vital role in the power infrastructure of data centers. They provide a stable and safe link from one infrastructure element to the next, while retaining the ability to reconfigure the system as required. This flexibility allows operators to upgrade or reconfigure systems as workloads and hardware requirements change.

The Impact of Connectors on Efficiency

Connectors have a significant and measurable impact on the overall efficiency of data centers. Every connector interface has the potential to introduce resistance, which in turn can reduce efficiency, create heat, and even cause equipment failure. In applications that transfer high currents, such as those found in data centers, even a slight increase in electrical resistance has significant effects. Therefore, connector design is critical for energy efficiency.

Perhaps the greatest challenge for data center infrastructure is power density. In conventional applications, the simplest way to transmit more power through a connector is to make it larger. However, as the demand for data increases, so does the number of AI accelerators and GPUs in server racks. Combined with the energy demands of the latest chipsets, these increases require a higher power density.

To meet these requirements, connector manufacturers must find new methods to increase the power density. This is achieved by making each terminal within the connector more efficient. Using high-conductivity materials and designs that maximize the contact area between connectors will ensure that resistance is as low as possible.

For example, Amphenol Positronic [ORV3 power connectors](#), which conform to Open Compute Project (OCP) standards, are designed to provide solutions for power shelf and server applications within data centers. The connectors' single-row, slimline housing design makes them ideal for 1U rack-mounted power supplies or distribution systems. The contacts of the ORV3 series are machined from high-conductivity copper alloy to ensure the lowest possible electrical resistance. Additionally, their closed-entry design ensures maximum contact area between male and female contacts.

The ORV3 contacts are recessed within the connector insulator for improved safety, addressing a critical concern for high-power applications. The insulator is manufactured from high-performance thermoplastic with a maximum operating temperature of 125°C to ensure it can withstand high ambient temperatures without damage.

While choosing the right connector directly affects data center efficiency and operating costs, it also impacts heat. As power demands rise, so do temperatures inside server racks. Low-resistance connectors help reduce heat generation, improving both efficiency and reliability.

Rising Temperatures

The modern data center requires extensive data transmission and power connectivity infrastructure. However, the transmission and consumption of electrical energy create heat. As the workloads created by AI continue to grow, so will the heat they generate. High-density servers, GPU accelerators, and memory devices are all forcing designers to face thermal loads that traditional technologies are ill-equipped to handle. In this environment, connectors must play their part by minimizing their heat contribution and providing new and innovative approaches to thermal management.

Heat generation is a function of the equation for power loss ($P=I^2R$). Known as Joule's law, this equation shows that energy loss (P) is proportional to the resistance (R) of the circuit multiplied by the square of the current (I). This energy will be lost as heat, increasing the temperature of the cables and connectors carrying the power while also heating the environment around them.

The Effects of Heat

Heat is more than just an unwanted side effect of power delivery. It can significantly impact the operation of any data center and severely affect the long-term reliability of electronic systems. Thermal cycling can cause materials to expand and contract, leading to wear and degradation, especially in critical components like insulation on cables and connectors.

Heat also affects semiconductor life cycles. A well-known rule, based on the Arrhenius equation model, states that for every 10°C rise above a device's optimal operating temperature, mean time between failures (MTBF) decreases by as much as 50 percent. These elevated temperatures also affect computational performance. Modern processors include thermal protection mechanisms, such as thermal throttling, which reduces clock speed when temperatures rise. This delay could be disastrous for any safety-critical application that relies on AI analysis.

Conventional Thermal Management

Thermal management is a familiar challenge for designers that has become even more critical with the high workload demands of the latest AI accelerators. Conventional air cooling—using fans and blowers to carry heat away from critical components—is an effective but expensive method of thermal management. Studies by the IEA suggest that as much as 40 percent of the energy consumption in modern data centers is devoted to cooling (**Figure 4**).⁷



Figure 4: Cooling is one of the biggest engineering challenges for modern data centers. (Source: Maryna/stock.adobe.com)

In the pursuit of data center infrastructure that reduces heat production, connectors can positively contribute to thermal management. Many of the connectors used in rack-based systems are linear, which have the potential to disrupt the free flow of air through servers. In some applications, it is possible to reorient connectors on printed circuit boards (PCBs) for better airflow. For other designs where this is not possible, the design of the connector itself can play a role. Amphenol Positronic's Scorpion power connector system addresses airflow needs with a modular design that lets users configure contact arrangements to better distribute thermal load across the connector. Equally important is the option to incorporate cavities along the connector to allow airflow through the insulator. This can help reduce the connector's temperature while allowing air to flow freely over the PCB.

Even with these innovative features, the extreme thermal loads created by the latest chipsets challenge the effectiveness of air cooling. The low thermal capacity of air and the high energy costs created by large cooling systems have forced engineers to explore liquid cooling technologies. The higher thermal capacity of liquids makes them significantly more effective for cooling purposes.

The Advantages of Liquid Cooling

From our earliest experiences at school and home, we are taught that one should never mix liquids and electricity. Therefore, placing an entire printed circuit board into liquid can appear counterintuitive. However, non-conductive or dielectric liquids are entirely safe for use with electronic systems, and immersing a PCB in such a liquid provides uniform cooling without the need for fans. This technique not only reduces energy and maintenance costs but also contributes to a far quieter workspace.

The efficiency of immersion cooling is complicated by the requirement for specialized enclosures. A key requirement of any data center

is its ability to be scaled or reconfigured to respond to changing demands. The infrastructure that supports immersion cooling is difficult to retrofit, so it is better suited for greenfield installations or dedicated AI elements within a larger conventional data center.

An alternative, cold plate cooling, is well-suited for scalable installations. This technique, also known as direct-to-chip cooling, places thermally conductive plates directly onto the heat-producing component. Coolant is then circulated through channels inside the plate to draw heat away from the component's surface.

Cold plate cooling offers several advantages for operators. It can be easily retrofitted to existing equipment and is self-contained, allowing users to introduce the technology as required. This technique also works well alongside conventional air-cooling systems, making it an attractive solution for phased introduction.

New Challenges for Connectors

Liquid cooling places new requirements on connector design, combining the easy operation of electrical connectors with the reliability required of liquid transfer. Unlike conventional connectors, these components need to withstand hydraulic pressure and resist corrosion. They must also be compact to meet data centers' ever-increasing density requirements.

The Amphenol LTW [SnapQD Liquid Cooling](#) connector series is specifically engineered for cold plate cooling. The series features quick-disconnect couplings to support the flexibility required by data center operators, while delivering robust sealing and low fluid resistance. The connectors' small size helps minimize their impact when retrofitting these systems into existing installations. Moreover, they are certified to OCP standards, making them an ideal solution for integration into next-generation racks that require both liquid and electrical connectivity in dense configurations.

Conclusion

As AI workloads become more demanding, data centers must deliver more speed, power, and thermal efficiency without wasting energy or space. These demands make connector choice a critical detail. The right connector reduces losses and supports scalability across data, power, and thermal systems.

Amphenol designs connector solutions that meet the demands of high-speed, high-power, and thermally intensive AI workloads, supporting scalable, efficient data center infrastructure. The company has created families of products that exceed expected standards, providing solutions for the latest AI applications.

Sources

1. <https://www.mckinsey.com/featured-insights/sustainable-inclusive-growth/charts/data-center-demands>
2. <https://www.datacenterknowledge.com/hyperscalers/hyperscalers-will-command-60-of-global-data-center-capacity-by-2030-report>
3. <https://www.businessinsider.com/elon-musk-xai-data-center-colossus-power-memphis-2025-4>
4. <https://iea.blob.core.windows.net/assets/6b2fd954-2017-408e-bf08-952fdd62118a/Electricity2024-Analysisandforecastto2026.pdf>
5. <https://www.cso.ie/en/releasesandpublications/ep/p-dcmec/datacentresmeteredelectricityconsumption2022/>
6. <https://www.irishtimes.com/business/2024/01/24/data-centres-to-consume-one-third-of-states-electricity-by-2026/>
7. <https://www.iea-4e.org/wp-content/uploads/2025/05/Data-Centre-Energy-Use-Critical-Review-of-Models-and-Results.pdf>