Networking Technologies in AI Backend Networks: The Battle of InfiniBand™ vs. Ethernet

Ethernet protocol has traditionally dominated data center networks, in both cloud and enterprise environments. Its long-term evolution and robust ecosystem of suppliers and experts with decades of experience have given it a significant edge over other networking protocols. However, the rise of Generative AI (GenAI), exemplified by large language models (LLMs) like ChatGPT, and the market dominance of leading graphics processing unit (GPU) suppliers like NVIDIA, have brought the InfiniBand networking protocol into the spotlight.

InfiniBand is a high-speed networking technology and industry-standard specification primarily designed for high-performance computing (HPC) applications. NVIDIA's acquisition of Mellanox (a leading InfiniBand supplier) and their subsequent promotion of the protocol for their AI platforms have propelled InfiniBand to become the leading networking protocol in AI backend networks.

That doesn't mean Ethernet is giving up without a fight, however. This analysis examines these two competing technologies in AI backend networks and provides insights into the future of AI networking applications.

AI clusters: A fabric of intelligence

AAI cloud data centers comprise two critical networks—frontend and backend. The frontend network, dominated by Ethernet, serves as the client-facing interface that connects AI to the broader cloud infrastructure and ultimately to the outside world. The backend network, where InfiniBand has gained prominence, houses the AI cluster, a fabric of interconnected GPUs that deliver AI functionality. Figure 1 illustrates the frontend and backend networks in an AI data center.

AI functionality can be predictive AI or GenAI. Predictive AI utilizes data to forecast or infer likely predictions and outcomes. GenAI

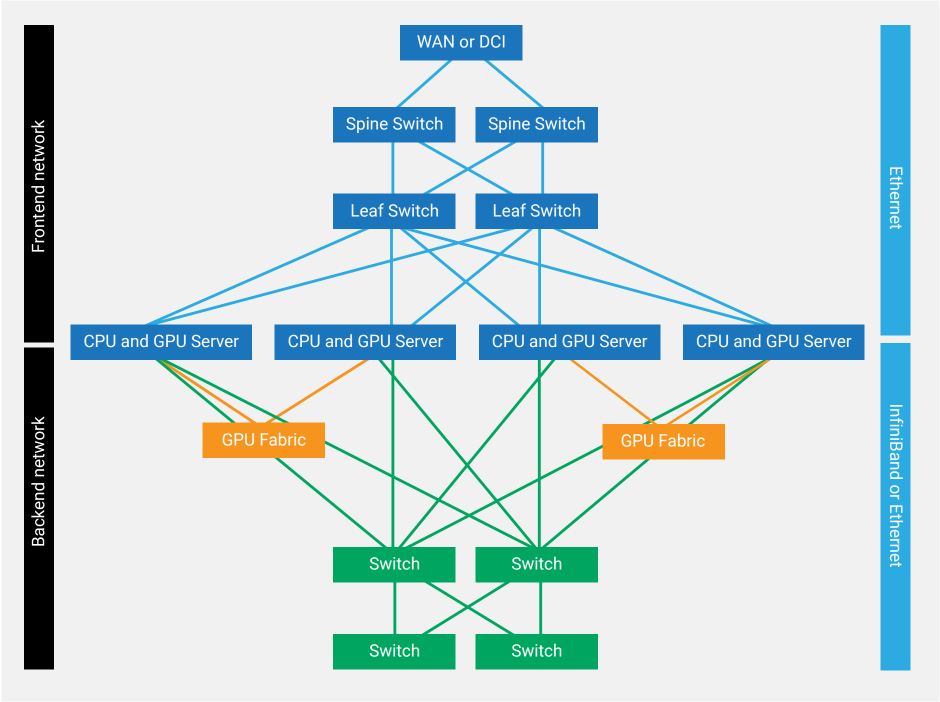


Figure 1: AI data centers include both frontend and backed networks.

utilizes data to generate something new, such as text, music or graphics. Both types of AI functionality necessitate "on-the-job" training of AI models before they can provide services. This training process occurs within the GPU fabric of the AI backend network.

A prime example of AI training is a machine learning (ML) model that analyzes extensive datasets to learn parameters, enabling them to recognize patterns, make decisions or predict outcomes. In 2020, OpenAI released the Generative Pre-Trained Transformer 3 (GPT-3), a state-of-the-art LLM with 175 billion parameters and, while official numbers have not been made available, GPT-4 is rumored to include 1.8 trillion parameters.[1] Meanwhile, Meta's deep learning recommendation model (DLRM) boasts several trillion parameters and is poised for further growth. This training process is computationally intensive, requiring powerful processors, vast amounts of memory, and high-speed data storage and transfer capabilities.

The backend network plays a critical role in both the training and inference phases of AI workloads. During training, the backend network facilitates the transfer of large volumes of data between storage, compute nodes, and other components. This requires high throughput, low latency, and the ability to handle data traffic bursts without congestion or loss. When the network underperforms, expensive AI clusters are underutilized. Therefore, the infrastructure interconnecting these compute resources must be as efficient and cost-effective as possible.

The impact of network performance on AI

Primarily designed for high-performance computing (HPC) environments, the InfiniBand protocol offers exceptionally low latency and high bandwidth, making it ideal for applications like AI training that require a predictable and lossless network fabric. Compared to traditional Ethernet, InfiniBand exhibits lower latency and fewer congestion-related packet losses. In terms of bandwidth, Ethernet and InfiniBand offer comparable performance, with Ethernet enjoying an edge in market adoption. Let's delve deeper into latency, packet losses, and bandwidth considerations.

Latency

Latency, in simple terms, is the time it takes for data to transmit from one point to another on a network. In AI networks, latency occurs at two locations—at the network switch and the network interface card (NIC). Tail latency, the delay experienced by the slowest packets, is especially concerning as it can significantly impact AI training times.

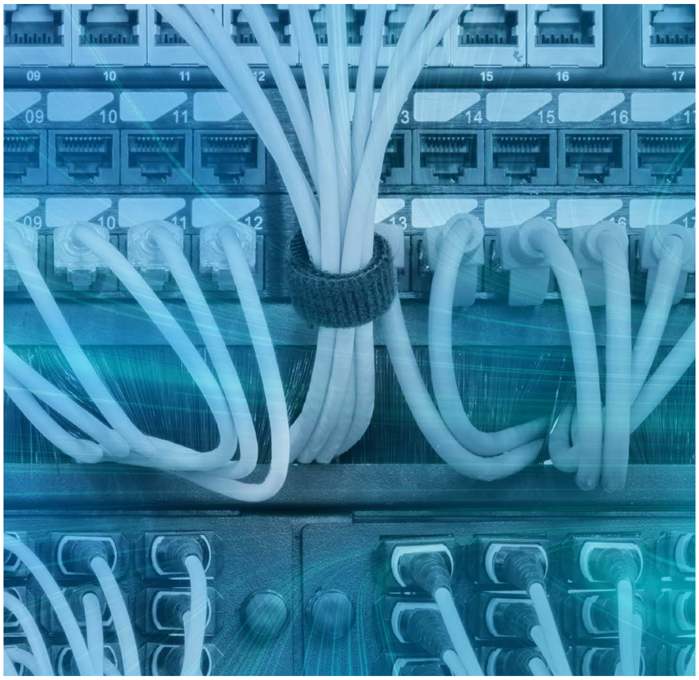
Traditional Ethernet switches typically use a “store-and-forward” switching technology. This involves receiving and storing the entire data packet within the switch, checking it for errors, and transmitting it to the destination. Other switch types use the faster “cut through” technology where the switch starts forwarding the packet as soon as it reads the destination address, without waiting for the entire packet to be received. This means minimal error checking, but lower latency compared to “store-and-forward.” Many modern Ethernet AI switches can dynamically adjust between these modes based on network conditions and requirements, striking a balance between speed and reliability.

At the NIC level, InfiniBand uses Remote Direct Memory Access (RDMA) technology that allows two servers to read and/or write to each other’s memory without involving the server’s processor, cache or operating system (i.e., kernel). Network, storage, and compute applications benefit from this technology. By bypassing the networking kernel, RDMA frees up the CPU, lowers latency and provides much faster data transfer between two servers. Modern Ethernet NICs support AI implementations using RDMA over Converged Ethernet (RoCE). This enables direct memory exchange over an Ethernet network, effectively allowing Ethernet to achieve latency levels comparable to InfiniBand.

Packet Loss

Packet losses—and the subsequent re-transmission of data packets—significantly impact the overall performance of the AI backend network, making a lossless network protocol crucial for optimal operation. InfiniBand is designed as a lossless networking protocol, incorporating congestion management, buffer control, and quality of service (QoS) mechanisms.

While Ethernet is not inherently lossless, it can be configured to operate in a lossless manner by implementing specific techniques and methods. These include priority flow control (PFC), acknowledgment congestion control (ACC) and enhanced transmission protocols like the aforementioned RoCE to bring Ethernet close to the lossless performance of InfiniBand.



Bandwidth

Last, but not least, is the importance of network bandwidth. While both Ethernet and InfiniBand currently support 800G as their fastest data rate, Ethernet is poised to achieve 1.6T by the beginning of 2027 through ongoing standards development, with the development of 3.2T also in sight. A key factor is the market adoption of transceivers supporting these faster speeds.

InfiniBand currently uses 800G with fiber-optic transceivers featuring 2x400G ports, enabling 400G bandwidth per GPU in an AI network. In contrast, the broader and more established Ethernet ecosystem already offers fiber-optic transceivers featuring a single native 800G port.

The outlook ahead

New “Ethernet-boosting” technologies have been developed and are in the pipeline. The Ultra Ethernet Consortium (UEC), established in 2023 by leading Ethernet vendors and major users, is working to modernize RoCE to improve bandwidth, latency, tail latency and scalability. The Ultra Ethernet Transport (UET) protocol is central to the UEC’s efforts. This new protocol optimizes next-generation AI and HPC networks through multiple innovations that increase network utilization and reduce tail latency, both critical factors for accelerating AI training times. Furthermore, the UEC is defining these specifications while preserving the following key Ethernet advantages, as outlined in a UEC whitepaper[1]:

- **A robust, multi-vendor ecosystem,** comprising interoperable Ethernet switches, NICs, cables, transceivers, optics, management tools and software from numerous participants.
- **Proven addressing and routing scale,** leveraging the scalability of IP networks for rack-scale, building-scale, and data center-scale deployments.
- **A comprehensive suite of tools for testing,** measuring, deploying, and efficiently operating Ethernet networks.
- **A proven track record of cost reduction** driven by a competitive ecosystem and economies of scale.
- **A demonstrated ability to rapidly advance,** illustrated by IEEE Ethernet standards which have quickly and regularly progressed across various physical and optical layers.

With these advancements, it appears that Ethernet will likely do in AI backend networks what it has done in frontend networks—dominate. Chip giants like Intel and AMD support Ethernet with their new-generation AI nodes, and major switch manufacturers (including NVIDIA) are implementing enhanced Ethernet into their systems. In fact, one of the world's largest AI clusters was recently built using Ethernet. Leading market research firms in the AI networking environment appear to have also already reached a consensus, predicting that Ethernet-based AI implementations will surpass those using InfiniBand by 2027.

By Hans-Juergen Neithammer,
Solutions Architect, Data Center Solutions

This article was originally published as a chapter in CommScope's 2026 Data Center eBook, available for free here.

[1] Number of Parameters in GPT-4 (Latest Data) by Josh Howarth, published on Exploding Topics, June 17, 2025

EB-117988.3-EN (03/26)