Dell Technologies' Open Ethernet Solutions for Generative AI

Navigating New Frontiers in IT Infrastructure

'AI Fabric (backend switching for GPU-to-GPU connectivity) is forecasted to grow from $1.2B (2022) to $15.2B (2027), five-year CAGR of 65%.

Ethernet is expected to attain 32% revenue share and 37% port shipment for AI Fabrics (2027)'

Dell'Oro Research[1]

GPU performance is highly dependent on network performance. With many AI workloads running on large clusters of servers requiring constant communication between compute nodes and storage, they need robust networking to avoid bottlenecks. If networking performance is insufficient for the workload, GPUs will become idle, and training and inference times will increase, slowing data processing and time to insights.

Introduction: Generative AI's Network Demands

As generative AI (GenAI) solutions continue to evolve, pushing the boundaries of data processing and computational needs, IT infrastructures are being challenged to find ways to support the immense requirements of these environments. These models, particularly large language models (LLMs), demand not just more infrastructure but also carefully architected systems to manage the massive connectivity needs across GPU clusters. Traditional network solutions are quickly becoming bottlenecks, threatening the viability and success of GenAI initiatives. AI fabrics require low latency, lossless performance and maximum bandwidth. Massive data and application processing requirements drive increased requirements for both front- and back-end fabrics.
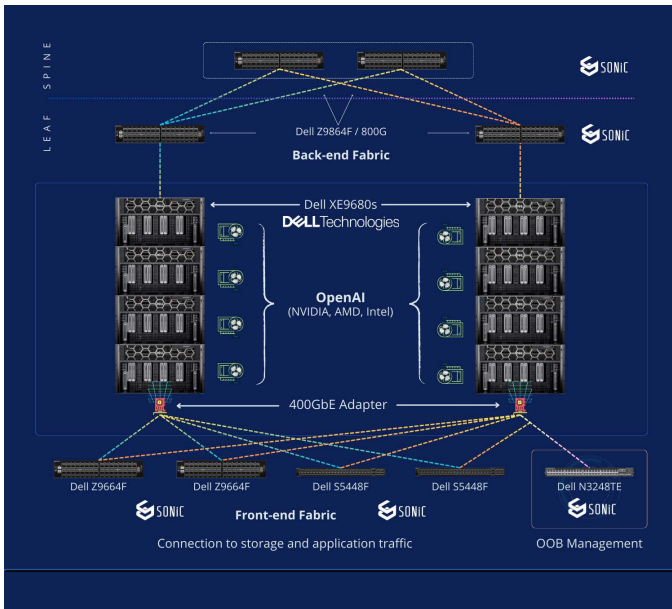
Challenges and Needs in GenAI Infrastructure Deployment

Deploying GenAI technologies introduces a range of challenges, from the technical complexities associated with new architectures to the scarcity of skilled professionals capable of managing such deployments. Solutions that rely on proprietary technologies, like InfiniBand, add another layer of complexity, limiting resource availability and complicating integration with existing monitoring or orchestration platforms. Additionally, the high costs, lengthy evaluation times and vendor lock-in associated with proprietary solutions pose significant barriers, especially in an era of supply chain uncertainties. These challenges underscore the urgent need for open, flexible, and robust GenAI infrastructure solutions that can accommodate the unique demands of GenAI workloads.

Dell Technologies' Approach to GenAI Networking

In response to these challenges, Dell Technologies has pioneered comprehensive and open Ethernet-powered solutions designed to meet the intricate demands of GenAI infrastructure. Leveraging its extensive experience in AI, modeling, and high-performance computing (HPC) environments, Dell Technologies offers a suite of solutions that address both front-end and back-end requirements. From modular compute systems optimized for acceleration, like the Dell PowerEdge XE servers, to AI-focused storage solutions such as PowerScale, Dell Technologies provides the essential components for a successful GenAI deployment. Central to this approach is the deployment of next-gen Ethernet fabrics powered by advanced network silicon. With the **Dell PowerSwitch Z9864-ON** delivering **800GbE** of nonblocking network performance critical for GenAI applications, customers can deploy AI clusters with low latency and high throughput using high bandwidth switching and new features found in **Dell's Enterprise SONiC Distribution by Dell Technologies** like Advanced Routing, RoCEv2, Enhanced Hashing, and Priority Flow Control, for enhanced fabric performance and better congestion monitoring.

GenAI Fabric Architecture example

Innovations in GenAI Networking by Dell Technologies

Dell Technologies stands at the forefront of GenAI networking innovation, offering solutions that meet the requirements of GenAI environments today, and tomorrow, from the edge to the core to the cloud. By focusing on open and extensible solutions, leveraging merchant silicon and open-source based software, Dell Technologies ensures maximum performance and flexibility.

The use of commercially available open source network operating systems like SONiC, coupled with Dell Technologies' active participation in, and contribution to the **Ultra Ethernet Consortium (UEC)**, underscores its commitment to open standards and collaborative development in the Ethernet space. These efforts are aimed at ensuring that Ethernet continues to play a pivotal role in supporting the next generation of AI environments.

The Dell PowerSwitch Z-series, utilizing state-of-the-art silicon, provides the backbone for a high-performance, scalable network capable of supporting thousands of nodes, thereby addressing the connectivity challenges inherent in GenAI applications.

Accelerating GenAI Deployment with Dell Technologies

The surge in GenAI has brought with it a host of challenges for IT infrastructures, demanding a new approach to networking that is both innovative and flexible. Dell Technologies answers this call with open, Ethernet-powered solutions that not only meet the immediate needs of GenAI deployments but also lay the foundation for future advancements.

To help take the guesswork out of AI hardware solutions, Dell offers lab-validated reference architectures optimized for AI workloads. These Validated Designs include architectural concepts, full solution overviews, performance and other lab validations proving the solution's capabilities on the workload for which it was designed. Go from AI-possible to AI-proven with validated solutions that make it easier to deliver faster, deeper insights.

By choosing Dell Technologies, organizations gain a partner with the expertise, an end-to-end solution stack, and the commitment to ensure the success of their GenAI initiatives. With Dell Technologies, businesses are equipped to navigate the complexities of GenAI architectures, ensuring their projects are not only viable but positioned for success.

Accelerate deployment and time-to-value for your GenAI environments, reducing risk and operational complexity with Dell Technologies. We invite you to explore how an open, flexible and sustainable network solution can transform your GenAI initiatives, propelling your business into a new era of innovation and efficiency.

Learn more about Dell Networking

Contact a Dell Technologies Expert

Read the Analyst Brief from ESG

Read the Analyst Brief from IDC

D∢LL Technologies