



Dell Technologies Boosts AI for Enterprises with Dell AI Factory Advancements

November 18, 2024

New AI infrastructure, solutions and services improve performance and accelerate time to value for AI projects, while helping organizations better manage data and improve efficiency

ATLANTA, Nov. 18, 2024 /PRNewswire/ -- SC24



News Summary

- AI infrastructure additions, including new liquid and air-cooled Dell PowerEdge XE servers and Dell Data Lakehouse enhancements, efficiently deliver advanced compute at scale and simplify data access for AI
- Supporting the NVIDIA GB200 Grace Blackwell NVL4 Superchip will be a new Dell PowerEdge XE server designed for the Dell IR7000, supporting up to 144 GPUs in a 500U standard rack
- New solutions accelerate the development and deployment of AI applications
- Dell Professional Services for sustainability, data management and networking help organizations more quickly reach AI goals by reducing complexity

Full story

Dell Technologies (NYSE: DELL) continues to make enterprise AI adoption easier with the Dell AI Factory, expanding the world's broadest AI solutions portfolio.¹ Powerful new infrastructure, solutions and services accelerate, simplify and streamline AI workloads and data management.

"Getting AI up and running across a company can be a real challenge," said Arthur Lewis, president, Infrastructure Solutions Group, Dell Technologies. "We're making it easier for our customers with new AI infrastructure, solutions and services that simplify AI deployments, paving the way for smarter, faster ways to work and a more adaptable future."

New AI infrastructure offers high-density, efficient compute and eases access to data to power AI workloads

Introduced earlier this year, the [Dell Integrated Rack Scalable Systems \(IRSS\)](#) program is a turnkey factory integration program that delivers fully loaded, plug-and-play rack-scale systems with Dell Smart Cooling. IRSS further simplifies deployment of energy efficient AI infrastructure with one-call service and support options for the entire rack. Once set up, Dell takes care of the packaging waste and recycling, and can help with responsible recycling of your old hardware.

Dell has expanded the program to include the **Dell PowerEdge XE9685L** and **Dell PowerEdge XE7740** servers in the standard 19-inch Dell Integrated Rack 5000 (IR5000). The Dell IR5000 is designed for high-density applications in a space-efficient form factor, and delivers high performance without compromising on energy efficiency.

- The **Dell PowerEdge XE9685L** is a dense, 4U liquid-cooled server designed for AI, machine learning, high performance computing (HPC) and other data-intensive workloads. The dual 5th Gen AMD EPYC™ CPUs paired with the [NVIDIA H200](#) or B200 platforms, along with up to 12 PCIe gen 5.0 slots, offer customizable configurations to meet specific compute needs, optimized storage connectivity and maximum IO throughput for demanding workloads. This platform offers industry-leading GPU-density with up to 96 NVIDIA GPUs per rack.
- The 4U air-cooled **Dell PowerEdge XE7740** uses dual Intel® Xeon® 6 with P-cores and up to 8 double-wide accelerators, including Intel Gaudi® 3 AI accelerators or [NVIDIA H200 NVL](#), or up to 16 single-wide accelerators, such as the [NVIDIA L4 Tensor Core GPU](#). This flexibility allows enterprises to right-size their server configuration for the workload at hand, from fine-tuning or inferencing generative AI models, to extracting value from large datasets.

Dell plans to support the upcoming NVIDIA GB200 Grace Blackwell NVL4 Superchip with a new Dell PowerEdge XE server designed for the [Dell IR7000](#), supporting up to 144 GPUs per rack in a 500U standard rack. The IR7000 rack supports large-scale HPC and AI workloads requiring high power and liquid cooling with the ability for near 100% heat capture.

Updates to the **Dell Data Lakehouse** provide enterprises with modern architectures for efficiently managing and analyzing data for AI tasks. This platform is built on AI-optimized hardware and a full-stack software suite, and will expand to include Apache Spark for distributed data processing at scale. For businesses managing massive data volumes, the new Spark integration will provide significant efficiency gains, with a unified approach for data analytics, management, processing and analysis for faster, more actionable insights.

"The continued reimagining and improvement of AI and HPC technologies is critical to accelerating discoveries across industries," said Dan Stanzione, executive director, Texas Advanced Computing Center (TACC). "With Dell's AI solutions we are building complex HPC systems that allow innovation to

thrive in research communities. Ultimately, these advances have the potential to reshape society and drive progress that benefits everyone."

Leading AI partner ecosystem solutions turbocharge AI adoption

Dell is collaborating with partners across the AI ecosystem to power and simplify AI deployments:

- Updates to the **Dell AI Factory with NVIDIA** deliver accelerated performance and reduced time-to-outcomes for AI operations and use case deployment. The NVIDIA HGX H200 and H100NVL support options provide up to 1.9x higher performance compared to NVIDIA HGX H100.
- **Dell Agentic RAG with NVIDIA**, part of the Dell AI Factory with NVIDIA, allows customers to perform complex queries and accelerate their retrieval augmented generation (RAG) operations. The design helps organizations with large data sets use AI agents to improve RAG workflow performance, handle complex queries and deliver higher quality outcomes. The solution leverages Dell PowerEdge, Dell PowerScale and NVIDIA AI Enterprise software and generative AI tools including [NeMo Retriever microservices](#) and the NVIDIA AI Blueprint for multimodal PDF data extraction.
- **Dell Validated Designs for AI PCs** are open-source guides designed to boost the development of AI applications on Dell AI PCs with NPU technology. Developers can easily customize the modular designs to integrate features like LLMs, vision, text and speech into applications. They can deploy the AI applications across multiple platforms, regardless of silicon. With this scalable approach, businesses can automate routine processes, save time, reduce costs and improve data security with on-device AI.

Reduce complexity and ease AI adoption with Dell Services

Finding talent to develop strategies and implement AI continues to be a challenge for organizations. Dell Professional Services help organizations reach their AI goals more efficiently with:

- **Advisory and Implementation Services for Sustainable Data Centers** provide the expertise to create and implement strategies designed to improve energy efficiency which can help lower emissions in the data center with intelligent power and cooling management to help customers achieve their sustainability goals.
- **Data Management Services** deliver an AI-ready catalog by discovering, classifying and refining data, giving organizations simplified access to organized, high-quality data.
- **Design Services for AI Networking** provide customers with a network design optimized for AI workloads that delivers faster speeds, reduced latency and enhanced scalability.²
- **Implementation Services for ServiceNow Now Assist** integrate GenAI capabilities into service management workflows via Now Assist, which streamlines content gathering through AI-driven summaries to automate outcomes and improve productivity.

"Enterprises are pursuing AI to remain competitive in today's digital landscape but they need to harness their proprietary data to differentiate," said Dave Vellante, chief analyst, theCUBE Research. "We're seeing a resurgence in interest for offerings like Dell's AI infrastructure, solutions and services that can help customers prepare their data to extract insights, automate tasks and transform processes. As we continue into this new era of technological innovation, it will be critical for enterprises to continuously learn and invest in innovation to realize the benefits AI promises."

Availability

- The Dell PowerEdge XE9685L will be globally available Q1 CY2025.
- The Dell PowerEdge XE7740 will be globally available Q2 CY2025.
- Updates to the Dell Data Lakehouse are globally available now.
- Dell Validated Designs for AI PCs are globally available [now](#).
- Dell Generative AI Solutions with NVIDIA – GPU update will be available Q4 CY2024.
- Dell Generative AI Solutions with NVIDIA – Enterprise-scale RAG is globally available now.
- Dell Data Management Services are available in select countries now.
- Dell Services for Sustainable Data Centers are available in select countries now.
- Dell Design Services for AI Networking are available in select countries now.
- Dell Implementation Services for ServiceNow Now Assist are available in select countries now.

Additional resources

- Connect with Dell on [X](#) and [LinkedIn](#)
- [Blog](#): Accelerating AI Innovation: New Servers and Integrated Rack Solutions for the Future
- [Blog](#): Revolutionizing AI PC Adoption: Embrace the New Era
- [Blog](#): Your AI Future with Dell Services for Sustainable Data Centers
- [Learn more](#) about the Dell AI Factory

About Dell Technologies

[Dell Technologies](#) (NYSE: DELL) helps organizations and individuals build their digital future and transform how they work, live and play. The company provides customers with the industry's broadest and most innovative technology and services portfolio for the AI era.

Copyright © 2024 Dell Inc. or its subsidiaries. All Rights Reserved. Dell Technologies and Dell are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners.

¹ Based on Dell analysis, July 2024. Dell Technologies offers hardware solutions with AI-powered features and solutions to support AI workloads from AI PCs and workstations to servers for high-performance computing, data storage, cloud native software-defined infrastructure, networking switches, data protection, HCI and services.

² Compared to traditional networks.

 View original content to download multimedia: <https://www.prnewswire.com/news-releases/dell-technologies-boosts-ai-for-enterprises-with-dell-ai-factory-advancements-302307886.html>

SOURCE Dell Technologies

Media Relations: Media.Relations@Dell.com