Dell AI Factory Transforms Data Centers with Advanced Cooling, High Density Compute and AI Storage Innovations

October 15, 2024

New integrated rack, server and storage enhancements power high-performance AI workloads

Round Rock, TEXAS, Oct. 15, 2024 /PRNewswire/ --

News summary

- Dell Integrated Rack 7000 offers liquid cooled high density compute infrastructure with scalable power management
- Dell PowerEdge XE9712 is an NVIDIA GB200 NVL72-based platform targeted toward the largest AI GPU clusters for large language model (LLM) training and real-time inference
- Dell PowerEdge M7725 is a direct-to-chip liquid cooled 5th Gen AMD EPYC-based system for high-performance, dense compute
- Dell PowerScale advancements deliver increased capacity, accelerate AI workload performance and improve data discoverability
- Dell Data Lakehouse and Dell Services simplify how customers access and manage their data
- Dell Generative AI Solutions with Intel provide AI optimized solutions for any size workload

Full story
Dell Technologies (NYSE: DELL) introduces new integrated rack-scalable systems, server, storage and data management innovations to the Dell AI Factory, powering high density computing and AI workloads at scale.

"Today's data centers can't keep up with the demands of AI, requiring high density compute and liquid cooling innovations with modular, flexible and efficient designs," said Arthur Lewis, president, Infrastructure Solutions Group, Dell Technologies. "These new systems deliver the performance needed for organizations to remain competitive in the fast-evolving AI landscape."

The future of accelerated compute with leading cooling innovations
The **Dell Integrated Rack 7000 (IR7000)** handles accelerated computing demands with superior density, more sustainable power management and advanced cooling technologies. This Open Compute Project (OCP) standards-based rack is ideal for large-scale deployment and features a futureproof design for multigeneration and heterogenous technology environments.

Key features include:

- **Designed for density**, the 21-inch Dell IR7000 is designed to support industry-leading CPU and GPU density.
- **Future-ready and efficient**, the rack features wider, taller server sleds to accommodate the latest, larger CPU and GPU architectures. This rack was purpose built for liquid cooling natively, capable of cooling future deployments of up to 480KW, and is able to capture nearly 100% of heat created.
- **Engineered for greater choice and flexibility**, this integrated rack offers support for both Dell and off-the-shelf networking.
- **Deployments are simple and energy-efficient** with Dell Integrated Rack Scalable Systems (IRSS). IRSS delivers innovative rack-scale infrastructure optimized for AI workloads, making the setup process seamless and efficient with a fully integrated plug-and-play rack scale system.

Dell Technologies introduces AI-ready platforms designed for the Dell IR7000:

- Part of the Dell AI Factory with NVIDIA, the **Dell PowerEdge XE9712** offers high-performance, dense acceleration for LLM training and real-time inferencing of large-scale AI deployments. Designed for industry-leading GPU density with NVIDIA GB200 NVL72, this platform connects up to 36 NVIDIA Grace CPUs with 72 NVIDIA Blackwell GPUs in a rack-scale design. The 72 GPU NVLink domain acts as a single GPU for up to 30x faster real-time trillion-parameter LLM inferencing. The liquid cooled NVIDIA GB200 NVL72 is up to 25x more efficient than the air-cooled NVIDIA H100-powered systems.
- The **Dell PowerEdge M7725** provides high performance dense compute ideal for research, government, fintech and higher education environments. Designed to be deployed in the IR7000 rack, the Dell PowerEdge M7725 delivers more compute with improved serviceability scaling between 24K-27K cores per rack, with 64 or 72 two socket nodes, powered by 5th Gen

AMD EPYC CPUs Front IO slots enables high speed IO connectivity and provides seamless connectivity for demanding applications. The server's energy-efficient form factor allows for more sustainable deployments through both direct liquid cooling (DLC) to CPUs and air cooling via quick connect to the integrated rack.

Unstructured storage and data management innovations for the AI era
Dell Technologies unstructured data storage portfolio innovations improve AI application performance and deliver simplified global data management.

Dell PowerScale, the world's first Ethernet storage certified for NVIDIA DGX SuperPOD, delivers new updates that enhance data management strategies, improve workload performance and offer greater support for AI workloads.[1]

- **Enhanced discoverability:** Unlock data insights for faster smarter decision-making using PowerScale metadata and the Dell Data Lakehouse. A forthcoming Dell open-source document loader for NVIDIA NeMo services and RAG frameworks is designed to help customers improve data ingestion time and decrease compute and GPU cost.
- **Denser storage:** Customers can fine tune their AI models by training them on larger datasets with new 61TB drives that increase capacity and efficiency while reducing data center storage footprint by half.[2]
- **Improved AI performance:** AI workload performance is enhanced through front-end NVIDIA InfiniBand capabilities and 200GbE Ethernet adapter support that delivers up to 63% faster throughput.[3]

With new enhancements to the Dell Data Lakehouse data management platform, customers can save time and improve operations with new features like disaster recovery, automated schema discovery, comprehensive management APIs, and self-service full stack upgrades.

Customers can simplify their data-driven journey and quickly scale their AI and business use cases with **Optimization Services for Data Cataloging** and **Implementation Services for Data Pipelines**. These services increase accessibility to high-quality data through discovery, organization, automation and integration.

Dell Generative AI Solutions with Intel for modern workflows
As part of the Dell AI Factory, **Dell Generative AI Solutions with Intel** offers jointly engineered, tested and validated platforms for seamless AI deployment. Featuring the Dell PowerEdge XE9680 and Intel ® Gaudi 3 ® AI accelerators with Dell storage, networking, services and an open-source software stack, these preconfigured, flexible and high performing solutions support a range of GenAI use cases including content creation, digital assistants, design and data creation, code generation and more.

Availability

- The Dell IR7000 will be globally available Q1 CY2025.
- The Dell PowerEdge XE9712 is sampling for select customers now.
- The Dell PowerEdge M7725 will be globally available Q1 CY2025.
- Dell PowerScale updates will be available in Q4 CY2024.
- Dell Data Lakehouse updates will be available in 1H CY2025.
- Dell Generative AI Solutions with Intel will be available in Q4 CY2024.

Additional resources

- Connect with Dell on X and LinkedIn
- Blog: Leading the Charge: Dell's OCP Solutions Propel AI Innovation
- Blog: Unstructured Storage Innovations for AI and Data Management
- Blog: Empowering GenAI Choice: Dell AI Factory Expands with Intel
- Join Dell for its "Making AI Real with Data" event
- 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 internal analysis, Mar. 2024.
2. Based on 61TB SSD availability vs previously max availability of 30TB SSDs. October, 2024
3. Based on 200GbE ethernet connectivity vs previously available 100GbE. October, 2024

View original content to download multimedia:https://www.prnewswire.com/news-releases/dell-ai-factory-transforms-data-centers-with-advanced-cooling-high-density-compute-and-ai-storage-innovations-302275793.html

SOURCE Dell Technologies

Media Relations: Media.Relations@Dell.com