

Accelerate Insights

Discover how the Dell PowerEdge accelerated server portfolio can help you unleash your AI advantage, accelerate your insights, and boost your outcomes.

Deliver the PowerEdge difference

Emerging adoption of AI, ML as well as growth of HPC and VDI is adding complexity to data center operations, as workforce grows globally and remotely, as well as demanding use cases becoming more mainstream.

AI, ML, and HPC applications benefit greatly from optimal performance with servers and complementing GPU accelerators in a production environment, as less optimized all-CPU-based systems have difficulty handling the unique hardware and software requirements.

Consider an infrastructure which can deliver the capabilities to make organizations successful with AI and other demanding workloads. A modern architecture approach where one of the biggest innovations is improved performance with the addition of dense acceleration, at scale can be achieved with Dell PowerEdge servers and innovation:



Maximize Performance
with faster graphics, AI, HPC and vGPU capabilities



Drive productivity
across GPU applications with More processing density



Support business and IT agility with seamless scalability



Drive sustainable operations for improved business results



Assure secure operations across infrastructure, workloads and processes

AI for faster insights and decision-making



AI/Machine Learning

Unlocking more value from customers, operations and data can benefit greatly from an infrastructure designed to model, learn and drive discovery and insights for business growth.



HPC

Realize a performance approach to HPC applications with accelerated compute for genomics sequencing and complex modeling, solve more with faster model training.



NLP and Analytics

Optimize decision-making by accelerating and automating language and data processing to new levels of insights in chatbots, fraud detection, digital banking.



Modeling & Simulation

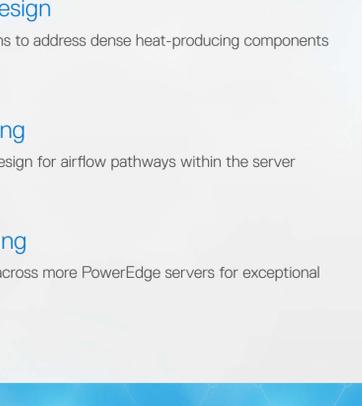
Leverage AI to discover and simulate complex processes and problems. Deploy Digital Twin simulations, drive faster designs for CAD/CAM/CAE.

Capitalize on innovation to realize business benefits

Combine Dell Technologies PowerEdge servers and innovation, alongside with partner technology to extract more value from your infrastructure and your data.

Accelerate Insights

Engineered to optimize the latest technology advances for predictable profitable outcomes



Focus on Acceleration

Support for the most complete portfolio of GPUs, delivering maximum performance for HPC, AI-ML/DL training and inferencing, language and analytics and VDI workloads



Thoughtful Thermal Design

New thermal solutions and designs to address dense heat-producing components



Dell Multi Vector Cooling

Streamlined, advanced thermal design for airflow pathways within the server



Dell Direct Liquid Cooling

Extending liquid cooling support across more PowerEdge servers for exceptional heat removal capability

Explore the Dell PowerEdge Acceleration Optimized server portfolio

No-compromise accelerated AI

XE9680¹ is designed to drive business insights in the most demanding Deep Learning and modeling applications, from large natural language processing models and recommendation engines to complex research and academia problems.

- Highest performance for HPC and Enterprise
- 8x NVIDIA® H100 or A100 Tensor Core GPUs
- Air-cooled operation

Ideal workloads

Large language Models, Natural Language processing, Large Recommendation engine training, Modeling & Simulation, molecular dynamics and genomic sequencing

Applicable GPUs:

NVIDIA H100 SXM or A100 SXM

Dense acceleration

XE9640¹ boosts insights from your growing data sets with AI acceleration technology designed for optimal performance, fastest time-to-value, in a liquid-cooled environment.

- Mainstream 2U enables highest performance AI operations
- 4x Intel® Data Center GPU Max Series
- Liquid-cooled CPU and GPU operation

Ideal workloads

Natural language processing, Large Recommendation engine training, Modeling & Simulation, Artificial Intelligence, ML/DL Training for object recognition

Applicable GPUs:

Intel® Data Center GPU Max 1550 OAM

Purpose-built performance

XE8640¹, helps businesses unlock AI performance with purpose-built performance in a highly dense server for AI, removing traditional computational boundaries of real-time insights.

- Optimized balance of performance for diverse applications
- 4x NVIDIA H100 Tensor core GPUs with NVLink
- Air-cooled operation

Ideal workloads

Medium data set language Models, Natural Language processing, Modeling & Simulation, Artificial Intelligence, ML/DL Training and Inferencing, image recognition

Applicable GPUs:

NVIDIA H100 SXM

Purpose-built scale up server for GPU applications

R760xa¹ maximizes results from AI to Modeling & Simulation applications with maximum flexibility and the latest 4th Generation Intel® Xeon® Scalable processors.

R750xa¹ is optimized to tackle GPU workloads and deliver outstanding performance for demanding and emerging applications.

- Maximize performance
- Front-to-back air-cooled design
- R760xa supports up to 12 Single-wide GPUs or 4 Double-wide GPUs, up to 350W
- Supports all GPU cards

Ideal workloads

AI & ML training and inferencing, data analytics, PC, VDI & Performance graphics

Applicable GPUs:

NVIDIA H100, A100, L40, A40, A30, A16, A10, A2; L4* AMD MI210, MI100.

Superior AI, ML and HPC processing

XE8545¹ delivers optimized CPU and GPU performance for AI and ML training and inferencing by pairing the maximum core count AMD EPYC™ processors, highest performing NVIDIA A100 GPUs, and NVLINK to maximize the time to value.

- Supercharge AI/ML and HPC performance
- Interconnected 4-way NVLINK architecture
- GPU Virtualization

Ideal workloads

AI & ML training and inferencing, HPC, GPU virtualization

Applicable GPUs:

NVIDIA A100 SXM

One-to-one acceleration

R940xa¹ is optimized to tackle workloads that are compute-intensive, combining up to 4 CPUs, up to 112 cores, with four GPUs in a powerful 1:1 ratio to drive artificial intelligence, machine learning and deep learning workloads

- Accelerate applications
- Scale dynamically
- Streamline IT operations

Ideal workloads

GPU database acceleration, data analytics, artificial intelligence, machine learning

Applicable GPUs:

NVIDIA A100, A16

¹Available 1H2023

NVIDIA-Certified Dell Systems brings together NVIDIA GPUs and NVIDIA networking in servers and hyperconverged infrastructure from Dell Technologies in optimized configurations.

These systems are validated for performance, manageability, security, and scalability and are backed by enterprise-grade support from NVIDIA and Dell Technologies.

- Deliver infrastructure to drive a diverse range of accelerated workloads for the enterprise
- Excellent performance
- Reduce time to deployment
- Secure, no-compromise operations and workflows
- Designed for single to multi-node configs, optimal Scale-out and clusters



Achieve more, deliver quick results and maximize efficiency

Dell Validated Designs are purpose-designed with IT's transformation journey in mind to run intelligent applications and processes in the digital business.

Along with Dell PowerEdge servers, Dell Technologies partners and collaborates with industry leaders including Intel, Microsoft, NVIDIA, and others to optimize IT for your critical business workloads together with emerging technologies such as AI, machine learning, and blockchain.

- Validated Designs for AI, including Deep Learning with NVIDIA and Cloudera
- Validated Designs for Data Analytics
- Validated Designs for HPC
- Validated Designs for VDI

Driving a faster time-to-value from your workloads requires optimized approaches which complement your strategy.

Dell PowerEdge accelerated GPU servers for AI deliver the foundation of quality and reliability, along with a worry-free infrastructure optimized for demanding visualization and accelerated business outcomes.

Learn more at

www.Dell.com/Poweredge

www.Dell.com/en-us/dt/servers/specialty-servers/poweredge-xe-servers.htm