

NVIDIA and Dell Technologies Deliver AI and HPC Performance in Leaps and Bounds With Hopper, at SC22

Dell PowerEdge XE9680, based on the NVIDIA HGX H100 architecture, is the first eight-GPU Dell system, built to supercharge advanced workloads.

Author: Dion Harris

Whether focused on tiny atoms or the immensity of outer space, supercomputing workloads benefit from the flexibility that the largest systems provide scientists and researchers.

To meet the needs of organizations with such large AI and high performance computing (HPC) workloads, Dell Technologies today unveiled the Dell PowerEdge XE9680 system — its first system with eight NVIDIA GPUs interconnected with NVIDIA NVLink — at SC22, an international supercomputing conference running through Friday.

The Dell PowerEdge XE9680 system is built on the NVIDIA HGX H100 architecture and packs eight NVIDIA H100 Tensor Core GPUs to serve the growing demand for large-scale AI and HPC workflows.

These include large language models for communications, chemistry and biology, as well as simulation and research in industries spanning aerospace, agriculture, climate, energy and manufacturing.

The XE9680 system is arriving alongside other new Dell servers announced today with NVIDIA Hopper architecture GPUs, including the Dell PowerEdge XE8640.

“Organizations working on advanced research and development need both speed and efficiency to accelerate discovery,” said Ian Buck, vice president of Hyperscale and High Performance Computing, NVIDIA. “Whether researchers are building more efficient rockets or investigating the behavior of molecules, Dell Technologies’ new PowerEdge systems provide the compute power and efficiency needed for massive AI and HPC workloads.”

“Dell Technologies and NVIDIA have been working together to serve customers for decades,” said Rajesh Pohani, vice president of portfolio and product management for PowerEdge, HPC and Core Compute at Dell Technologies. “As enterprise needs have grown, the forthcoming Dell PowerEdge servers with NVIDIA Hopper Tensor Core GPUs provide leaps in performance, scalability and security to accelerate the largest workloads.”

Fresh off setting world records in the MLPerf AI training benchmarks earlier this month, NVIDIA H100 is the world’s most advanced GPU. It’s packed with 80 billion transistors and features major advances to accelerate AI, HPC, memory bandwidth and interconnects at data center scale.

H100 is the engine of AI factories that organizations use to process and refine large datasets to produce intelligence and accelerate their AI-driven businesses. It features a dedicated Transformer Engine and fourth generation NVIDIA NVLink interconnect to accelerate exascale workloads.

Each system built on the NVIDIA HGX H100 platform features four or eight Hopper GPUs to deliver the highest AI performance with 3.5x more energy efficiency compared with the prior generation, saving development costs while accelerating discoveries.

Dell systems power the work of leading organizations, and the forthcoming Hopper-based systems will broaden Dell’s portfolio of solutions for its customers around the world.

With its enhanced, air-cooled design and support for eight NVIDIA H100 GPUs with built-in NVLink connectivity, the PowerEdge XE9680 is purpose-built for optimal performance to help modernize

operations and infrastructure to drive AI initiatives.

The PowerEdge XE8640, Dell's new HGX H100 system with four Hopper GPUs, enables businesses to develop, train and deploy AI and machine learning models. A 4U rack system, the XE8540 delivers faster AI training performance and increased core capabilities with up to four PCIe Gen5 slots, NVIDIA Multi-Instance GPU (MIG) technology and NVIDIA GPUDirect Storage support.

The Dell PowerEdge XE9680 and XE8640 will be available from Dell starting in the first half of 2023.

Customers can now try NVIDIA H100 GPUs on Dell PowerEdge servers on NVIDIA LaunchPad , which provides free hands-on experiences and gives companies access to the latest hardware and NVIDIA AI software.

To take a first look at Dell's new servers with NVIDIA H100 GPUs at SC22, visit Dell in booth 2443.

Original URL: <https://blogs.nvidia.com/blog/2022/11/14/dell-hopper-hpc-ai-sc/>