Supermicro Offers End-to-End Portfolio of NVIDIA GPU Systems

March 18, 2019

SuperServers Optimized for the Full Range of NVIDIA GPUs and Market Segments Including the World's Most Powerful AI Training and Inferencing Systems On-Display at GTC 2019

SAN JOSE, Calif., March 18, 2019 /PRNewswire/ -- **Super Micro Computer, Inc. (SMCI)**, a global leader in enterprise computing, storage, networking solutions and green computing technology, is showcasing the industry's most comprehensive portfolio of NVIDIA® GPU servers at the GPU Technology Conference (GTC) 2019 at the San Jose Convention Center, Booth #829 from March 18-21.

Leveraging NVIDIA GPU technology with Supermicro's system engineering expertise to design the most advanced server systems, Supermicro offers the industry a comprehensive portfolio of NVIDIA GPU systems. From NVIDIA Tesla®, Quadro® and GRID® to the NVIDIA HGX-2 server architecture, Supermicro offers optimized SuperServers for the full range of market segments and applications, including VDI, professional visualization, transcoding, IoT, HPC, AI inferencing and deep learning.

"Our industry-leading portfolio of systems optimized for NVIDIA GPUs continues to grow stronger and stronger as Supermicro produces innovative GPU servers to address the accelerating market demand for a wide range of AI solutions," said Charles Liang, president and CEO of Supermicro. "By optimizing our server designs to fully leverage NVIDIA's latest Tensor Core GPUs, our GPU servers not only maximize system-level performance and efficiency but also offer the most flexible selection of features including the most advanced networking and storage options."

"Enterprise customers will see a tremendous boost in performance and efficiency by leveraging NVIDIA's Tensor Core GPUs housed in Supermicro's GPU-optimized SuperServers," said Ian Buck, vice president and general manager of Accelerated Computing at NVIDIA. "These servers deliver dramatically higher throughput for compute-intensive AI workloads while minimizing power consumption."

AI Inferencing

To address the rapidly emerging high-throughput inference market driven by technologies such as 5G, Smart Cities and IoT devices that generate huge amounts of data and require real-time decision making, Supermicro's new SuperServer 6049GP-TRT provides the superior performance required to vertically scale the technology of modern AI. To achieve maximum GPU density and performance, this 4U server supports up to 20 NVIDIA T4 Tensor Core GPUs, three terabytes of memory, and 24 hot-swappable 3.5" drives. This system also features four 2000-watt Titanium level efficiency (2+2) redundant power supplies to help optimize the power efficiency, uptime and serviceability.

NVIDIA T4 is the world's most efficient GPU for supercharging mainstream enterprise servers for a full range of accelerated applications, including AI training, inference and machine learning workloads. Enterprise customers will benefit from a dramatic boost in throughput, utilization and power efficiency from the NVIDIA T4 GPUs in Supermicro's new high-density servers.

Supermicro has an entire family of 4U GPU systems that support the ultra-efficient NVIDIA T4, which is designed to accelerate deep learning training, inference and machine learning workloads in any mainstream server. The hardware-accelerated transcode engine in NVIDIA T4 delivers multiple HD video streams in real-time and allows integrating deep learning into the video transcoding pipeline to enable a new class of smart video applications. To achieve responsiveness, these models are deployed on powerful Supermicro servers with NVIDIA GPUs to deliver maximum throughput for inference workloads.

NGC-Ready Container Solutions

Supermicro's SuperServer 4029GP-TVRT is NGC-Ready, which helps simplify the actual deployment of AI and HPC applications in the data center. Customers can now run GPU-accelerated software from the NGC container registry, including its expanded HPC and AI software library with new machine learning and analytics containers, with confidence on the 4029GP-TVRT, which is a 4U system with eight NVIDIA V100 Tensor Core GPUs with NVIDIA NVLink™ high-speed interconnect technology.

AI Training and HPC

HPC applications are continuing to grow in complexity as they unlock new scientific insights. Supermicro's new NVIDIA HGX-2 based SuperServer, 9029GP-TNVRT, supports 16 NVIDIA V100 Tensor Core 32GB GPUs connected via NVIDIA NVLink and NVSwitch™ to leverage over 80,000 CUDA cores and deliver unmatched performance for accelerating AI and HPC on premises and in the cloud. This new system can deliver up to 2 PetaFLOPS of performance and occupies just ten units of rack space.

From natural speech by computers to autonomous vehicles, rapid progress in AI has transformed entire industries. To enable these capabilities, AI

models are exploding in size. HPC applications are similarly growing in complexity as they unlock new scientific insights. Supermicro's HGX-2 based SuperServer (SYS-9029GP-TNVRT) will provide a superset design for datacenters accelerating AI and HPC in the cloud. With fine-tuned optimizations, this SuperServer will deliver the highest compute performance and memory for rapid model training.

Data Center Gaming Applications

Supermicro's new SBI-4119MG-X SuperBlade® server modules optimized for NVIDIA GPUs are ideal for data center deployments where total cost of ownership (TCO) is a top priority. This blade solution is suitable for a wide range of data center applications including gaming, and one 8U SuperBlade enclosure supports twenty of these blades and up to 40 GPUs.

For comprehensive information on Supermicro product lines featuring NVIDIA GPUs, please go to https://www.supermicro.com/products/nfo/gpu.cfm.

Follow Supermicro on Facebook and Twitter to receive the latest news and announcements.

About Super Micro Computer, Inc. (SMCI)
Supermicro®, the leading innovator in high-performance, high-efficiency server technology is a premier provider of advanced Server Building Block Solutions® for Data Center, Cloud Computing, Enterprise IT, Hadoop/Big Data, HPC and Embedded Systems worldwide. Supermicro is committed to protecting the environment through its "We Keep IT Green®" initiative and provides customers with the most energy-efficient, environmentally-friendly solutions available on the market.

SMCI-F

View original content to download multimedia:http://www.prnewswire.com/news-releases/supermicro-offers-end-to-end-portfolio-of-nvidia-gpu-systems-300813447.html

SOURCE Super Micro Computer, Inc.

Michael Kalodrich, Super Micro Computer, Inc., PR@supermicro.com