Supermicro Unveils 2 PetaFLOPS SuperServer Based on New NVIDIA HGX-2, the World's Most Powerful Cloud Server Platform for AI and HPC

May 30, 2018

Designed for the Next-Generation of AI, New HGX-2 System with 16 Tesla V100 GPUs and NVSwitch leverages over 80,000 Cuda Cores to deliver unmatched performance for deep learning and compute workloads

TAIPEI, Taiwan, May 30, 2018 /PRNewswire/ -- **Super Micro Computer, Inc. (NASDAQ: SMCI)**, a global leader in enterprise computing, storage, networking solutions and green computing technology, today announced that it is among the first to adopt the NVIDIA® HGX-2 cloud server platform to develop the world's most powerful systems for artificial intelligence (AI) and high-performance computing (HPC).

Supermicro is hosting a Platinum Sponsor Booth at the GPU Technology Conference (GTC) Taiwan 2018, Taipei Marriott Hotel on May 30-31.

"To help address the rapidly expanding size of AI models that sometimes require weeks to train, Supermicro is developing cloud servers based on the HGX-2 platform that will deliver more than double the performance," said Charles Liang, president and CEO of Supermicro. "The HGX-2 system will enable efficient training of complex models. It combines 16 Tesla V100 32GB SXM3 GPUs connected via NVLink and NVSwitch to work as a unified 2 PetaFlop accelerator with half a terabyte of aggregate memory to deliver unmatched compute power."

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 systems will provide a superset design for datacenters accelerating AI and HPC in the cloud. With fine-tuned optimizations, Supermicro's HGX-2 server will deliver the highest compute performance and memory for rapid model training.

"As AI model complexity and size are exploding, researchers and data scientists need new levels of GPU-accelerated computing," said Ian Buck, vice president and general manager of accelerated computing at NVIDIA. "HGX-2 provides the power to handle these massive new models for faster training of advanced AI, while saving significant cost, space and energy in the datacenter."

For comprehensive information on Supermicro NVIDIA GPU system product lines, 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. (NASDAQ: SMCI)
Supermicro® (NASDAQ: SMCI), 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.

Supermicro, Server Building Block Solutions, and We Keep IT Green are trademarks and/or registered trademarks of Super Micro Computer, Inc.

All other brands, names and trademarks are the property of their respective owners.

SMCI-F

C View original content with multimedia:http://www.prnewswire.com/news-releases/supermicro-unveils-2-petaflops-superserver-based-on-new-nvidia-hgx-2-the-worlds-most-powerful-cloud-server-platform-for-ai-and-hpc-300655316.html

SOURCE Super Micro Computer, Inc.

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