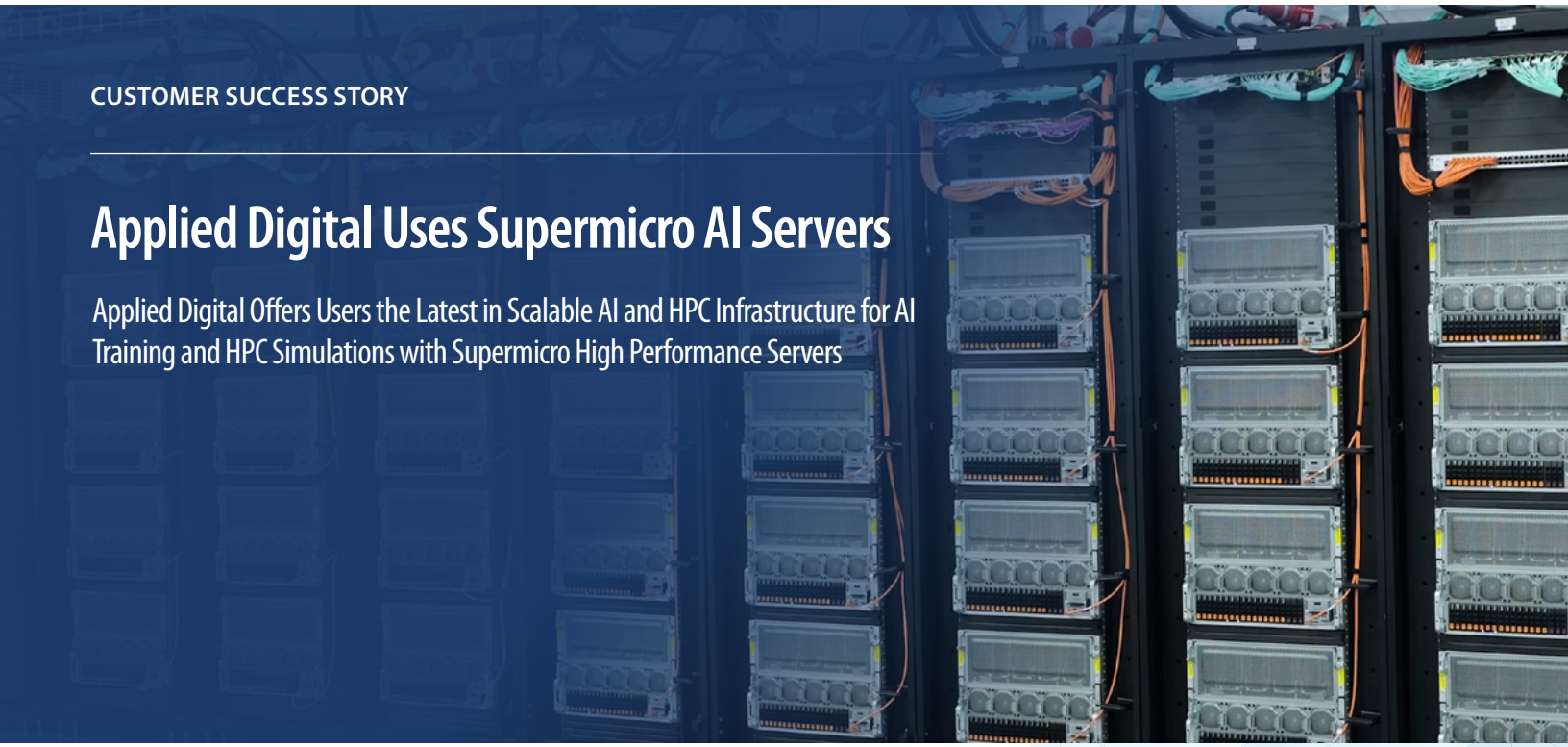
CUSTOMER SUCCESS STORY

Applied Digital Uses Supermicro AI Servers

Applied Digital Offers Users the Latest in Scalable AI and HPC Infrastructure for AI Training and HPC Simulations with Supermicro High Performance Servers



INTRODUCTION

Applied Digital Corporation (Nasdaq: APLD) is a designer, builder, and operator of next-generation digital infrastructure designed for Artificial Intelligence (AI) and High-Performance Computing (HPC) applications, cloud services, and data center hosting. With many data centers located across North America, Applied Digital is a leading cloud service provider that works closely with its customers to offer the latest in server technology with the latest GPUs. Applied Digital's Cloud Services, Applied Digital Cloud, provide bare metal and supercomputing-as-a-service for workloads requiring no-compromise performance at a large scale.

In many cases, Applied Digital locates its data centers close to renewable energy, which reduces the carbon footprint associated with high-end computing. In addition, Applied Digital is committed to operational efficiency to keep costs low, using renewable energy when available, and providing white glove service when working with customers to solve their unique needs.

INDUSTRY

AI Cloud Service Provider

Data Center Colocation Services

CHALLENGES

The Applied Digital HPC team has deep experience in the architecture, deployment, and optimization of parallel computing clusters for the latest server technology for a range of applications. With expertise in offering hardware and services for the most demanding compute intensive applications, Applied Digital was in a unique position to offer the AI/ML market the additional computing capacity that is needed, which continues to be scarce.

Applied Digital needed to evaluate which servers were available that would allow their customers to train their AI models and run HPC applications that were available in quantity in short timeframes. The requirements for the new servers were to have the most performant GPUs with fast interconnects between the GPUs. This architecture allows for quick training of AI models without having the CPU involved at every step. These servers are subsequently scaled over a 3.2Tbps NDR GPU-Direct RDMA fabric, allowing thousands of GPUs to operate in parallel.

To give customers an environment where they can execute their AI and HPC applications, Applied Digital needed to partner with a systems vendor to acquire the latest technology that could supply the best AI servers on the market.

SOLUTION

Applied Digital Cloud decided the most optimal system for their users would be the Supermicro SYS-821GE-TNHR with dual 4th Gen Intel® Xeon® Platinum processors (8462Y+). These servers are designed for NVIDIA HGX™ H100 8-GPU that are built for high-performance computing and AI training that require substantial memory bandwidth and computational power. The NVIDIA H100 Tensor Core GPU delivers 67 TFLOPS of FP32 performance, as well as 4 PFLOPS of Tensor Core optimized FP8 performance per GPU for large-scale AI inference. The system hosts eight H100 Tensor Core GPUs with 900GB/s NVLink GPU-GPU interconnect on the NVIDIA HGX baseboard. Applied Digital utilizes 100 Gigabit Ethernet for in-band management and object storage, and 400G NDR InfiniBand with 8x NVIDIA ConnectX-7 NICs per system for GPU fabric, GPUDirect support, and converged flash file system traffic. Inside the CPU tray, Applied Digital opted for 2 TB of system RAM to stage large datasets to ensure smooth data flow to the GPUs and prevent bottlenecks. Leveraging NVIDIA's reference architecture and networking topology, Applied Digital Cloud scales to thousands of H100 GPUs working in a single parallel compute cluster.

System	Processor	GPUs	Memory
 **SYS-821GE-TNHR**	Dual 4th Gen Intel Xeon Platinum (8462Y+)	8× NVIDIA HGX H100 (80 GB each)	2 TB
 **AS -4124GS-TNR**	Dual AMD EPYC (7513)	8× NVIDIA A100 (80 GB each)	1 TB
 **SYS-111E-WR**	Single 4th Gen Intel Xeon Scalable (5415+)	-	64 GB
 **AS -2014TP-HTR**	Single AMD EPYC (7534P)	-	1 TB

CHALLENGES

- Adding Large Scale AI Training to Offerings
- Increased AI and ML Abilities
- Enterprise Class Servers and AI Hardware
- Exceeding performance, value, and service over traditional cloud HPC offerings

SOLUTION

Additional Supermicro systems include:

- AS -2014TP-HTR (cluster infrastructure)
- SYS-111E-WR for UFM (cluster infrastructure)
- AS -4124GS-TNR with PCIe GPU (non-clustered general-purpose GPU cloud)

Supermicro switches acquired:

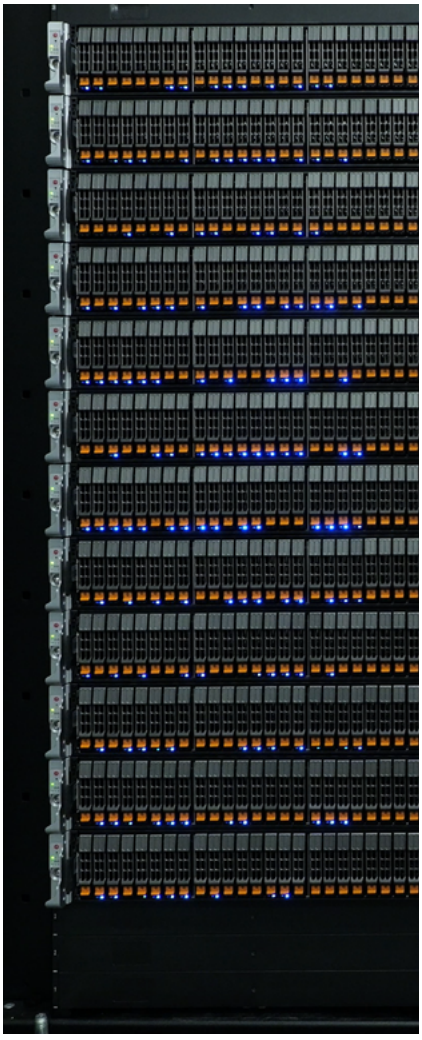
- SSE-SN3700-CS2FC (32× port 100 GbE switch)
- SSE-F3548SR (48× port 10/25 GbE switch)

Applied Digital takes pride in providing some of the most advanced AI clusters to the market for AI and HPC workloads. We work closely with Supermicro to deliver a range of systems to give our customers the most advanced and performant solutions to speed up their demanding workloads."

– **Don Allen**
Sales, *Applied Digital*

BENEFITS

Applied Digital immediately realized the benefits of using the Supermicro 8 GPU servers. With the AI training market exploding, Applied Digital was able to offer new services for AI training and HPC simulations. By giving customers access to the most advanced and powerful computing capability available today, HPC and AI applications run faster, and more complex AI pipelines can be used. Besides offering hardware-based solutions, Applied Digital experts can work with customers to determine the software that is most optimized for their particular workloads.



SUPER MICRO COMPUTER, INC.

Supermicro is a global leader in high performance, green computing server technology and innovation. We provide our global customers with application-optimized servers and workstations customized with blade, storage, and GPU solutions. Our products offer proven reliability, superior design, and one of the industry's broadest array of product configurations, to fit all computational need.

www.supermicro.com

APPLIED DIGITAL

Applied Digital (Nasdaq: APLD) designs, develops, and operates next-generation data centers across North America to provide digital infrastructure solutions to the rapidly growing high-performance computing (HPC) industry.

https://www.applieddigital.com

Printed in USA ♻ Please Recycle 50_2025-AppliedDigital_R04