



THE BIOS ANNA HGX GPU SERVER

Optimised for Deep Learning workloads

Introducing, the BIOS ANNA HGX - a new high density 4U GPU server which features an NVIDIA HGX A100 8 GPU baseboard. Designed for the most demanding AI workloads, this high performing, fast-to-market server is based on NVIDIA A100™ Tensor Core GPUs. These servers can deliver up to 5 PetaFLOPS of AI performance in a single 4U system, with the help of the newest version of NVIDIA® NVLink™ and NVIDIA NVSwitch technologies.

DISAGGREGATED HARDWARE

Ideal for large-scale deep learning training and neural network model applications. The new 4U GPU system features the NVIDIA HGX A100 8-GPU baseboard, up to six NVMe U.2 and two NVMe M.2, 10 PCI-E 4.0x16 I/O, with Supermicro's unique AIOM support invigorating the 8-GPU communication and data flow between systems through the latest technology stacks such as NVIDIA NVLink and NVSwitch, GPUDirect RDMA, GPUDirect Storage, and NVMe-oF on InfiniBand.

KEY FEATURES

Up to 8 NVIDIA A100 GPUs

PCI-E Gen 4

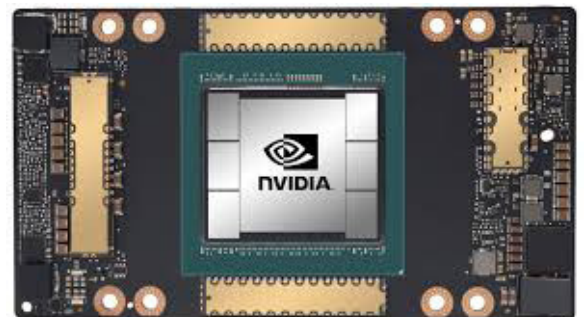
2 Processors

32 DIMM slots

4x 2000W Redundant Titanium Level PSU

FEATURING LATEST GEN GPUS

The NVIDIA A100 GPU is architected to not only accelerate large complex workloads, but also efficiently accelerate many smaller workloads.



A100 powers the NVIDIA data centre platform that includes Mellanox HDR InfiniBand (IB), NVSwitch, HGX A100 and the Magnum IO SDK for scaling up.



BUILT ON SUPERMICRO

For the most demanding AI workloads, Supermicro builds the highest-performance, fastest-to-market servers based on NVIDIA A100™ Tensor Core GPUs.

Supermicro supports a range of customer needs with optimized systems for the new HGX™ A100 8-GPU and HGX™ A100 4-GPU platforms. With the newest version of NVIDIA® NVLink™ and NVIDIA NVSwitch™ technologies, these servers can deliver up to 5 PetaFLOPS of AI performance in a single 4U system. Supermicro can also support the new NVIDIA A100 PCI-E GPUs in a range of systems, with up to 8 GPUs in a 4U server.

BIOS ANNA HGX	
COMPUTE	Dual AMD EPYC™ 7002 Series Processors
GPU	Supports up to 8 double-width GPUs Direct connect for maximum performance
MEMORY	8TB Registered ECC DDR4 3200MHz SDRAM in 32 DIMMs
STORAGE CAPACITY	Up to 24 Hot-swap 2.5" drive bays; 4x 2.5" SATA drives and 4x 2.5" NVMe drives
EXPANSION SLOTS	9 PCI-E 4.0 x16 (or 10 PCI-E 4.0 x16 slots without NVMe devices)
NETWORKING	2 GbE LAN ports (rear)
POWER SUPPLY	2000W (2+2) Redundant Power Supplies Titanium Level (96%+)

ABOUT BIOS IT

BIOS IT is a global design house, systems builder and integrated solution provider for enterprise performance computing. We construct bespoke clusters and appliances by hand picking the best components and newest technologies based on specific customer goals. We support well-known organisations at the top of their research fields, across science, engineering, academia and finance disciplines. Our comprehensive range of products and services include: high-performance enterprise servers, storage and networking, on-premise or in the cloud, with associated services, support, hosting and software.

Custom configurations are also available. For more information contact BIOS IT..



AMERICAS: 1-800-654-BIOS | EMEA: +44 (0) 203 178 6467 | APAC: +61 (0)2 8866 3343

www.bios-it.com

All products and companies referred to herein are trademarks or registered trademarks of their respective companies or mark holders.