Air & Liquid-Cooled B300 Systems
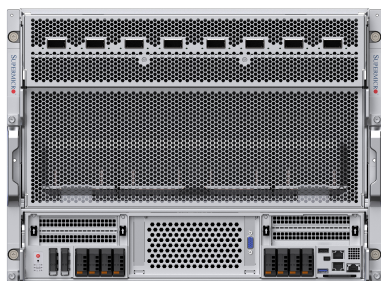
Supermicro NVIDIA HGX™ B300 8-GPU Systems, 8U Air-Cooled and 4U Liquid-Cooled

Air Cooling

Liquid Cooling

The Proven Platform for Large-Scale AI Workloads, Upgraded with Ultra Performance

The Supermicro NVIDIA HGX platform is the building block of the world's largest AI training clusters, delivering the immense computational output required for delivering the transformative AI applications of today. Now featuring NVIDIA Blackwell Ultra, Supermicro's NVIDIA HGX B300 solutions enable rapid deployment of the industry's highest-performing AI infrastructure to tackle the largest-scale AI training, real-time AI reasoning, agentic AI applications, multimodal AI inference, and physical AI applications.

Supermicro integrates Blackwell Ultra generational advancements with system-level and rack-level engineering to create a solution that maximizes efficiency and performance at scale. Both 8U air-cooled and 4U liquid-cooled configurations are available, providing up to 7.5x performance gains over NVIDIA Hopper systems, along with 288GB of HBM3e memory per GPU* to tackle massive AI model sizes.

Air-Cooled 8U
NVIDIA HGX B300 8-GPU System

Coming Soon

Liquid-Cooled 4U
NVIDIA HGX B300 8-GPU System

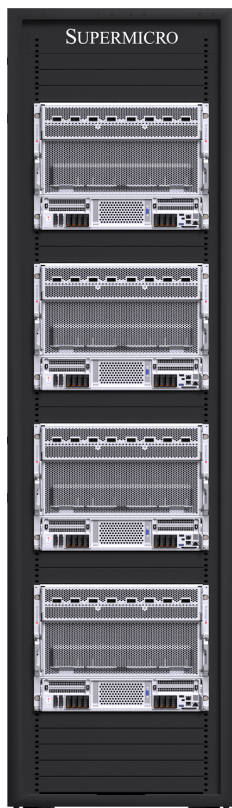| System	SYS-822GS-NB3RT / AS -8126GS-NB3RT
Overview	8U air-cooled system with front I/O NICs, DPUs, storage, and management
CPU	Dual Intel® Xeon® 6700 series processors with P-cores (SYS-822GS-NB3RT) Dual AMD EPYC™ 9005/9004 Series Processors (AS -8126GS-NB3RT)
GPU	NVIDIA HGX B300 8-GPU (288GB HBM3e per GPU*) 1.8TB/s NVLink GPU-GPU interconnect with NVSwitch
Memory	32 DIMMs, up to 8TB DDR5-5200 or up to 4TB DDR5-6400 (SYS-822GS-NB3RT) 24 DIMMs, up to 6TB DDR5-6400 (AS -8126GS-NB3RT)
Storage	8 Hot-swap E1.S NVMe storage drive bays and 2 M.2 NVMe boot drives
Networking	8 integrated NVIDIA ConnectX®-8 SuperNICs, up to 800Gb/s 2 dual-port NVIDIA BlueField®-3 DPUs
Power	6x 6600W Redundant (3 + 3) Titanium Level (96%) power supplies

System	SYS-422GS-NB3RT-LCC/ALC
Overview	4U liquid-cooled system with front I/O NICs, DPUs, storage, and management
CPU	Dual Intel® Xeon® 6700 series processors with P-cores
GPU	NVIDIA HGX B300 8-GPU (288GB HBM3e per GPU*) 1.8TB/s NVLink GPU-GPU interconnect with NVSwitch
Memory	32 DIMMs, up to 8TB DDR5-5200 or up to 4TB DDR5-6400
Storage	8 Hot-swap E1.S NVMe storage drive bays and 2 M.2 NVMe boot drives
Networking	8 integrated NVIDIA ConnectX®-8 SuperNICs, up to 800Gb/s 2 dual-port NVIDIA BlueField®-3 DPUs
Power	6x 6600W Redundant (3 + 3) Titanium Level (96%) power supplies

*Physical GPU memory

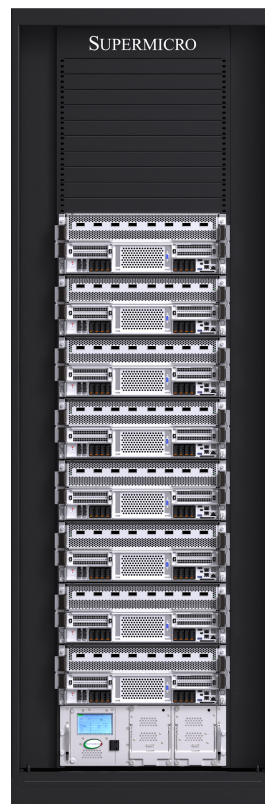**SUPERMICRO**

Air-Cooled Rack

Networking
- NVIDIA Quantum-X800 InfiniBand or NVIDIA Spectrum-X Ethernet for up to 800Gb/s compute fabric
- Ethernet leaf switches for in-band management
- Out-of-band 1G/10G IPMI switch
- Non-blocking network

Compute
- 4x SYS-822GS-NB3RT or AS -8126GS-NB3RT
- 32x NVIDIA B300 GPUs per rack
- 9.2TB of HBM3e per rack
- Flexible storage options with local or dedicated storage fabric with full NVIDIA GPUDirect RDMA and Storage or RoCE support

Liquid-Cooled Rack

Networking
- NVIDIA Quantum-X800 InfiniBand or NVIDIA Spectrum-X Ethernet for up to 800Gb/s compute fabric
- Ethernet leaf switches for in-band management
- Out-of-band 1G/10G IPMI switch
- Non-blocking network

Compute
- 8x SYS-422GS-NB3RT-LCC/ALC
- 64x NVIDIA B300 GPUs per rack
- 18.4TB of HBM3e per rack
- Flexible storage options with local or dedicated storage fabric with full NVIDIA GPUDirect RDMA and Storage or RoCE support

Liquid-Cooling
- Supermicro 250kW capacity Cooling Distribution Unit (CDU) with redundant PSU and dual hot-swap pumps
- Vertical Cooling Distribution Manifold (CDM)

Air Cooling

Liquid Cooling

Coming Soon

72-Node Air-Cooled Cluster

Overview	Fully integrated air-cooled 72-node cluster with up to 576 NVIDIA B300 GPUs
Compute Fabric Leaf	Up to 8x NVIDIA Quantum-X800 InfiniBand Switches or 8x NVIDIA Spectrum-4 Ethernet Switches, up to 800Gb/s
Compute Fabric Spine	Up to 4x NVIDIA Quantum-X800 InfiniBand Switches or 4x NVIDIA Spectrum-4 Ethernet Switches, up to 800Gb/s
In-band Management Switch	3x NVIDIA Spectrum SN4600 100GbE Ethernet Switch
Out-of-band Management Switch	2x SSE-G3748R-SMIS, 48-port 1Gbps Ethernet ToR management switch 1x SSE-F3548SR, 48-port 10Gbps Ethernet ToR management switch
Rack	20x 48U x 750mm x 1400mm
PDU	38x 415V/60A 3Ph

*Recommended configuration. Other network switch options and rack dimensions and layouts are available.

72-Node Liquid-Cooled Cluster

Overview	Fully integrated liquid-cooled 72-node cluster with up to 576 NVIDIA B300 GPUs
Compute Fabric Leaf	Up to 8x NVIDIA Quantum-X800 InfiniBand Switches or 8x NVIDIA Spectrum-4 Ethernet Switches, up to 800Gb/s
Compute Fabric Spine	4x NVIDIA Quantum-X800 InfiniBand Switches or 4x NVIDIA Spectrum-4 Ethernet Switches, up to 800Gb/s
In-band Management Switch	3x NVIDIA Spectrum SN4600 100GbE Ethernet Switch
Out-of-band Management Switch	2x SSE-G3748R-SMIS, 48-port 1Gbps Ethernet ToR management switch 1x SSE-F3548SR, 48-port 10Gbps Ethernet ToR management switch
Rack and PDU	11x 48U x 800mm x 1470mm PDU: 56x 415V 60A 3Ph
Liquid Cooling	9x Supermicro 250kW capacity CDU with redundant PSU and dual hot-swap pumps

SUPERMICRO