Liquid-Cooled 2-OU B300 System

NVIDIA HGX™ B300 System Optimized for OCP ORV3 Design with up to 144 GPUs in a Rack

Hyperscale AI Platform in Ultra-Compact Design

Built to the 21-inch OCP Open Rack V3 (ORV3) specification, Supermicro's 2-OU liquid-cooled NVIDIA HGX B300 system sets a new standard for GPU density and efficiency. Each compact node features eight NVIDIA Blackwell B300 GPUs operating at up to 1,100W TDP, cooled by state-of-the-art liquid cooling with blind-mate manifold connections and modular GPU/CPU tray architecture. This innovative design delivers exceptional serviceability while dramatically reducing power consumption and rack footprint—enabling hyperscale and cloud providers to maximize compute performance in space-constrained data centers. Data Center Building Block Solutions®, combined with Supermicro's expertise in on-site deployments, provide a total solution encompassing liquid-cooling technology, network topology and cabling, power delivery, and thermal management to accelerate time-to-online.

144 GPUs per Rack: Industry-Leading Density at Scale

Scale seamlessly from node to rack to cluster with Supermicro's complete NVIDIA HGX B300 rack-level architecture. A single OCP ORV3 rack supports up to 18 compact 2-OU nodes with 144 GPUs total, delivering the industry's highest GPU density while maintaining full serviceability through modular design. Each rack integrates with NVIDIA Quantum-X800 InfiniBand switches and Supermicro's 1.8MW in-row coolant distribution units (CDUs), creating a complete liquid-cooled infrastructure optimized for maximum performance and energy efficiency. At full scale, eight HGX B300 compute racks, three NVIDIA Quantum-X800 networking racks, and two Supermicro CDUs form a SuperCluster with 1,152 GPUs—delivering unmatched performance density for the most demanding AI workloads in hyperscale data centers and AI factories.

Liquid-Cooled 2-OU
NVIDIA HGX B300 8-GPU System

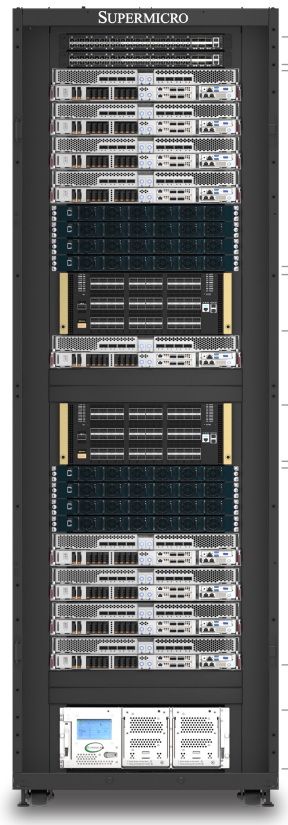
System	SYS-222GS-NB30T-ALC
Overview	2-OU Liquid-cooled system with front I/O NICs, DPUs, storage, and management
CPU	Dual Intel® Xeon® 6700 series processors with P-cores
Memory	32 DIMMs, up to 8TB DDR5-5200 or up to 4TB DDR5-6400
GPU	NVIDIA HGX B300 8-GPU (288GB HBM3e per GPU*) 1.8TB/s NVLink GPU-GPU interconnect with NVSwitch
NVLink	5th Generation NVIDIA NVLink at 1.8TB/s
Networking	8 integrated NVIDIA ConnectX®-8 SuperNICs, up to 800Gb/s 2 dual-port NVIDIA BlueField®-3 DPUs
Storage	8 front hot-swap E1.S NVMe drive bays 2 M.2 NVMe slots
Networking	8 integrated NVIDIA ConnectX®-8 SuperNICs, up to 800Gb/s 2 dual-port NVIDIA BlueField®-3 DPUs
Power Supply	Shared power through 4+4 rack power shelves with ORV3 busbar design

*Physical GPU memory

NVIDIA Blackwell Architecture Solutions

DATASHEET

72-GPU per Rack



Management Networking

- In-band and out-of-band management switches

Compute Nodes and Power Shelves

- 5x SYS-222GS-NB3OT-ALC
- 4x power shelves with OCP ORV3 44-OU rack busbar
- 72x NVIDIA Blackwell Ultra B300 GPUs per rack (total of 9 compute nodes)

Compute Fabric Networking

- 2x NVIDIA Quantum-X800 InfiniBand switches for up to 800Gb/s compute fabric (leaf and spine switch)
- Dedicated storage fabric options with full NVIDIA GPUDirect RDMA and Storage or RoCE support

Compute Nodes and Power Shelves

- 4x SYS-222GS-NB3OT-ALC
- 4x power shelves with OCP ORV3 44-OU rack busbar

Liquid-Cooling

- Supermicro 250kW capacity Coolant Distribution Unit (CDU) with redundant PSU and dual hot-swap pumps

144-GPU per Rack

Management Networking

- In-band and out-of-band management switches

Compute Nodes and Power Shelves

- 18x SYS-222GS-NB3OT-ALC
- 10x power shelves with OCP ORV3 48-OU rack busbar
- 144x NVIDIA Blackwell Ultra B300 GPUs per rack

Compute Fabric Networking

- Centralized networkng racks with NVIDIA Quantum-X800 InfiniBand switches for up to 800Gb/s compute fabric (leaf and spine switches)
- Dedicated storage fabric options with full NVIDIA GPUDirect RDMA and Storage or RoCE support

Liquid-Cooling

- Supermicro 1.8MW capacity Coolant Distribution Unit (CDU) with redundant PSU and n+1 hot-swap pumps

288-GPU Scalable Unit

1152-GPU Scalable Unit

36-Node SuperCluster

Overview	Fully integrated Liquid-cooled 36-node cluster with up to 288 NVIDIA B300 GPUs
Compute Fabric Leaf	4x NVIDIA Quantum-X800 InfiniBand switches, up to 800Gb/s
Compute Fabric Spine	2x NVIDIA Quantum-X800 InfiniBand switches, up to 800Gb/s
In-band Management Switch	1x NVIDIA Spectrum SN5610 Ethernet switch
Out-of-band Management Switch	2x NVIDIA Spectrum SN2201 Ethernet switches
Rack Dimension	44-OU x 750mm x 1200mm
Power Shelves with busbar	32x 1U 33kW (6x 5.5kW PSUs) power shelves with built-in capacitor
Liquid Cooling	4x in-rack Supermicro 4U 250kW capacity CDU with redundant PSU and n+1 hot-swap pumps Optional: 1x 1.8MW capacity in-row CDU

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

144-Node SuperCluster

Overview	Fully integrated liquid-cooled 144-node cluster with up to 1152 NVIDIA B300 GPUs
Compute Fabric Leaf	16x NVIDIA Quantum-X800 InfiniBand switches, up to 800Gb/s
Compute Fabric Spine	8x NVIDIA Quantum-X800 InfiniBand switches, up to 800Gb/s
In-band Management Switch	3x NVIDIA Spectrum SN5610 Ethernet switch
Out-of-band Management Switch	9x NVIDIA Spectrum SN2201 Ethernet switches
Rack Dimension	48-OU x 750mm x 1200mm
Power Shelves with busbar	80x 1U 33kW (6x 5.5kW PSUs) power shelves with built-in capacitor
Liquid Cooling	2x Supermicro 1.8MW capacity CDUs with redundant PSU and n+1 hot-swap pumps

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