NVIDIA HGX B300

Also known as: HGX B300 8-GPU baseboard · B300 baseboard · 935-26287-0070-000 · Blackwell Ultra platform

HGX B300 is the 8-GPU Blackwell Ultra baseboard NVIDIA sells to OEMs, the platform under every non-DGX B300 server. It carries 2.3 TB of HBM3e across eight 288 GB GPUs on a 14.4 TB/s NVLink 5 fabric, and unlike HGX B200 it integrates the scale-out network on the platform itself: eight ConnectX-8 SuperNICs at 800 Gb/s per GPU over PCIe Gen6, so the NIC choice OEMs used to make is now fixed by NVIDIA. Supermicro, Dell, HPE, Lenovo, Aivres and others wrap their own CPUs, chassis and cooling around it; the same baseboard ships in air-cooled 8U boxes at the 1,100 W GPU bin and in 2-OU/4U direct-liquid chassis, with the full 1,400 W bin reserved for liquid builds. It is the default OEM quote for new Blackwell capacity in 2026.

8x B300 SXM GPUs · Blackwell Ultra, fully NVLinked	**2.3 TB HBM3e total** GPU memory · 288 GB per GPU	**up to 64 TB/s aggregate** Memory bandwidth · 8 TB/s per GPU
14.4 TB/s aggregate NVLink · NVLink 5, 1.8 TB/s per GPU	**8x ConnectX-8 on-platform** Networking · 800 Gb/s per GPU, PCIe Gen6	**1,100 W air / up to 1,400 W liquid** GPU TDP · per GPU, bin set by OEM chassis

COMPUTE	01/05
GPU	NVIDIA B300 (Blackwell Ultra) · SXM form factor, 8 per baseboard
FP4 Tensor	up to 144 PFLOPS · 8-GPU platform, NVIDIA figure; dense is 120 PFLOPS
FP8 Tensor	72 PFLOPS · 8-GPU platform, NVIDIA figure
Attention throughput	~2x HGX B200 · doubled SFU throughput
GPU TDP	1,100 W (air) / up to 1,400 W (liquid) · configurable per GPU

MEMORY	02/05
Memory type	HBM3e, 12-Hi stacks · per GPU
Memory per GPU	288 GB · 50% more than HGX B200's 180 GB
Total baseboard memory	2.3 TB · 8x 288 GB
Bandwidth per GPU	up to 8 TB/s
Aggregate bandwidth	up to 64 TB/s · per node

INTERCONNECT	03/05
NVLink generation	NVLink 5 · same generation as HGX B200
GPU-to-GPU bandwidth	1.8 TB/s per GPU · bidirectional
Aggregate NVLink bandwidth	14.4 TB/s · across all 8 GPUs
Switch fabric	NVSwitch · full any-to-any GPU mesh
Scale-out NICs	8x ConnectX-8 SuperNICs, on-platform · 800 Gb/s per GPU (2x 400G), PCIe Gen6 direct to GPU
vs. HGX B200	NICs integrated, not OEM-selected · B200 platforms left NIC choice to the OEM

POWER & COOLING	04/05
GPU-only draw (air)	8.8 kW · 8x 1,100 W bin
GPU-only draw (liquid)	up to 11.2 kW · 8x 1,400 W bin
Full server draw	12 to 16 kW · adds CPUs, DRAM, NVMe, NICs, fans; DGX B300 reference is 14.5 kW
Cooling options	Air (8U class) or direct-to-chip liquid (2-OU/4U class) · liquid unlocks the top TDP bin and rack density

OEM PLATFORMS

05/05

Ships from	Supermicro, Dell, HPE, Lenovo, Aivres, and other NVIDIA-certified OEMs · each builds its own chassis around the baseboard
Example SKUs	Supermicro SYS-822GS-NBR (8U air), SYS-422GS-NB3RT-ALC (4U liquid), SYS-222GS-NB3OT-ALC (2-OU liquid)
CPU choice	AMD EPYC 9005/9004 or Intel Xeon · OEM- and SKU-dependent
Baseboard part number	935-26287-0070-000 · sold as a discrete part through distribution
vs. DGX B300	Same platform, OEM picks chassis/CPU/cooling · DGX is NVIDIA's fixed-config, air-cooled-only equivalent

PROCUREMENT

Typical lead time

Reported 4 to 8 weeks through OEMs with established NVIDIA allocation as of mid-2026 (Supermicro cited fastest at 2 to 4 weeks, Dell and Lenovo 6 to 8 in constrained periods), and 8 to 14 weeks more typical for buyers without an existing relationship. The 288 GB memory tier is the most allocation-constrained Blackwell part.

Pricing

$430,000 to $550,000+ for a fully configured 8-GPU HGX B300 server, reported. Aivres quotes its HGX B300 platform starting near $430,000; liquid-cooled and heavier-networked builds run higher. Individual B300 modules have been reported around $53,000. NVIDIA's own DGX B300 (same platform, fixed config) reports at $300,000 to $500,000 depending on source and configuration, which for the first time makes DGX price-competitive with OEM builds.

- Unlike HGX B200, reported OEM pricing does not consistently undercut DGX: DGX B300's MGX-standard build reportedly lowered NVIDIA's own cost. Quote both paths before assuming OEM is cheaper.
- The eight ConnectX-8 NICs are on the platform, so the scale-out fabric decision (Quantum-X800 InfiniBand vs Spectrum-X800 Ethernet) and its switch/optics budget is locked in the moment you pick B300.
- Air-cooled 8U SKUs run the 1,100 W GPU bin; the 1,400 W bin requires direct-liquid chassis. Confirm which bin an OEM quote assumes, since it moves both performance and facility requirements.
- New OEM allocation is skewing to B300 over B200; expect B300 to be the default quote even when 180 GB per GPU would suffice.

FIELD NOTES

- HGX B300 is a baseboard, not a server, but it fixes more of the design than HGX B200 did: GPUs, NVSwitch fabric and the eight ConnectX-8 SuperNICs are all NVIDIA's. The OEM now differentiates only on CPU, memory, storage, chassis and cooling.
- Air-cooled B300 exists (1,100 W bin, 8U class) despite widespread claims that Blackwell Ultra is liquid-only. What is true: the full 1,400 W bin needs direct liquid, and four 8U air boxes per rack already demand ~60 kW rack power, so air works only in high-density halls.
- The 800 Gb/s per-GPU scale-out only pays off on an 800G-class fabric (Quantum-X800 or Spectrum-X SN5600-class). Budget switches and optics with the node order; at 8x 800 Gb/s per node the fabric BOM rivals mid-five-figures per node.
- Per-GPU HBM capacity, not FLOPS, is the usual reason to specify B300 over B200: 288 GB holds model shards and KV cache that would force tensor-parallel spill on 180 GB parts, and that single change often removes a whole parallelism dimension from the deployment.
- Two 'HGX B300' servers from different OEMs still differ meaningfully in CPU platform, DRAM ceiling, storage layout and serviceability. Compare full system BOMs, not the shared baseboard spec sheet.

QUESTIONS WE GET ON THIS PART

How much does an HGX B300 server cost?

Reported street pricing for a fully configured 8-GPU HGX B300 server runs roughly $430,000 to $550,000 or more, with Aivres quoting from about $430,000. This excludes the 800G switch fabric, optics, and rack power and cooling infrastructure.

What is the difference between HGX B300 and HGX B200?

B300 carries 288 GB per GPU (2.3 TB per baseboard) versus B200's 180 GB (1.44 TB), raises the GPU power ceiling to 1,400 W liquid-cooled, adds roughly 1.5x the dense FP4 throughput, and integrates eight ConnectX-8 SuperNICs at 800 Gb/s per GPU on the platform, where HGX B200 left NIC selection to the OEM.

Is there an air-cooled HGX B300 server?

Yes. Supermicro ships an 8U air-cooled HGX B300 system (SYS-822GS-NBR) running the GPUs at the 1,100 W TDP bin, alongside 2-OU and 4U direct-liquid systems. The full 1,400 W GPU bin is only available in liquid-cooled chassis.

What is the difference between HGX B300 and DGX B300?

Same 8-GPU Blackwell Ultra platform. HGX B300 is the baseboard OEMs build their own servers around, with choice of CPU, chassis and air or liquid cooling. DGX B300 is NVIDIA's fixed-configuration, air-cooled 10U appliance on the same platform with NVIDIA-direct support.

What is the HGX B300 baseboard part number?

The 8-GPU HGX B300 baseboard circulates in distribution as NVIDIA part number 935-26287-0070-000, though in practice nearly all volume ships inside complete OEM servers rather than as bare baseboards.

REFERENCES

[1] https://docs.nvidia.com/enterprise-reference-architectures/hgx-ai-factory/latest/components.html

[2] https://www.supermicro.com/en/accelerators/nvidia

[3] https://smicro.eu/nvidia-baseboard-8x-b300-935-26287-0070-000-1

[4] https://www.arccompute.io/arc-blog/aivres-nvidia-hgx-b200-b300-gpu-servers-air-and-liquid-cooled-performance-at-scale

[5] https://slyd.com/resources/server-comparison

[6] https://acecloud.ai/blog/nvidia-hgx-b300/

[7] https://www.sabrepc.com/blog/computer-hardware/nvidia-hgx-b300

[8] https://tech-insider.org/nvidia-blackwell-gpu-pricing/

Latest version of this datasheet: https://gpusmith.com/hardware/systems/nvidia-hgx-b300. Specifications compiled from the references above; verify against vendor documentation before purchase.