

Supermicro Super AI Station

Introducing the new class of development platform purpose-built to power the age of AI



Featuring the NVIDIA GB300 Grace™ Blackwell Ultra Desktop Superchip

- Up to 20 PFLOPS of NVFP4 performance
- Closed-loop liquid cooled desktop tower for a quiet and shared working space friendly design
- Optional NVIDIA RTX PRO™ 4000 Blackwell or RTX PRO 2000 Blackwell GPU for graphics rendering
- 748GB total coherent memory to excel at AI inferencing, model training, and fine-tuning
- Equipped to handle inference of models with up to 1 trillion parameters
- 1600W Titanium (94%) level power supply for stable, efficient, and reliable performance

The Ultimate AI Development Platform

The Supermicro Super AI Station (ARS-511GD-NB-LCC) brings data center level performance in a space-efficient desktop tower form factor, bridging a crucial gap between consumer-grade PCIe-based GPU workstations and data-center-grade AI servers. Powered by NVIDIA's high-end GB300 Grace Blackwell Ultra Desktop Superchip, this system comes pre-installed with the NVIDIA AI Enterprise (NVAIE) software stack and supports NVIDIA's extensive catalog of CUDA accelerated data science libraries and tools. These combined capabilities make the Super AI Station the ideal platform for AI researchers, application developers, engineers, and data scientists.

Securely Build, Train, and Evaluate Models

The Super AI Station is well equipped to handle inference of models up to 1 trillion parameters in size, and the training or fine-tuning of domain-specific models, making it ideal for quickly iterating, building, and evaluating prototypes for new and innovative AI applications. This is a powerful on-prem platform that keeps proprietary data and intellectual property on-site without any cloud egress, satisfying compliance rules.

Bring the Data Center to Your Desk

Purpose built for research labs, AI & deep-tech startups, and enterprise teams, the Super AI Station utilizes conventional office power outlets and is equipped with a 1600W Titanium level (94%) power supply. With its closed-loop liquid cooling system, the Super AI Station is both quiet and shared workspace friendly, which combined with space-efficient 5U desktop tower design makes this system a true AI supercomputer for the office.

Develop Autonomous AI Agents In House

With the rise of agentic AI and long-running autonomous agents, developers need powerful, secure on-prem solutions that handle compute-heavy systems and support the software frameworks required to build them. Combined with the Super AI Station, NVIDIA's NemoClaw enables users to run powerful autonomous agents safely 24/7 with frontier-class, trillion-parameter models without sacrificing control. With massive memory and data center-class performance at your desk, the Super AI Station delivers a secure, scalable, plug-and-play compute platform for building autonomous agents locally – free from data privacy and security risks.

Unprecedented Compute Performance

The NVIDIA GB300 Grace Blackwell Ultra Desktop Superchip combines an NVIDIA B300 GPU and NVIDIA Grace CPU; the same technologies that power the NVIDIA GB300 NVL72 rack-scale solution. The powerful GPU features 252GB of HBM3e GPU memory and delivers up to 20 PFLOPS of NVFP4 performance. The NVIDIA GB300 Grace Blackwell Ultra Desktop Superchip uses a coherent memory architecture which allows the CPU and GPU to share a single, unified 748GB memory pool

to support large AI training datasets. The CPU and GPU are interconnected via NVIDIA's NVLink™-Chip-2-Chip (NVLink™-C2C), which provides 900GB/s bi-directional bandwidth. It also features an integrated dual-port 400Gb/s NVIDIA ConnectX®-8 SuperNIC™ for extremely fast and efficient high-bandwidth, low-latency networking that is perfect for data-hungry AI workloads.



Supermicro Super AI Station ARS-511GD-NB-LCC	
Form Factor	5U Tower
Processor & GPU Support	NVIDIA GB300 Blackwell Ultra Desktop Superchip, featuring: • 1 NVIDIA Grace™ 72-core ARM Neoverse V2 CPU • 1 NVIDIA Blackwell Ultra (B300) GPU Optional: 1 NVIDIA Blackwell GPU for graphics rendering
Memory	Total Coherent Memory: 748GB • System: 496GB LPDDR5X • GPU: 252GB HBM3e
Storage	4 M.2 PCIe 5.0 NVMe
Networking	Dual-port QSFP 400GbE LAN via NVIDIA ConnectX®-8 SuperNIC™ 1GbE BMC LAN 10GbE LAN
Power Supply	1600W Titanium (94%) Level Power Supply Requires NEMA 5-20 outlet (in North America)
Cooling	Closed-loop liquid cooling
Operating System	Ubuntu 24.04 LTS with NVIDIA AI Developer Tools

Private Cloud & Multi-User Support

The Super AI Station can also serve as a centralized compute node for multiple team members to fine-tune and run IP-specific models on-demand. The NVIDIA GB300 Grace Blackwell Ultra Desktop Superchip supports Multi-Instance GPU (MIG)

technology to partition into as many as seven instances for local development with multiple users, each fully isolated with its own high-bandwidth memory, cache, and compute cores.