



NVIDIA DGX Station for Windows Puts a Trillion-Parameter AI Supercomputer on Every Enterprise Desk

News Summary:

- NVIDIA announces DGX Station for Windows -- the world's most powerful deskside AI supercomputer for developing and running agents on Windows -- built on the NVIDIA GB300 Grace Blackwell Ultra Desktop Superchip, coming in Q4 this year.
- DGX Station brings frontier AI agents to Windows -- enabling enterprise developers, researchers, engineers, designers and data scientists to build and deploy AI across the workflows and applications their business runs on.
- DGX Station will support NVIDIA OpenShell on Windows, built on new Windows security and containment primitives.

NVIDIA GTC Taipei--NVIDIA today announced [NVIDIA DGX Station™ for Windows](#), the world's most powerful deskside AI supercomputer designed to build, run and connect always-on AI agents to Windows applications and workflows, capable of running frontier AI models of up to 1 trillion parameters locally.

Historically, heavy-duty enterprise AI workloads -- training, fine-tuning, large-scale inference and multi-agent development -- have required powerful AI systems in the data center that run on Linux, while the vast majority of Fortune 500 companies use Windows for everyday productivity, creative, design and engineering applications.

Building on the [NVIDIA DGX Station](#) system design, DGX Station for Windows bridges this gap as the first deskside AI supercomputer to bring NVIDIA GB300 Grace Blackwell-class AI infrastructure directly into the Windows ecosystem -- providing the compute needed to build, run and connect powerful AI agents to the applications and infrastructure Windows users already harness.

"As enterprises scale AI agents across their organizations, they need AI infrastructure that can connect directly to the applications and workflows that power their business," said Chris Marriott, vice president of enterprise platforms at NVIDIA. "DGX Station delivers supercomputing-class AI directly into Windows, where millions already design, engineer, research and create every day."

"For decades, Microsoft and NVIDIA have partnered to advance the most powerful computing platforms in the world," said Pavan Davuluri, executive vice president of Windows + Devices at Microsoft. "Today, we're taking that collaboration to the next level, scaling the full power of Windows from thin-and-light PCs to data-center-class workstations with DGX Station powered by GB300. This unlocks a new class of AI performance on Windows, the platform enterprises trust for security, manageability and compatibility."

Introducing DGX Station for Windows

The DGX Station state-of-the-art design is powered by the NVIDIA GB300 Grace Blackwell Ultra Desktop Superchip, connecting a powerful NVIDIA Blackwell Ultra GPU to a high-performance 72-core NVIDIA Grace™ CPU via the NVIDIA NVLink™-C2C interconnect for best-in-class system communication and performance.

It features up to 748GB of coherent memory and up to 20 petaflops of FP4 performance, and can be paired with an NVIDIA RTX PRO™ 6000 Blackwell Workstation GPU for frontier AI compute with ray-traced visualization and simulation.

In addition, DGX Station for Windows features the NVIDIA ConnectX®-8 SuperNIC™, optimized to supercharge hyperscale AI computing workloads. With support for networking at up to 800Gb/s, the ConnectX-8 SuperNIC enables extremely fast network data transfers for AI workloads and high-speed connectivity of multiple DGX Station systems for even larger workloads.

A Window Into the World of Always-On AI Agents

Enterprise AI is evolving from simple chatbot interactions to agentic inference that operate continuously, reason in real time and connect directly to enterprise applications and workflows.

Developed in collaboration with Microsoft, DGX Station for Windows serves as dedicated agent infrastructure -- enabling enterprises to build and run agents with frontier intelligence locally, supporting AI models of up to 1 trillion parameters. DGX Station can also run agents at scale, with hundreds of agents executing on tasks simultaneously.

Powerful AI agents can be built and connected to 3D design and engineering applications, giving developers, designers and engineers intelligent assistants that understand their tools and processes to automate repetitive tasks and accelerate productivity.

For enterprise IT teams, DGX Station for Windows brings a secure, managed platform to GB300 deployments -- extending

the same Windows security, compliance and fleet management infrastructure organizations already rely on. AI agents deploy and operate within this managed environment, governed through familiar Microsoft tools.

Linux workloads get the same level of manageability support through Windows Subsystem for Linux. Enterprise-grade features -- including deployment and system updates -- help organizations maintain security, compliance and operational readiness across their fleets.

NVIDIA and Microsoft's collaboration to deliver agents in the Windows experience extends from frontier to personal agents with [NVIDIA RTX Spark™](#) -- bringing the full spectrum of NVIDIA AI to slim laptops and small desktop PCs.

Secure Agent Development and Deployment With NVIDIA OpenShell

Autonomous agents need to be developed and deployed in a secure runtime that governs how they act, how they operate tools and how they interact with the rest of the system. DGX Station is an ideal platform for Windows users to build and run always-on, autonomous agents locally and securely before scaling to data center AI factories.

[NVIDIA OpenShell™](#) is an open source, secure-by-design runtime for autonomous agents. Building on the new Windows security and containment primitives, it creates an individual, isolated sandbox for each agent and separates application-layer operations from infrastructure-layer policy enforcement.

This means security and privacy policies are out of the agent's reach and applied at the system level. Instead of relying on behavioral system prompts, OpenShell uses the new Windows security and containment primitives with the aim to enforce constraints on the environment the agent runs in, so it cannot override policies, or leak credentials or private data.

DGX Station for Windows Workflows

DGX Station for Windows is built for the full spectrum of enterprise AI workflows -- from autonomous agent deployment and frontier model development to high-throughput inference, data science and physical AI, all with the full Microsoft and Windows enterprise manageability stack.

- **AI Agents:** Build and run multiple frontier agents in parallel, and connect them directly to enterprise applications and workflows.
- **AI Development:** Pretrain, fine-tune and iterate on large AI models within a Windows environment, with access to Linux AI toolchains via Windows Subsystem for Linux.
- **Data Science:** Ingest large datasets into up to 748GB of coherent memory, eliminating data movement bottlenecks and accelerating every step from data preparation to machine learning and analytics.
- **AI Inference:** Run high-throughput inference on AI models, and run large AI models of up to 1 trillion parameters.
- **Physical AI:** Pair the GB300 Superchip with an additional NVIDIA RTX PRO Blackwell Workstation GPU to combine frontier AI compute with ray-traced visualization and simulation in a single desktide system -- delivering the performance needed for agents to perceive, simulate and interact in virtual-to-physical environments.

DGX Station for Windows can serve as a dedicated AI supercomputer for a single developer or a shared local compute node for entire teams -- with workloads scaling seamlessly to GB300 in the data center or cloud.

Global Ecosystem and Availability

DGX Station for Windows is expected to be available from ASUS, Dell Technologies, GIGABYTE, HP, [MSI](#) and Supermicro, coming in Q4 this year.

Watch NVIDIA founder and CEO Jensen Huang's [keynote](#) and learn more at [NVIDIA GTC Taipei](#).

About NVIDIA

[NVIDIA](#) (NASDAQ: NVDA) is the world leader in AI and accelerated computing.

Certain statements in this press release including, but not limited to, statements as to: expectations with respect to growth, performance, availability, and benefits of NVIDIA's products, services and technologies, and related trends and drivers; expectations with respect to NVIDIA's third party arrangements, including with its collaborators and partners; expectations with respect to technology developments, and related trends and drivers; projected market growth and trends; expectations with respect to AI and related industries; and other statements that are not historical facts are forward-looking statements within the meaning of Section 27A of the Securities Act of 1933, as amended, and Section 21E of the Securities Exchange Act of 1934, as amended, which are subject to the "safe harbor" created by those sections based on management's beliefs and assumptions and on information currently available to management and are subject to risks and uncertainties that could cause results to be materially different than expectations. Important factors that could cause actual results to differ materially include: global economic and political conditions; NVIDIA's reliance on third parties to manufacture, assemble, package and test NVIDIA's products; the impact of technological development and competition; development of new products and technologies or enhancements to NVIDIA's existing products and technologies; market acceptance of NVIDIA's products or NVIDIA's partners' products; design, manufacturing or software defects; changes in consumer preferences or demands; changes in industry standards and interfaces; unexpected loss of performance of NVIDIA's products or technologies when integrated into systems; NVIDIA's ability to realize the potential benefits of business investments or acquisitions; and changes in applicable laws and regulations, as well as other factors detailed from time to time in the most recent reports

NVIDIA files with the Securities and Exchange Commission, or SEC, including, but not limited to, its Annual Report on Form 10-K and Quarterly Reports on Form 10-Q. Copies of reports filed with the SEC are posted on the company's website and are available from NVIDIA without charge. These forward-looking statements are not guarantees of future performance and speak only as of the date hereof, and, except as required by law, NVIDIA disclaims any obligation to update these forward-looking statements to reflect future events or circumstances.

© 2026 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, ConnectX, DGX Station, NVIDIA Grace, NVIDIA RTX PRO, NVIDIA RTX Spark NVLink and SuperNIC are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and/or other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Allie Courtney
Corporate Communications
NVIDIA Corporation
press@nvidia.com