

NVIDIA and Partners Build America's AI Infrastructure and Create Blueprint to Power the Next Industrial Revolution

US Government Labs and Nation's Leading Companies Investing in Advanced AI Infrastructure to Power AI Factories and Accelerate US AI Development

News Summary

- Seven new systems across Argonne and Los Alamos National Laboratories to be released, accelerating the Department of Energy's mission of driving technological leadership across U.S. security, science and energy applications.
- NVIDIA AI Factory Research Center in Virginia to host the first Vera Rubin infrastructure and lay the groundwork for NVIDIA Omniverse DSX, a blueprint for multi-generation, gigawatt-scale build-outs using NVIDIA Omniverse libraries.
- Leading U.S. companies across server makers, cloud service providers, model builders, technology suppliers and enterprises are investing in advanced AI infrastructure.

GTC Washington, D.C.—NVIDIA today announced that it is working with the U.S. Department of Energy's national labs and the nation's leading companies to build America's AI infrastructure to support scientific discovery, economic growth and power the next industrial revolution.

"We are at the dawn of the AI industrial revolution that will define the future of every industry and nation," said Jensen Huang, founder and CEO of NVIDIA. "It is imperative that America lead the race to the future — this is our generation's Apollo moment. The next wave of inventions, discoveries and progress will be determined by our nation's ability to scale AI infrastructure. Together with our partners, we are building the most advanced AI infrastructure ever created, ensuring that America has the foundation for a prosperous future, and that the world's AI runs on American innovation, openness and collaboration, for the benefit of all."

NVIDIA AI Advances Scientific Research at National Labs

NVIDIA is accelerating seven new systems by providing the AI infrastructure to drive scientific research and innovation at two U.S. Department of Energy (DOE) facilities — Argonne National Laboratory and Los Alamos National Laboratory (LANL).

NVIDIA is collaborating with Oracle and the DOE to build the U.S. Department of Energy's largest AI supercomputer for scientific discovery. The Solstice system will feature a record-breaking 100,000 NVIDIA Blackwell GPUs and support the DOE's mission of developing AI capabilities to drive technological leadership across U.S. security, science and energy applications.

Another system, Equinox, will include 10,000 NVIDIA Blackwell GPUs expected to be available in 2026. Both systems will be located at Argonne, and will be interconnected by NVIDIA networking and deliver a combined 2,200 exaflops of AI performance.

Argonne is also unveiling three powerful NVIDIA-based systems — Tara, Minerva and Janus — set to expand access to AI-driven computing for researchers across the country. Together, these systems will enable scientists and engineers to revolutionize scientific discovery and boost productivity.

"Argonne's collaboration with NVIDIA and Oracle represents a pivotal step in advancing the nation's AI and computing infrastructure," said Paul K. Kearns, director of Argonne National Laboratory. "Through this partnership, we're building platforms that redefine performance, scalability and scientific potential. Together, we are shaping the foundation for the next generation of computing that will power discovery for decades to come."

LANL, based in New Mexico, announced the selection of the NVIDIA Vera Rubin platform and the [NVIDIA Quantum-X800 InfiniBand](#) networking fabric for its next-generation Mission and Vision systems, to be built and delivered by HPE. The Vision system builds on the achievements of LANL's Venado supercomputer, built for unclassified research. Mission is the fifth Advanced Technology System (ATS5) in the National Nuclear Security Administration's Advanced Simulation and Computing program, which LANL supports, and is expected to be operational in late 2027 and designed to run classified applications.

The Vera Rubin platform will deliver advanced accelerated computing capabilities for these systems, enabling researchers to process and analyze vast datasets at unprecedented speed and scale. Paired with the Quantum-X800 InfiniBand fabric, which delivers high network bandwidth with ultralow latency, the platform enables scientists to run complex simulations to advance areas spanning materials science, climate modeling and quantum computing research.

“Our integration of the NVIDIA Vera Rubin platform and Quantum X800 InfiniBand fabric represents a transformative advancement of our lab — harnessing this level of computational performance is essential to tackling some of the most complex scientific and national security challenges,” said Thom Mason, director of Los Alamos National Laboratory. “Our work with NVIDIA helps us remain at the forefront of innovation, driving discoveries to strengthen the resilience of our critical infrastructure.”

NVIDIA AI Factory Research Center and GigaScale AI Factory Blueprint

NVIDIA also announced the build-out of an AI Factory Research Center at Digital Realty in Virginia. This facility, powered by the NVIDIA Vera Rubin platform, will accelerate breakthroughs in generative AI, scientific computing and advanced manufacturing and serve as a foundation for pioneering research in [digital twins](#) and large-scale simulation.

The center lays the groundwork for [NVIDIA Omniverse DSX](#) — a blueprint for multi-generation, gigawatt-scale build-outs using [NVIDIA Omniverse™](#) libraries — that will set a new standard of excellence for AI infrastructure. By integrating virtual and physical systems, NVIDIA is creating a scalable model for building intelligent facilities that continuously optimize for performance, energy efficiency and sustainability.

With this new center, NVIDIA and its partners are collaborating to develop Omniverse DSX, which will integrate autonomous control systems and modular infrastructure to power the next generation of AI factories. NVIDIA is collaborating with companies to enable the gigawatt-scale rollout of hyperscale AI infrastructure:

- **Engineering and construction partners** [Bechtel](#) and Jacobs are working with NVIDIA to integrate advanced digital twins into validated designs across complex architectural, power, mechanical and electrical systems.
- **Power, cooling and energy equipment partners** including [Eaton](#), GE Vernova, [Hitachi](#), Mitsubishi Electric, [Schneider Electric](#), [Siemens](#), Siemens Energy, Tesla, [Trane Technologies](#) and [Vertiv](#) are contributing to the center. Power and system modeling enable AI factories to dynamically interact with utility networks at gigawatt scale. Liquid-cooling, rectification and power-conversion systems optimized for NVIDIA Grace Blackwell and Vera Rubin platforms are also modeled in the earlier [NVIDIA Omniverse Blueprint for AI factory digital twins](#).
- **Software and agentic AI solutions providers** including [Cadence](#), Emerald AI, Phaidra, [PTC](#), Schneider Electric ETAP, [Siemens](#) and Switch have built digital twin solutions to model and optimize AI factory lifecycles, from design to operation. AI agents continuously optimize power, cooling and workloads, turning the NVIDIA Omniverse DSX blueprint for AI factory digital twins into a self-learning system that boosts grid flexibility, resilience and energy efficiency.

Building the Next Wave of US Infrastructure

Leading U.S. companies across server makers, cloud service providers, model builders, technology suppliers and enterprises are investing in advanced AI infrastructure to power AI factories and accelerate U.S. AI development.

System makers Cisco, Dell Technologies, HPE and Supermicro are collaborating with NVIDIA to build secure, scalable AI infrastructure by integrating NVIDIA GPUs and AI software into their full-stack systems. This includes the newly announced [NVIDIA AI Factory for Government](#) reference design, which will accelerate AI deployments for the public sector and highly regulated industries.

In addition, Cisco is launching the new Nexus N9100 switch series powered by [NVIDIA Spectrum-X™](#) Ethernet switch silicon. The switches' integration with the existing Cisco Nexus management framework will allow customers to seamlessly deploy and manage the new high-speed NVIDIA-powered fabrics using the same trusted tools and operational models they already rely on.

Cisco will now offer an [NVIDIA Cloud Partner](#)-compliant AI factory with the Cisco Cloud reference architecture based on this switch. The N9100 Series switches will be orderable before the end of the year.

Leading Cloud Providers and Model Builders Accelerate AI

Cloud providers and model builders are continuing to invest in AI infrastructure to create a diverse ecosystem for AI innovation, ensuring the U.S. remains at the forefront of AI advancements and their practical applications across industries globally.

The following companies are expanding their commitments to further bolster U.S.-based AI innovation:

- [Akamai](#) is launching Akamai Inference Cloud, a distributed platform that expands AI inference from core data centers to the edge — targeting 20 initial locations across the globe, including five U.S. states, and plans for further expansion — accelerated by NVIDIA RTX PRO™ Servers.
- CoreWeave is establishing CoreWeave Federal, a new business focused on providing secure, compliant, high-performance AI cloud infrastructure and services to the U.S. government running on NVIDIA GPUs and validated designs. The initiative includes anticipated FedRAMP and related agency authorizations of the CoreWeave platform.
- [Global AI](#), a new NVIDIA Cloud Partner, has placed its first big purchase for 128 NVIDIA GB300 NVL72 racks (featuring 9,000+ GPUs), which will be the largest GB300 NVL72 deployment in New York.
- Google Cloud is offering new A4X Max VMs with [NVIDIA GB300 NVL72](#) and G4 VMs with NVIDIA RTX PRO 6000

Blackwell GPUs, as well as bringing the NVIDIA Blackwell platform on premises and in air-gapped environments with Google Distributed Cloud.

- [Lambda](#) is building a new 100+ megawatt AI factory in Kansas City, Missouri. The supercomputer will initially feature more than 10,000 NVIDIA GB300 NVL72 GPUs to accelerate AI breakthroughs from U.S.-based researchers, enterprises and developers.
- [Microsoft](#) is using NVIDIA RTX PRO 6000 Blackwell GPUs on Microsoft Azure, and has recently announced the deployment of a large-scale Azure cluster using NVIDIA GB300 NVL72 for OpenAI. In addition, Microsoft is adding Azure Local support for NVIDIA RTX™ GPUs in the coming months.
- Oracle recently [launched](#) Oracle Cloud Infrastructure Zettascale10, the industry's largest AI supercomputer in the cloud, powered by NVIDIA AI infrastructure.
- Together AI, in partnership with 5C, already operates an AI factory in Maryland featuring NVIDIA B200 GPUs and is bringing a new one online soon in Memphis, Tennessee, featuring NVIDIA GB200 and GB300 systems. Both locations are set for near-term expansion, and new locations will be coming up in 2026 to accelerate the development and scaling of AI-native applications.
- xAI is working on its massive Colossus 2 data center in Memphis, Tennessee, which will house over half a million NVIDIA GPUs — enabling rapid, frontier-level training and inference of next-generation AI models.

US Enterprises Build AI Infrastructure for Industries

Beyond cloud providers and model builders, U.S. organizations are looking to build and offer AI infrastructure for themselves and others that will accelerate workloads across a variety of industries such as pharmaceutical and healthcare.

Lilly is building the pharmaceutical industry's most powerful AI factory with an [NVIDIA DGX SuperPOD™](#) with [NVIDIA DGX™ B300 systems](#), featuring [NVIDIA Spectrum-X Ethernet](#) and [NVIDIA Mission Control™](#) software, which will allow the company to develop and train large-scale biomedical foundation models that aim to accelerate drug discovery and design. This builds on Lilly's use of [NVIDIA RTX PRO Servers](#) to power drug discovery and research by accelerating enterprise AI workloads.

Mayo Clinic — with access to 20 million digitized pathology slides and one of the world's largest patient databases — has created an AI factory powered by DGX SuperPOD with DGX B200 systems and NVIDIA Mission Control. This delivers the AI computational power needed to advance healthcare applications such as medical research, digital pathology and personalized care for better patient outcomes.

Learn more about how NVIDIA and partners are advancing AI innovation in the U.S. by watching the [NVIDIA GTC Washington, D.C., keynote by Huang](#).

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: the next wave of inventions, discoveries and progress being determined by our nation's ability to scale AI infrastructure; together with our partners, NVIDIA building the most advanced AI infrastructure ever created, ensuring that America has the foundation for a prosperous future, and that the world's AI runs on American innovation, openness and collaboration, for the benefit of all; the benefits, impact, performance, and availability of NVIDIA's products, services, and technologies; expectations with respect to NVIDIA's third party arrangements, including with its collaborators and partners; expectations with respect to technology developments; 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 product 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; 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..

Many of the products and features described herein remain in various stages and will be offered on a when-and-if-available basis. The statements above are not intended to be, and should not be interpreted as a commitment, promise, or legal obligation, and the development, release, and timing of any features or functionalities described for our products is subject to change and remains at the sole discretion of NVIDIA. NVIDIA will have no liability for failure to deliver or delay in the delivery

of any of the products, features or functions set forth herein.

© 2025 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, BlueField, ConnectX, NVIDIA DGX, NVIDIA DGX SuperPOD, NVIDIA Omniverse, NVIDIA RTX, NVIDIA RTX PRO and NVIDIA Spectrum-X are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated. Features, pricing, availability and specifications are subject to change without notice.

Kristin Uchiyama
Enterprise and Edge Computing
+1-408-486-2248
kuchiyama@nvidia.com