



NVIDIA DGX Platform

The best of NVIDIA AI—all in one place.



Firms that plan to adopt next-generation AI are 2.6X more likely to increase revenue by 10 percent or more.¹

The capabilities of AI is rapidly accelerating, from recommender systems that understand customer behavior and create tailored experiences to smarter chatbots that provide contextual conversations. Organizations need a full-stack technology solution that can enable them to handle the latest advances of large language models and generative AI and scale innovation across their organization. With the **NVIDIA DGX™ platform**, every organization can now aspire to create cutting-edge breakthroughs using AI, while remaining efficient and agile, even in unprecedented times.

NVIDIA DGX Platform: The Proven Standard for Enterprise AI

Built from the ground up for enterprise AI, the NVIDIA DGX platform incorporates the best of NVIDIA software, infrastructure, and expertise in a modern, unified AI development and training solution. Every aspect of the DGX platform is infused with NVIDIA AI expertise, featuring world-class software, record-breaking NVIDIA-accelerated infrastructure in clouds or on premises, and direct access to NVIDIA DGXperts to speed the ROI of AI for every enterprise.

The Leading Platform for AI Development

AI software and systems that seamlessly spans from cloud to on-premises data centers.

The NVIDIA DGX platform integrates AI software and purpose-built hardware in a comprehensive solution for AI development. Accessible in NVIDIA DGX Cloud or on premises with NVIDIA DGX systems, the DGX platform is built from the ground up for enterprise AI, delivering industry-leading performance and scalability. NVIDIA Base Command™ is the software engine of the DGX platform, enabling advanced management and full-stack optimizations on premises and in the cloud. The DGX platform also includes the NVIDIA AI Enterprise software suite, a large library of fully supported and optimized software that unlocks developer productivity.

Thousands of Leading Companies Deploy the NVIDIA DGX Platform

8 of the Top 10
Global Telcos

7 of the Top 10
Global Pharma Companies

6 of the Top 10
U.S. Banks

9 of the Top 10
Global Car Manufacturers

10 of the Top 10
Global Universities

10 of the Top 10
U.S. Government Institutions

8 of the Top 10
Consumer Internet Companies

7 of the Top 10
US Aerospace and Defense Companies

¹ Accenture Research. [Breakthrough Innovation: Is your organization equipped for breakthrough innovation?](#) WEF 2023

Infused With NVIDIA AI Expertise

When you buy DGX, you get NVIDIA.

Every element in the DGX platform, from the engineering wonders within NVIDIA DGX B200 to the advanced AI software in NVIDIA Base Command, is infused with the AI expertise of NVIDIA. Our researchers have achieved AI breakthroughs using this platform for years and have completed millions of training runs with it. This vast institutional knowledge is incorporated directly into the design and evolution of the DGX platform, meaning the software tools and accelerated compute infrastructure you use today delivers the best of NVIDIA innovation and will keep getting better over time. Additionally, every DGX customer gets direct access to NVIDIA DGXperts to offer prescriptive guidance and help fast-track AI transformation.

Unprecedented Performance, Predictable Cost

Energy-efficient accelerated computing that fits the way you do business.

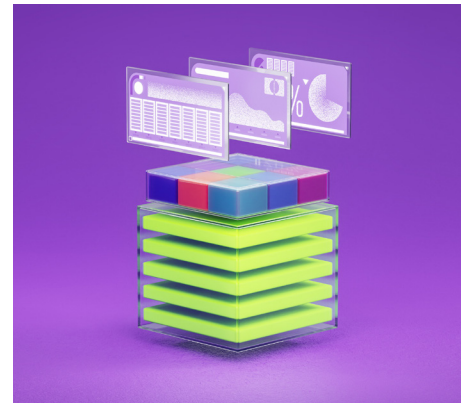
Since the introduction of NVIDIA DGX systems in 2016, DGX has represented the pinnacle of AI performance, with numerous record-breaking achievements in supercomputer performance and energy efficiency. With a platform experience that now transcends clouds and data centers, organizations can experience leading-edge DGX performance using integrated development and workflow management software that seamlessly blends cloud and on-premises systems into a consolidated interface. And no matter what deployment model you choose, the DGX platform provides an easy-to-understand, predictable cost model for AI development infrastructure.

Comprehensive AI Platform

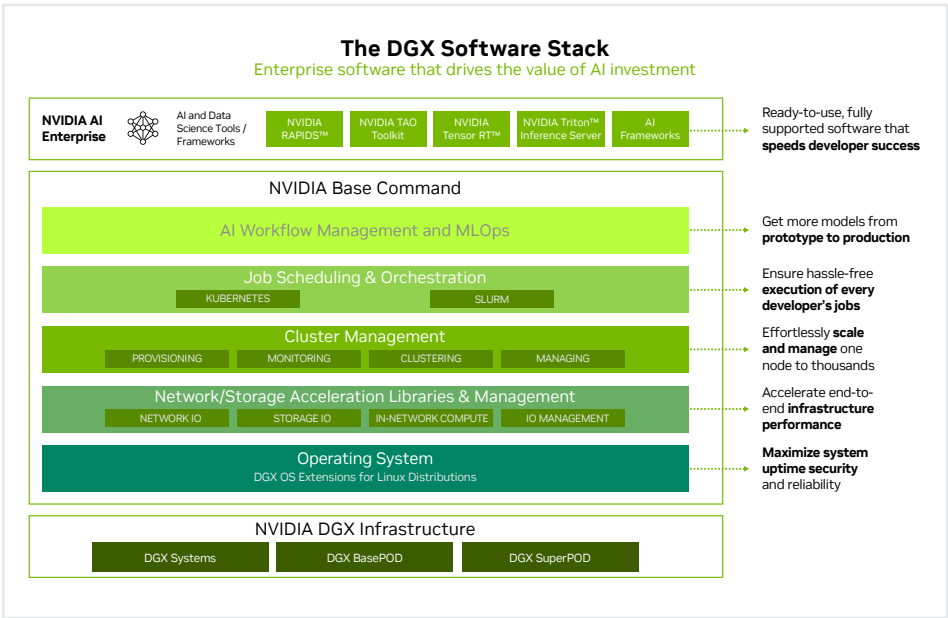
- > **NVIDIA DGX Cloud:** **NVIDIA DGX™ Cloud** is an AI platform designed for enterprise developers, offering a serverless experience. It delivers an integrated, full-stack solution co-engineered with leading cloud partners, integrating their best of breed architectures and incorporating the newest NVIDIA AI technologies across accelerated computing, network fabric, software and direct access to NVIDIA AI experts. DGX Cloud delivers industry-leading utilization efficiency with scale and productive work capacity for developers.
- > **NVIDIA Base Command:** **NVIDIA Base Command** powers the DGX platform, enabling organizations to leverage the best of NVIDIA software innovation. Enterprises can unleash the full potential of their DGX infrastructure with proven software that includes enterprise-grade orchestration and cluster management, libraries that accelerate compute, storage and network infrastructure, and an operating system optimized for AI workloads. Additionally, DGX infrastructure includes **NVIDIA AI Enterprise**, offering a suite of software optimized to streamline AI development and deployment.



NVIDIA DGX Cloud is an AI platform for enterprise developers, optimized for the demands of generative AI.



NVIDIA Base Command is the operating system of the DGX data center.

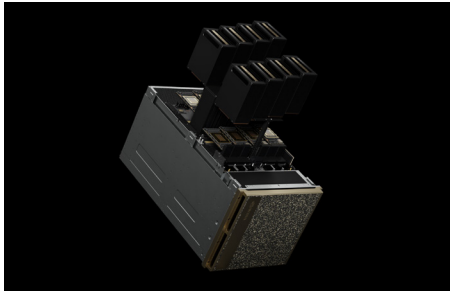


NVIDIA DGX AI Software Stack

- > **NVIDIA DGX SuperPOD:** **NVIDIA DGX SuperPOD™** is AI data center infrastructure that enables IT to deliver performance—without compromise—for every user and workload. As part of the NVIDIA DGX platform, DGX SuperPOD offers leadership-class accelerated infrastructure and scalable performance for the most challenging AI workloads, with industry-proven results.
- > **NVIDIA DGX BasePOD:** AI is powering mission-critical use cases in every industry—from healthcare to manufacturing to financial services. As part of the NVIDIA DGX platform, **NVIDIA DGX BasePOD™** provides the critical foundation on which business transformation is realized and AI applications are born.
- > **NVIDIA DGX B200:** **NVIDIA DGX B200** delivers a unified platform for training, fine-tuning, and inference in a single solution optimized for any enterprise and powered by the NVIDIA Blackwell architecture. The latest addition to the NVIDIA DGX platform, DGX B200 delivers leading-edge performance, offering 3X the AI training and 15X the inference performance of DGX H100. DGX B200 is ideal for businesses looking for a single platform for all of their develop-to-deploy pipelines.
- > **NVIDIA DGX H100:** **NVIDIA DGX H100** is an AI powerhouse that lets enterprises expand the frontiers of business innovation and optimization. DGX H100 is powered by the groundbreaking NVIDIA H100 Tensor Core GPU.



NVIDIA DGX SuperPOD is full-cycle, industry-leading infrastructure for the fastest path to AI innovation at scale.



NVIDIA DGX B200 is the latest iteration of NVIDIA DGX systems, providing a single unified platform for develop-to-deploy pipelines.

Toward an AI Center of Excellence

With the growth in AI and its use in day-to-day operations by companies, many are now starting to recognize the importance of developing AI centers of excellence (CoE). With the DGX platform, companies can leverage DGX Cloud to create their AI CoE and get immediate access to world-class AI infrastructure. Or they can choose to deploy DGX SuperPOD, DGX BasePOD, or individual DGX B200 or DGX H100 systems in their own data center or choose from a variety of **managed service providers, colocation options, or AI consulting partners.**

Trusted AI Experts for the Most Challenging Problems

More than a software or servers, DGX is a complete hardware and software platform backed by thousands of AI experts at NVIDIA. Leveraging the DGX platform gives you direct access to **NVIDIA DGXperts**, a global team of AI-fluent practitioners that offer prescriptive guidance and design expertise to help fast-track AI transformation. This ensures mission-critical applications get up and running quickly and stay running smoothly, dramatically improving time to insights.

“We trained our large language models (LLM) more effectively with NVIDIA DGX SuperPOD’s powerful performance. We considered using other platforms, but it was difficult to find an alternative that provides full-stack environments—from the hardware level to the inference level.”

Hwijung Ryu, LLM Development Team Lead, KT Corporation

Ready to Get Started?

To learn more about the NVIDIA DGX platform visit:
nvidia.com/dgx

© 2024 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA, the NVIDIA logo, Base Command, DGX, DGX BasePOD, and DGX SuperPOD are trademarks and/or registered trademarks of NVIDIA Corporation and affiliates in the U.S. and other countries. All other trademarks and copyrights are the property of their respective owners. 3193770. MAR24

Partner
Logo

