

NVIDIA Puts Grace Blackwell on Every Desk and at Every Al Developer's Fingertips

NVIDIA Project DIGITS With New GB10 Superchip Debuts as World's Smallest Al Supercomputer Capable of Running 200B-Parameter Models

CES—NVIDIA today unveiled NVIDIA[®] Project DIGITS, a personal AI supercomputer that provides AI researchers, data scientists and students worldwide with access to the power of the NVIDIA Grace Blackwell platform.

Project DIGITS features the new NVIDIA GB10 Grace Blackwell Superchip, offering a petaflop of Al computing performance for prototyping, fine-tuning and running large Al models.

With Project DIGITS, users can develop and run inference on models using their own desktop system, then seamlessly deploy the models on accelerated cloud or data center infrastructure.

"Al will be mainstream in every application for every industry. With Project DIGITS, the Grace Blackwell Superchip comes to millions of developers," said Jensen Huang, founder and CEO of NVIDIA. "Placing an AI supercomputer on the desks of every data scientist, AI researcher and student empowers them to engage and shape the age of AI."

GB10 Superchip Provides a Petaflop of Power-Efficient Al Performance

The GB10 Superchip is a system-on-a-chip (SoC) based on the NVIDIA Grace Blackwell architecture and delivers up to 1 petaflop of AI performance at FP4 precision.

GB10 features an NVIDIA Blackwell GPU with latest-generation CUDA® cores and fifth-generation Tensor Cores, connected via NVLink®-C2C chip-to-chip interconnect to a high-performance NVIDIA Grace™ CPU, which includes 20 power-efficient cores built with the Arm architecture. MediaTek, a market leader in Arm-based SoC designs, collaborated on the design of GB10, contributing to its best-in-class power efficiency, performance and connectivity.

The GB10 Superchip enables Project DIGITS to deliver powerful performance using only a standard electrical outlet. Each Project DIGITS features 128GB of unified, coherent memory and up to 4TB of NVMe storage. With the supercomputer, developers can run up to 200-billion-parameter large language models to supercharge Al innovation. In addition, using NVIDIA ConnectX® networking, two Project DIGITS Al supercomputers can be linked to run up to 405-billion-parameter models.

Grace Blackwell Al Supercomputing Within Reach

With the Grace Blackwell architecture, enterprises and researchers can prototype, fine-tune and test models on local Project DIGITS systems running Linux-based NVIDIA DGX OS, and then deploy them seamlessly on NVIDIA DGX Cloud™, accelerated cloud instances or data center infrastructure.

This allows developers to prototype AI on Project DIGITS and then scale on cloud or data center infrastructure, using the same Grace Blackwell architecture and the NVIDIA AI Enterprise software platform.

Project DIGITS users can access an extensive library of NVIDIA AI software for experimentation and prototyping, including software development kits, orchestration tools, frameworks and models available in the NVIDIA NGC catalog and on the NVIDIA Developer portal. Developers can fine-tune models with the NVIDIA NeMo TM framework, accelerate data science with NVIDIA RAPIDS TM libraries and run common frameworks such as PyTorch, Python and Jupyter notebooks.

To build agentic AI applications, users can also harness <u>NVIDIA Blueprints</u> and <u>NVIDIA NIM</u>TM microservices, which are available for research, development and testing via the NVIDIA Developer Program. When AI applications are ready to move from experimentation to production environments, the NVIDIA AI Enterprise license provides enterprise-grade security, support and product releases of NVIDIA AI software.

Availability

Project DIGITS will be available in May from NVIDIA and top partners, starting at \$3,000. Sign up for notifications today.

About NVIDIA

NVIDIA (NASDAQ: NVDA) is the world leader in accelerated computing.

Certain statements in this press release including, but not limited to, statements as to: the benefits, impact, and performance of NVIDIA's products, services, and technologies, including NVIDIA Project DIGITS, NVIDIA Grace Blackwell, NVIDIA GB10 Grace Blackwell Superchip, NVIDIA Blackwell GPU, NVIDIA CUDA cores and fifth-generation Tensor Cores, NVIDIA NVLink-C2C chip-to-chip interconnect, NVIDIA Grace CPU, NVIDIA DGX Cloud, NVIDIA AI Enterprise software platform,

NVIDIA NeMo, NVIDIA RAPIDS libraries, NVIDIA Blueprints, and NVIDIA NIM microservices; Al being mainstream in every application for every industry; with Project DIGITS, the Grace Blackwell Superchip coming to millions of developers; and placing an Al supercomputer on the desks of every data scientist, Al researcher and student empowering them to engage and shape the age of Al are forward-looking statements that 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 conditions; our reliance on third parties to manufacture, assemble, package and test our products; the impact of technological development and competition; development of new products and technologies or enhancements to our existing product and technologies; market acceptance of our products or our partners' products; design, manufacturing or software defects; changes in consumer preferences or demands; changes in industry standards and interfaces; unexpected loss of performance of our products or technologies when integrated into systems; 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, ConnectX, CUDA, DGX, Project DIGITS, NGC, NVIDIA Grace, NVIDIA NeMo, NVIDIA NIM, NVIDIA RAPIDS and NVLink 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.

Kelly Musgrave Senior Manager, Consumer PR NVIDIA Corporation +1-650-421-3748 kmusgrave@nvidia.com