



NVIDIA Contributes Blackwell Platform Design to Open Hardware Ecosystem, Accelerating AI Infrastructure Innovation

NVIDIA GB200 NVL72 Design Contributions and NVIDIA Spectrum-X to Help Accelerate Next Industrial Revolution

OCP Global Summit—To drive the development of open, efficient and scalable data center technologies, NVIDIA today announced that it has contributed foundational elements of its NVIDIA Blackwell accelerated computing platform design to the [Open Compute Project](#) (OCP) and broadened NVIDIA Spectrum-X™ support for OCP standards.

At this year's OCP Global Summit, NVIDIA will be sharing key portions of the [NVIDIA GB200 NVL72](#) system electro-mechanical design with the OCP community — including the rack architecture, compute and switch tray mechanicals, liquid-cooling and thermal environment specifications, and [NVIDIA NVLink™](#) cable cartridge volumetrics — to support higher compute density and networking bandwidth.

NVIDIA has already made several official contributions to OCP across multiple hardware generations, including its [NVIDIA HGX™ H100](#) baseboard design specification, to help provide the ecosystem with a wider choice of offerings from the world's computer makers and expand the adoption of AI.

In addition, expanded [NVIDIA Spectrum-X Ethernet networking platform](#) alignment with OCP Community-developed specifications enables companies to unlock the performance potential of AI factories deploying OCP-recognized equipment while preserving their investments and maintaining software consistency.

"Building on a decade of collaboration with OCP, NVIDIA is working alongside industry leaders to shape specifications and designs that can be widely adopted across the entire data center," said Jensen Huang, founder and CEO of NVIDIA. "By advancing open standards, we're helping organizations worldwide take advantage of the full potential of accelerated computing and create the AI factories of the future."

Accelerated Computing Platform for the Next Industrial Revolution

NVIDIA's accelerated computing platform was designed to power a new era of AI.

GB200 NVL72 is based on the [NVIDIA MGX™ modular architecture](#), which enables computer makers to quickly and cost-effectively build a vast array of data center infrastructure designs.

The liquid-cooled system connects 36 [NVIDIA Grace™ CPUs](#) and 72 [NVIDIA Blackwell GPUs](#) in a rack-scale design. With a 72-GPU NVIDIA NVLink domain, it acts as a single, massive GPU and delivers 30x faster real-time trillion-parameter large language model inference than the NVIDIA H100 Tensor Core GPU.

The NVIDIA Spectrum-X Ethernet networking platform, which now includes the next-generation [NVIDIA ConnectX-8 SuperNIC™](#), supports OCP's Switch Abstraction Interface (SAI) and Software for Open Networking in the Cloud (SONiC) standards. This allows customers to use Spectrum-X's adaptive routing and telemetry-based congestion control to accelerate Ethernet performance for scale-out AI infrastructure.

[ConnectX-8 SuperNICs](#) feature accelerated networking at speeds of up to 800Gb/s and programmable packet processing engines optimized for massive-scale AI workloads. ConnectX-8 SuperNICs for OCP 3.0 will be available next year, equipping organizations to build highly flexible networks.

Critical Infrastructure for Data Centers

As the world transitions from general-purpose to accelerated and AI computing, data center infrastructure is becoming increasingly complex. To simplify the development process, NVIDIA is working closely with 40+ global electronics makers that provide key components to create AI factories.

Additionally, a broad array of partners are innovating and building on top of the Blackwell platform, including [Meta](#), which plans to contribute its Catalina AI rack architecture based on GB200 NVL72 to OCP. This provides computer makers with flexible options to build high compute density systems and meet the growing performance and energy efficiency needs of data centers.

"NVIDIA has been a significant contributor to open computing standards for years, including their high-performance computing platform that has been the foundation of our Grand Teton server for the past two years," said Yee Jiun Song, vice president of engineering at Meta. "As we progress to meet the increasing computational demands of large-scale artificial

intelligence, NVIDIA's latest contributions in rack design and modular architecture will help speed up the development and implementation of AI infrastructure across the industry."

[Learn more](#) about NVIDIA's contributions to the Open Compute Project at the 2024 OCP Global Summit, taking place at the San Jose Convention Center from Oct. 15-17.

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 Blackwell accelerated computing platform, NVIDIA Spectrum-X Ethernet networking platform, NVIDIA GB200 NVL72, NVIDIA NVLink, NVIDIA HGX H100, NVIDIA MGX modular architecture, NVIDIA Grace CPUs, NVIDIA H100 Tensor Core GPU, and NVIDIA ConnectX-8 SuperNIC; NVIDIA contributing foundational elements of its NVIDIA Blackwell accelerated computing platform design to the Open Compute Project (OCP) and broaden NVIDIA Spectrum-X support for OCP standards; the benefits and impact of NVIDIA's collaboration with third parties; third parties using or adopting our products or technologies; NVIDIA working alongside industry leaders to shape specifications and designs that can be widely adopted across the entire data center; by advancing open standards, NVIDIA helping organizations worldwide take advantage of the full potential of accelerated computing and create the AI factories of the future; as the world transitioning from general-purpose to accelerated and AI computing, data center infrastructure becoming increasingly complex; and the timing and themes of the 2024 OCP Global Summit 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.

© 2024 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, ConnectX, NVIDIA Grace, NVIDIA HGX, NVIDIA MGX, NVIDIA Spectrum-X, NVLink and SuperNIC 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