Computer Industry Joins NVIDIA to Build AI Factories and Data Centers for the Next Industrial Revolution

- Top Computer Manufacturers Unveil Array of Blackwell-Powered Systems Featuring Grace CPUs, NVIDIA Networking and Infrastructure
- Broad Portfolios Encompass Cloud, On-Premises, Embedded and Edge AI Systems
- Offerings Range from Single to Multi-GPUs, x86 to Grace, Air to Liquid Cooling

COMPUTEX—NVIDIA and the world’s top computer manufacturers today unveiled an array of NVIDIA Blackwell architecture-powered systems featuring Grace CPUs, NVIDIA networking and infrastructure for enterprises to build AI factories and data centers to drive the next wave of generative AI breakthroughs.

During his COMPUTEX keynote, NVIDIA founder and CEO Jensen Huang announced that ASRock Rack, ASUS, GIGABYTE, Ingrasys, Inventec, Pegatron, QCT, Supermicro, Wistron and Wwynn will deliver cloud, on-premises, embedded and edge AI systems using NVIDIA GPUs and networking.

“The next industrial revolution has begun. Companies and countries are partnering with NVIDIA to shift the trillion-dollar traditional data centers to accelerated computing and build a new type of data center — AI factories — to produce a new commodity: artificial intelligence,” said Huang. “From server, networking and infrastructure manufacturers to software developers, the whole industry is gearing up for Blackwell to accelerate AI-powered innovation for every field.”

To address applications of all types, the offerings will range from single to multi-GPUs, x86- to Grace-based processors, and air- to liquid-cooling technology.

Additionally, to speed up the development of systems of different sizes and configurations, the NVIDIA MGX™ modular reference design platform now supports NVIDIA Blackwell products. This includes the new NVIDIA GB200 NVL2 platform, built to deliver unparalleled performance for mainstream large language model inference, retrieval-augmented generation and data processing.

GB200 NVL2 is ideally suited for emerging market opportunities such as data analytics, on which companies spend tens of billions of dollars annually. Taking advantage of high-bandwidth memory performance provided by NVLink®-C2C interconnects and dedicated decompression engines in the Blackwell architecture speeds up data processing by up to 18x, with 8x better energy efficiency compared to using x86 CPUs.

Modular Reference Architecture for Accelerated Computing

To meet the diverse accelerated computing needs of the world’s data centers, NVIDIA MGX provides computer manufacturers with a reference architecture to quickly and cost-effectively build more than 100 system design configurations.

Manufacturers start with a basic system architecture for their server chassis, and then select their GPU, DPU and CPU to address different workloads. To date, more than 90 systems from over 25 partners have been released or are in development that leverage the MGX reference architecture, up from 14 systems from six partners last year. Using MGX can help slash development costs by up to three-quarters and reduce development time by two-thirds, to just six months.

AMD and Intel are supporting the MGX architecture with plans to deliver, for the first time, their own CPU host processor module designs. This includes the next-generation AMD Turin platform and the Intel® Xeon® 6 processor with P-cores (formerly codenamed Granite Rapids). Any server system builder can use these reference designs to save development time while ensuring consistency in design and performance.

NVIDIA’s latest platform, the GB200 NVL2, also leverages MGX and Blackwell. Its scale-out, single-node design enables a wide variety of system configurations and networking options to seamlessly integrate accelerated computing into existing data center infrastructure.

The GB200 NVL2 joins the Blackwell product lineup, which also includes NVIDIA Blackwell Tensor Core GPUs, GB200 Grace Blackwell Superchips and the GB200 NVL72.

An Ecosystem Unites

NVIDIA’s comprehensive partner ecosystem includes TSMC, the world’s leading semiconductor manufacturer and an NVIDIA foundry partner, as well as global electronics makers, which provide key components to create AI factories. These include manufacturing innovations such as server racks, power delivery, cooling solutions and more from companies such as Amphenol, Asia Vital Components (AVC), Cooler Master, Colder Products Company (CPC), Danfoss, Delta Electronics and LITEON.

As a result, new data center infrastructure can quickly be developed and deployed to meet the needs of the world’s
generative AI services to provide more personalized experiences to its customers. Solution platforms for AI-powered electric vehicle and robotics platforms, as well as a growing number of language-based enterprises — and further accelerated by Blackwell technology, NVIDIA Quantum-2 or Quantum-X800 InfiniBand networking, NVIDIA Spectrum™-X Ethernet networking and NVIDIA BlueField®-3 DPUs — in servers from leading systems makers Dell Technologies, Hewlett Packard Enterprise and Lenovo.

Enterprises can also access the NVIDIA AI Enterprise software platform, which includes NVIDIA NIM™ inference microservices, to create and run production-grade generative AI applications.

Taiwan Embraces Blackwell

Huang also announced during his keynote that Taiwan’s leading companies are rapidly adopting Blackwell to bring the power of AI to their own businesses.

Taiwan’s leading medical center, Chang Gung Memorial Hospital, plans to use the NVIDIA Blackwell computing platform to advance biomedical research and accelerate imaging and language applications to improve clinical workflows, ultimately enhancing patient care.

Foxconn, one of the world’s largest makers of electronics, is planning to use NVIDIA Grace Blackwell to develop smart solution platforms for AI-powered electric vehicle and robotics platforms, as well as a growing number of language-based generative AI services to provide more personalized experiences to its customers.

Additional Supporting Quotes

- R. Adam Norwitt, president and CEO at Amphenol: “NVIDIA’s groundbreaking AI systems require advanced interconnect solutions, and Amphenol is proud to be supplying critical components. As an important partner in NVIDIA’s rich ecosystem, we are able to provide highly complex and efficient interconnect products for Blackwell accelerators to help deliver cutting-edge performance.”
- Spencer Shen, chairman and CEO at AVC: “AVC plays a key role in NVIDIA products, providing efficient cooling for its AI hardware, including the latest Grace Blackwell Superchip. As AI models and workloads continue to grow, reliable thermal management is important to handle intensive AI computing — and we’re with NVIDIA every step of the way.”
- Jonney Shih, chairman at ASUS: “ASUS is working with NVIDIA to take enterprise AI to new heights with our powerful server lineup, which we’ll be showcasing at COMPUTEX. Using NVIDIA’s MGX and Blackwell platforms, we’re able to craft tailored data center solutions built to handle customer workloads across training, inference, data analytics and HPC.”
- Janel Wittmayer, president of Dover Corporation’s CPC: “CPC’s innovative, purpose-built connector technology enables the easy and reliable connection of liquid-cooled NVIDIA GPUs in AI systems. With a shared vision of performance and quality, CPC has the capacity and expertise to supply critical technological components to support NVIDIA’s incredible growth and progress. Our connectors are central to maintaining the integrity of temperature-sensitive products, which is important when AI systems are running compute-intensive tasks. We are excited to be part of the NVIDIA ecosystem and bring our technology to new applications.”
- Andy Lin, CEO at Cooler Master: “As the demand for accelerated computing continues to soar, so does demand for solutions that effectively meet energy standards for enterprises leveraging cutting-edge accelerators. As a pioneer in thermal management solutions, Cooler Master is helping unlock the full potential of the NVIDIA Blackwell platform, which will deliver incredible performance to customers.”
- Kim Fausing, CEO at Danfoss: “Danfoss’ focus on innovative, high-performance quick disconnect and fluid power designs makes our couplings valuable for enabling efficient, reliable and safe operation in data centers. As a vital part of NVIDIA’s AI ecosystem, our work together enables data centers to meet surging AI demands while minimizing environmental impact.”
- Ping Cheng, chairman and CEO at Delta Electronics: “The ubiquitous demand for computing power has ignited a new era of accelerated performance capabilities. Through our advanced cooling and power systems, Delta has developed innovative solutions capable of enabling NVIDIA’s Blackwell platform to operate at peak performance levels, while maintaining energy and thermal efficiency.”
- Etay Lee, vice president and general manager at GIGABYTE: “With our collaboration spanning nearly three decades, GIGABYTE has a deep commitment to supporting NVIDIA technologies across GPUs, CPUs, DPUs and high-speed networking. For enterprises to achieve even greater performance and energy efficiency for the compute-intensive workloads, we’re bringing to market a broad range of Blackwell-based systems.”
- Young Liu, chairman and CEO at Hon Hai Technology Group: “As generative AI transforms industries, Foxconn stands ready with cutting-edge solutions to meet the most diverse and demanding computing needs. Not only do we use the latest Blackwell platform in our own servers, but we also help provide the key components to NVIDIA, giving our customers faster time-to-market.”
- Jack Tsai, president at Inventec: “For nearly half a century, Inventec has been designing and manufacturing electronic products and components — the lifeblood of our business. Through our NVIDIA MGX rack-based solution powered by the NVIDIA Grace Blackwell Superchip, we’re helping customers enter a new realm of AI capability and performance.”
- Anson Chiu, president at LITEON Technology: “In pursuit of greener and more sustainable data centers, power management and cooling solutions are taking center stage. With the launch of the NVIDIA Blackwell platform, LITEON is releasing multiple liquid-cooling solutions that enable NVIDIA partners to unlock the future of highly efficient, environmentally friendly data centers.”
Barry Lam, chairman at Quanta Computer: “We stand at the center of an AI-driven world, where innovation is accelerating like never before. NVIDIA Blackwell is not just an engine; it is the spark igniting this industrial revolution. When defining the next era of generative AI, Quanta proudly joins NVIDIA on this amazing journey. Together, we will shape and define a new chapter of AI.”

Charles Liang, president and CEO at Supermicro: “Our building-block architecture and rack-scale, liquid-cooling solutions, combined with our in-house engineering and global production capacity of 5,000 racks per month, enable us to quickly deliver a wide range of game-changing NVIDIA AI platform-based products to AI factories worldwide. Our liquid-cooled or air-cooled high-performance systems with rack-scale design, optimized for all products based on the NVIDIA Blackwell architecture, will give customers an incredible choice of platforms to meet their needs for next-level computing, as well as a major leap into the future of AI.”

C.C. Wei, CEO at TSMC: “TSMC works closely with NVIDIA to push the limits of semiconductor innovation that enables them to realize their visions for AI. Our industry-leading semiconductor manufacturing technologies helped shape NVIDIA’s groundbreaking GPUs, including those based on the Blackwell architecture.”

Jeff Lin, CEO at Wistron: “As a key manufacturing partner, Wistron has been on an incredible journey alongside NVIDIA delivering GPU computing technologies and AI cloud solutions to customers. Now we’re working with NVIDIA’s latest GPU architectures and reference designs, such as Blackwell and MGX, to quickly bring tremendous new AI computing products to market.”

William Lin, president at Wiwynn: “Wiwynn is focused on helping customers address the rising demand for massive computing power and advanced cooling solutions in the era of generative AI. With our latest lineup based on the NVIDIA Grace Blackwell and MGX platforms, we’re building optimized, rack-level, liquid-cooled AI servers tailored specifically for the demanding workloads of hyperscale cloud providers and enterprises.”

To learn more about the NVIDIA Blackwell and MGX platforms, watch Huang’s COMPUTEX keynote.

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, performance, and availability of our products, services, and technologies, including NVIDIA Blackwell architecture-powered systems, NVIDIA networking and infrastructure for enterprises, NVIDIA MGX modular reference design platform, NVIDIA GB200 NVL2 platform, NVLink-C2C, NVIDIA Blackwell Tensor Core GPUs, GB200 Grace Blackwell Superchips, GB200 NVL72, NVIDIA Quantum-2 and Quantum-X800 InfiniBand networking, NVIDIA Spectrum-X Ethernet networking, and NVIDIA BlueField-3 DPUs; third parties using and adopting our technologies and products, our collaboration and partnership with third parties and the benefits and impact thereof, and the features, performance and availability of their offerings; generative AI being the defining technology of our time, and Blackwell being the engine that will drive this new industrial revolution; and the whole industry — from server, networking and infrastructure manufacturers to software developers — gearing up for Blackwell to accelerate AI-powered innovation for every field 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.

© 2024 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, BlueField, NVIDIA MGX, NVIDIA NIM, NVIDIA Spectrum, 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.

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