



NVIDIA Launches Accelerated Ethernet Platform for Hyperscale Generative AI

New NVIDIA Spectrum-X Networking Platform Combines NVIDIA Spectrum-4, BlueField-3 DPUs and Acceleration Software; World-Leading Cloud Service Providers Adopting Platform to Scale Out Generative AI Services

COMPUTEX—NVIDIA today announced [NVIDIA Spectrum-X™](#), an accelerated networking platform designed to improve the performance and efficiency of Ethernet-based AI clouds.

NVIDIA Spectrum-X is built on networking innovations powered by the tight coupling of the [NVIDIA Spectrum-4 Ethernet switch](#) with the [NVIDIA BlueField@-3 DPU](#), achieving 1.7x better overall AI performance and power efficiency, along with consistent, predictable performance in multi-tenant environments. Spectrum-X is supercharged by NVIDIA acceleration software and software development kits (SDKs), allowing developers to build software-defined, cloud-native AI applications.

The delivery of end-to-end capabilities reduces run-times of massive transformer-based generative AI models. This allows network engineers, AI data scientists and cloud service providers to improve results and make informed decisions faster.

The world's top hyperscalers are adopting NVIDIA Spectrum-X, including industry-leading cloud innovators.

As a blueprint and testbed for NVIDIA Spectrum-X reference designs, NVIDIA is building Israel-1, a hyperscale generative AI supercomputer to be deployed in its Israeli data center on Dell PowerEdge XE9680 servers based on the [NVIDIA HGX™ H100 eight-GPU platform](#), BlueField-3 DPUs and Spectrum-4 switches.

“Transformative technologies such as generative AI are forcing every enterprise to push the boundaries of data center performance in pursuit of competitive advantage,” said Gilad Shainer, senior vice president of networking at NVIDIA. “NVIDIA Spectrum-X is a new class of Ethernet networking that removes barriers for next-generation AI workloads that have the potential to transform entire industries.”

The NVIDIA Spectrum-X networking platform is highly versatile and can be used in various AI applications. It uses fully standards-based Ethernet and is interoperable with Ethernet-based stacks.

The platform starts with Spectrum-4, the world's first 51Tb/sec Ethernet switch built specifically for AI networks. Advanced RoCE extensions work in concert across the Spectrum-4 switches, BlueField-3 DPUs and NVIDIA LinkX optics to create an end-to-end 400GbE network that is optimized for AI clouds.

NVIDIA Spectrum-X enhances multi-tenancy with performance isolation to ensure tenants' AI workloads perform optimally and consistently. It also offers better AI performance visibility, as it can identify performance bottlenecks and it features completely automated fabric validation.

Acceleration software driving Spectrum-X includes powerful NVIDIA SDKs such as [Cumulus Linux](#), pure [SONiC](#) and [NetQ](#) — which together enable the networking platform's extreme performance. It also includes the [NVIDIA DOCA™ software framework](#), which is at the heart of BlueField DPUs.

NVIDIA Spectrum-X enables unprecedented scale of 256 200Gb/s ports connected by a single switch, or 16,000 ports in a two-tier leaf-spine topology to support the growth and expansion of AI clouds while maintaining high levels of performance and minimizing network latency.

Immediate Ecosystem Adoption

Companies offering NVIDIA Spectrum-X include Dell Technologies, Lenovo and [Supermicro](#).

Availability

NVIDIA [Spectrum-X](#), [Spectrum-4 switches](#), [BlueField-3 DPUs](#) and 400G LinkX optics are available now.

Learn more about NVIDIA Spectrum-X at [COMPUTEX](#).

About NVIDIA

Since its founding in 1993, [NVIDIA](#) (NASDAQ: NVDA) has been a pioneer in accelerated computing. The company's invention of the GPU in 1999 sparked the growth of the PC gaming market, redefined computer graphics, ignited the era of modern AI and is fueling the creation of the industrial metaverse. NVIDIA is now a full-stack computing company with data-center-scale offerings that are reshaping industry. More information at <https://nvidianews.nvidia.com/>.

Certain statements in this press release including, but not limited to, statements as to: the benefits, impact, performance,

features and availability of our products, collaborations, services and technologies, including NVIDIA Spectrum-X networking platform, Spectrum-4 switches, BlueField-3 DPUs, NVIDIA acceleration software and SDKs, Israel-1, NVIDIA HGX H100 eight-GPU platform, RoCE, LinkX, 400G LinkX optics, Cumulus Linux, pure SONiC, NetQ and NVIDIA DOCA; our collaborations with Dell Technologies, Lenovo and Supermicro, and the benefits, impact, features and availability thereof; the world's top hyperscalers adopting NVIDIA Spectrum-X; and transformative technologies such as generative AI forcing every enterprise to push the boundaries of data center performance in pursuit of competitive advantage 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.

© 2023 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, BlueField, NVIDIA DOCA, NVIDIA HGX, NVIDIA Spectrum 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.

Alex Shapiro
Enterprise Networking
1-415-608-5044
ashapiro@nvidia.com