NEWS RELEASE

Cisco Delivers AI Innovations across Neocloud, Enterprise and Telecom with NVIDIA

2025-10-28

Cisco offers a NVIDIA Cloud Partner-compliant reference architecture based on new data center switching solutions, providing customers ultimate flexibility as they build critical AI infrastructure.

News Summary:

- With the new Cisco N9100 series switch, Cisco now offers neocloud and sovereign cloud customers a NVIDIA Cloud Partner-compliant reference architecture, delivering a unified operating model with the flexibility of NX-OS or SONiC.
- Cisco is introducing the Cisco Cloud Reference Architecture for neocloud and sovereign cloud customers; the new offering is based on the tenets of the NVIDIA Cloud Partner reference architecture and utilizes Cisco Silicon One-based switches with embedded NVIDIA Spectrum-X capabilities.
- For enterprises, Cisco strengthens the Secure AI Factory with NVIDIA through advancements in compute, security, networking, observability, and new ecosystem partnerships.
- Cisco, NVIDIA, and other telecom industry partners unveil the first AI-native wireless stack for 6G, empowering telecom providers with a network transition path for the AI era.

WASHINGTON, Oct. 28, 2025 /PRNewswire/ -- **GTC** -- Cisco (NASDAQ: CSCO) today introduced major advancements to accelerate secure, scalable AI across market segments. Leading the announcements is the Cisco N9100, the first NVIDIA partner-developed data center switch based on NVIDIA Spectrum-X Ethernet switch silicon. With this switch, Cisco is offering a NVIDIA Cloud Partner-compliant reference architecture for neocloud and sovereign cloud deployments. For enterprise customers, Cisco Secure AI Factory with NVIDIA, strengthens protection and visibility across AI deployments with new security and observability integrations. To pave the way for next-generation connectivity in the telecom industry, Cisco, NVIDIA and additional partners unveiled the industry's first AI-native wireless stack for 6G. Together, these innovations offer neocloud, enterprise, and telecom customers the flexibility and interoperability to efficiently build, manage and secure AI infrastructure at scale.

"We're at the beginning of the largest data center build-out in history," said Jeetu Patel, President and

Chief Product Officer, Cisco. "The infrastructure that will power the agentic AI applications and innovation of the future requires new architectures designed to overcome today's constraints in power, computing, and network performance. Together, Cisco and NVIDIA are leading the way in defining the technologies that will power these AI-ready data centers in all their varieties, from emerging neoclouds, to global service providers, to enterprises, and beyond."

"NVIDIA Spectrum-X Ethernet delivers the performance of accelerated networking for Ethernet," said Gilad Shainer, SVP of Networking at NVIDIA. "Working with Cisco's Cloud Reference Architectures and NVIDIA Cloud Partner design principles, customers can choose to deploy Spectrum-X Ethernet using the newest Cisco N9100 series or Cisco Silicon One based switches to build open, high-performance AI networks."

A Portfolio for Any AI Workload
Back-end and front-end Ethernet-based networks must be flexible enough to keep pace with rapid AI innovation, integrate seamlessly with existing infrastructure, and be simple to deploy and manage. Orderable before the end of the year, the Cisco N9100 series switches offer a choice of Cisco NX-OS or SONiC operating systems, advancing Ethernet for AI networks and offering greater flexibility in how neocloud and sovereign cloud customers build their AI infrastructure. With the N9100 as a foundation, Cisco will offer a NVIDIA Cloud Partner-compliant reference architecture. Cisco's portfolio of Nexus data center switching solutions provides a unified operating model through Cisco Nexus Dashboard, across Silicon One, Cloud-scale ASICs, and now switches built on Spectrum-X Ethernet switch silicon.

Additionally, for neocloud and sovereign cloud customers, the Cisco Cloud Reference Architecture is based on the design tenets of NVIDIA's Cloud Partner reference architecture and utilizes Cisco's Silicon One and Cloud-scale ASIC offerings. The reference architecture will also include the recently introduced Cisco 8223 based on the Silicon One P200 for scale-across networks, NVIDIA BlueField-4 DPUs, and NVIDIA ConnectX-9 SuperNICs.

Cisco Secure AI Factory with NVIDIA: Built for Performance, Security, Resiliency
Since its unveiling at GTC in March 2025, the Cisco Secure AI Factory with NVIDIA has led the industry in offering enterprises a comprehensive architecture for AI infrastructure that puts security and observability at the forefront without sacrificing performance. With Cisco AI PODs and Cisco Silicon One-powered Nexus switching as a foundation, Cisco today is delivering new capabilities and features across:

- **Security and Observability** - Cisco AI Defense now integrates with NVIDIA NeMo Guardrails to deliver robust cybersecurity for AI applications. Cisco AI Defense is orderable for on-premises data-plane deployment enabling security and AI teams to protect AI models and applications, limiting the sensitive data that leaves their organization's data centers. Also available, Splunk Observability Cloud helps teams to monitor the performance, quality, security, and cost of their AI application stack—including real-time insights into AI infrastructure health with Cisco AI PODs—while Splunk Enterprise Security extends this visibility to protect AI workloads.
- ● **Core AI Infrastructure -** Cisco Isovalent is now validated for inference workloads on AI PODs, enabling enterprise grade, high-performance Kubernetes networking. Cisco Nexus Hyperfabric AI with a new cloud-managed Cisco G200 Silicon One switch that delivers high-density 800G Ethernet, is now orderable as a deployment option in AI PODs. Cisco UCS 880A M8 rack servers with NVIDIA HGX B300, and the Cisco UCS X-Series modular servers with NVIDIA RTX PRO 6000 Blackwell Server Edition GPUs are also now orderable as part of AI PODs. This enables high-performance GPU support for a wide range of workloads including generative AI fine-tuning, inference and more.
- **Ecosystem expansion** - NVIDIA Run:ai software is available through Cisco and its partners, enabling intelligent AI workload and GPU orchestration capabilities. Nutanix Kubernetes Platform

(NKP) solution is now a supported Kubernetes platform, and Nutanix Unified Storage (NUS) solution is now a supported storage option, with Nutanix Enterprise AI (NAI) solution as the interoperable software component that simplifies building and operating containerized inference services.
- **Government-Ready** - Cisco is collaborating with NVIDIA and aligning to the new NVIDIA AI Factory for Government, a full-stack end-to-end reference design for AI workloads deployed in highly regulated environments.

The First AI-native Wireless Stack – with Cisco at its Core
As AI moves from smartphones to more connected things – augmented reality glasses, connected cars and robotics – wireless networks face mounting demand to support billions of connections at unprecedented scale and efficiency. To meet this challenge, Cisco, NVIDIA, and additional telecom partners have developed the first American AI-RAN stack for mobile networks that integrates sensing and communication, with multiple pre-6G applications being showcased at NVIDIA GTC DC. It allows telecom providers to infuse AI into their mobile networks, starting with 5G advanced services and establishes the groundwork for 6G. The stack combines Cisco's user plane function and 5G core software with the NVIDIA AI Aerial platform, creating a foundation that enables physical AI and integrated sensing with unmatched efficiency and security.

Cisco and NVIDIA: Moving AI Forward, Together
Cisco and NVIDIA's collaboration continues to accelerate, driven by a shared vision of an AI-powered future that is scalable, observable, and secure. The advancements announced today are a testament to Cisco's relentless pursuit of innovation, that will accelerate AI adoption across enterprises, neoclouds and telecom providers.

Industry Reactions:
"The real challenge in AI infrastructure isn't just performance—it's maintaining operational sanity as you scale from dozens to thousands of GPUs. Cisco's approach with NX-OS and Nexus Dashboard creates a single pane of glass across our entire AI fabric, whether we're optimizing inference latency in the front-end or maximizing training throughput in the back-end. That operational simplicity translates directly to faster deployments and lower TCO." **– Xiaohe Hu, CEO, Infrawaves**

"Cisco's N9100 series powered by NVIDIA Spectrum-X Ethernet switch silicon, provides a solution for high-performance, open infrastructure to meet our AI cloud demands. The capability to run NX-OS or SONiC under a unified operating model on Nexus Dashboard delivers more flexibility to our customers with operational simplicity. Its enterprise-grade networking with the scale and agility of the cloud — exactly what the next generation of AI workloads requires." **– Yih Leong Sun, Head of Infra, GMI Cloud**

"As the demand for computing power continues to grow, our GPU clusters are expanding rapidly in scale. In ultra-large-scale GPU networks, we face various challenges such as congestion management and load balance. Cisco's NCP compliant reference architecture with N9100 Series switch provides us with the desired performance and openness out of the box. With the Cisco Nexus platform, we can fully leverage the AI networking capabilities of NVIDIA Spectrum-X Ethernet switch silicon without incurring additional operational or development costs and seamlessly integrate with our existing systems. We look forward to partnering with Cisco on this exciting journey ahead." **– Junfeng Cheng, Head of Networking Infrastructure, Xiaohongshu (RedNote)**

"World Wide Technology (WWT) clients know and trust Cisco networking in the enterprise data center. Bringing that together with the innovation of NVIDIA Spectrum-X Ethernet technology extends the value of our clients' investment in the data center to now include AI workloads. This will be critical as our clients look to scale AI in the enterprise data center." **— Neil Anderson, VP and CTO Cloud, Infrastructure and AI Solutions at World Wide Technology (WWT)**

"BlueSky Compute is excited to be one of the first Neoclouds to deploy Cisco's N9100 series switches for the scale-out fabric in our AI training clusters- powering our vision to turn AI into ROI for the world's enterprises by building larger, more interconnected B300 clusters, faster. Cisco's NCP compliant reference architecture based on N9100 Series switch while maintaining the operational simplicity with Nexus Dashboard is a game-changer." **– Ian Hartley, CEO Bluesky Compute**

"We've deployed thousands of GPUs across our customers' AI infrastructure, and network complexity has been our biggest scaling challenge. Cisco's Nexus N9100 series platform based on NVIDIA Spectrum-X Ethernet switch silicon directly addresses this with NCP RA compliance - giving us the performance we need with the openness we require. The unified Nexus operating model is particularly compelling as it maintains consistent operations. This flexibility to meet customers where they are, rather than forcing architectural decisions, is reshaping how customers approach AI infrastructure." **– Thomas Berger, Computacenter**

"Shanghaj Lichan Co., Ltd., as an NVIDIA Cloud Partner, offers a comprehensive full-stack solution that integrates both hardware and software to power AI-driven cloud services. Our services include end-to-end capabilities, from consulting and planning, to testing, deployment, and implementation, as well as ongoing operations and maintenance. We are excited about the upcoming launch of the Cisco N9100 Series . We believe its unified operating model with Nexus Dashboard and NCP-compliant reference architecture will simplify deployments and scale our AI infrastructure more efficiently. This solution is designed to accelerate innovation, reduce costs, and deliver large-scale capabilities to our customers faster than ever before." **– Wendy Wu, Chairman, Shanghai Lichan Technology Co., Ltd.**

Additional Resources:

- Executive blog: Cisco Drives AI Networking Innovation with NVIDIA by Will Eatherton, SVP, Data Center, Internet & Cloud Infrastructure Engineering, Cisco
- Executive Blog: Avatar Cisco Nexus Delivers New AI Innovations with NVIDIA by Murali Gandluru, VP , Data Center Networking, Cisco
- Executive Blog: From AI Pilots to Production: Building Infrastructure That Makes AI Real by Jeremy Foster, SVP, Data Center Compute, Cisco
- Executive Blog: Leading the Next Era of Intelligent Connectivity by Masum Mir, SVP & GM, Cisco Provider Mobility
- Product blog: Cisco AI Defense Integrates with NVIDIA AI Enterprise Software to Secure AI Applications Using NVIDIA NeMo Guardrails
- Product blog: Unlocking AI Performance: Splunk Observability for Cisco Secure AI Factory with NVIDIA
- NVIDIA AI Factory for Government
- NVIDIA RTX PRO 6000 Blackwell Server Edition Series GPUs
- NVIDIA HGX B300
- NVIDIA BlueField DPUs

About Cisco
Cisco (NASDAQ: CSCO) is the worldwide technology leader that is revolutionizing the way organizations connect and protect in the AI era. For more than 40 years, Cisco has securely connected the world. With its industry leading AI-powered solutions and services, Cisco enables its customers, partners and communities to unlock innovation, enhance productivity and strengthen digital resilience. With purpose at its core, Cisco remains committed to creating a more connected and inclusive future for all. Discover more on The Newsroom and follow us on X at @Cisco.

http://www.cisco.com/go/trademarks. Third-party trademarks mentioned are the property of their respective owners. The use of the word 'partner' does not imply a partnership relationship between Cisco and any other company.

Disclaimer: Many of the products and features mentioned are still in development and will be made available as they are finalized, subject to ongoing evolution in development and innovation. The timeline for their release is subject to change.

View original content to download multimedia:https://www.prnewswire.com/news-releases/cisco-delivers-ai-innovations-across-neocloud-enterprise-and-telecom-with-nvidia-302597375.html

SOURCE Cisco Systems, Inc.