NEWS RELEASE

Cisco to Deliver Secure AI Infrastructure with NVIDIA

2025-03-18

Cisco Secure AI Factory with NVIDIA breaks new ground in AI infrastructure and security while accelerating and simplifying enterprise AI adoption

News Summary:

- Offering will empower customers to build and secure data centers to develop and run AI workloads.
- The Cisco Secure AI Factory with NVIDIA will embed security within all layers, from the application, to the workload, to the infrastructure using solutions like Cisco AI Defense and Hybrid Mesh Firewall.

SAN JOSE, Calif., March 18, 2025 /PRNewswire/ -- **GTC** -- Cisco [NASDAQ: CSCO] today unveiled an AI factory architecture with NVIDIA that puts security at its core. This collaboration with NVIDIA builds on the expanded partnership that was announced last month, and the companies have moved swiftly to provide validated reference architectures today. Together, the companies are developing the Cisco Secure AI Factory with NVIDIA to dramatically simplify how enterprises deploy, manage, and secure AI infrastructure at any scale.

"AI can unlock groundbreaking opportunities for the enterprise," said Chuck Robbins, Chair and CEO, Cisco. "To achieve this, the integration of networking and security is essential. Cisco and NVIDIA's trusted, innovative solutions empower our customers to harness AI's full potential simply and securely."

"AI factories are transforming every industry, and security must be built into every layer to protect data, applications and infrastructure," said Jensen Huang, founder and CEO, NVIDIA. "Together, NVIDIA and Cisco are creating the blueprint for secure AI—giving enterprises the foundation they need to confidently scale AI while safeguarding their most valuable assets."

Developing and delivering AI applications require high performing, scalable infrastructure and AI software tool chain. Securing this infrastructure and AI software requires a new architecture – one that embeds security at all layers of the AI stack and automatically expands and adapts as the underlying infrastructure changes. Cisco and NVIDIA's partnership on the NVIDIA Spectrum-X^{TM} Ethernet networking platform provides the foundation for the Cisco Secure AI Factory with NVIDIA. Cisco is integrating security solutions like Cisco Hypershield, to help protect AI workloads, and Cisco AI Defense, to help protect the development, deployment, and use of AI models and applications. Together, Cisco and NVIDIA will provide customers with the flexibility to design infrastructure for their specific AI needs without sacrificing operational simplicity or security.

Building a Secure AI Factory

AI factories – data centers purpose-built to power AI workloads – are designed to be more modular, scalable and agile, but organizations must also look beyond raw compute power. AI Factories must address new and complex security challenges. The recently published Cisco State of AI Security report analyzes dozens of AI-specific threat vectors and over 700 pieces of AI-related legislation to highlight key developments from a rapidly evolving AI security landscape. Organizations that strategically address both their AI infrastructure and security challenges simultaneously will be more agile, scale faster, and derive business value quicker.

Cisco Secure AI Factory with NVIDIA is expected to build on the companies' unique ability to offer flexible AI networking and full-stack technology options that leverage the planned joint architecture. The partnership will bring together technologies from Cisco, NVIDIA, and our ecosystem partners into a secure AI factory architecture for enterprise customers, including:

- **Compute**: Cisco UCS AI servers based on NVIDIA HGX and NVIDIA MGX for accelerated computing.
- **Networking**: Cisco Nexus Hyperfabric AI and Nexus networking solutions, powered by Silicon One and NVIDIA Spectrum-X Ethernet networking.
- **Storage**: High-performance storage from certified partners Pure Storage, Hitachi Vantara, NetApp, and VAST Data.
- **Software**: NVIDIA AI Enterprise software platform to streamline the development and deployment of production-grade agentic AI workloads.

The Cisco Secure AI Factory with NVIDIA includes security at all layers:

- **Securing the infrastructure:** Cisco Hybrid Mesh Firewall provides unified security management and consistent policy across multiple enforcement points, including network switches, traditional firewalls, and workload agents. This integrated approach ensures pervasive and consistent security, ranging from deep packet inspection to wide infrastructure coverage, detecting, blocking and containing adversaries. Cisco Hypershield (part of Hybrid Mesh Firewall) will, in the future, extend pervasive, zero-trust security enforcement to every AI node by integrating with NVIDIA BlueField-3 DPUs.
- **Securing the Workload:** Cisco Hypershield prevents adversary lateral movement and proactive vulnerability mitigation without the need for patching, all from a single management interface. By monitoring and controlling process executions, file access, and network activities, Hypershield delivers deep visibility and surgical runtime enforcement within AI workloads. Future enhancements will further strengthen workload protection through integration with NVIDIA BlueField-3's DOCA AppShield for real-time workload threat detection in AI-focused virtual machines and containers.
- **Securing the AI application:** Cisco AI Defense empowers security and AI teams with comprehensive tools to protect AI applications from safety (e.g. off-policy, toxic behavior) and security (i.e. prompt injection, data privacy) risks across the development lifecycle. AI Defense

integrates into existing CI/CD workflows to provide automated vulnerability testing and a common layer of runtime security across any number of models and applications. Additionally, AI Defense helps companies align to AI security standards with a single integration, including NIST, MITRE ATLAS, and OWASP LLM Top 10. Future enhancements include integration with NVIDIA AI Enterprise to streamline AI security workflows.

Cisco and NVIDIA each bring a unique understanding of customer AI infrastructure needs, and by combining their insights, can offer flexible deployment models alongside proven reference architectures. The Secure AI Factory will provide enterprise customers with scalable, high-performance AI infrastructure that supports customers at any stage of their journey and embeds security throughout.

Cisco Secure AI Factory with NVIDIA will have flexible deployment options, including:

- **Ready-to-deploy:** Utilizing Cisco Nexus Hyperfabric AI along with Cisco's security portfolio and NVIDIA technology, customers can deploy a vertically integrated AI solution that automates and simplifies the secure AI factory lifecycle from design to deployment and ongoing monitoring.
- **Build-your-own:** Featuring customizable modular components from Cisco, NVIDIA, and the companies' storage ecosystem partners, customers can incorporate their current infrastructure and build solutions that are designed precisely for their unique environments.

"In today's fast-moving market, businesses need more than just technology—they need end-to-end solutions that address their most pressing challenges. I see Cisco and NVIDIA combining their strengths to deliver integrated solutions that I believe will drive innovation, simplify deployment, and streamline operations," said Patrick Moorhead, Founder, CEO and Chief Analyst, Moor Insights & Strategy. "AI isn't easy but the combination of the two could be an 'easy button' for AI infrastructure. By making AI infrastructure easier to adopt and manage, they could empower enterprises to accelerate digital transformation and achieve their strategic goals with more confidence."

Cisco and NVIDIA: The journey to a validated and unified architecture

Moving quickly is crucial to meet today's demand for AI infrastructure, and Cisco and NVIDIA have made progress as part of the collaboration announced in February 2025. Cisco has developed new reference architectures with deployment options for Cisco Nexus Hyperfabric AI or Cisco Nexus 9000 Series Switches validated and based on the NVIDIA Enterprise Reference Architecture for HGX H200 and Spectrum-X.

AVAILABILITY

Solutions based on the Cisco Secure AI Factory with NVIDIA architecture are expected to be available for purchase before the end of calendar year 2025. Many of the individual technology components included in the architecture are available today.

ADDITIONAL RESOURCES

- Executive Blog Post: Embracing the AI Era: Cisco Secure AI Factory with NVIDIA, Jeetu Patel, Cisco's Executive Vice President and Chief Product Officer
- For more information on the Cisco Secure AI Factory with NVIDIA, click here.

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.

Products and features described in this release that are not currently available remain in varying stages of development and will be offered on a when-and-if-available basis. The delivery timeline of any future products and features is subject to change at the discretion of Cisco and its partners.

View original content to download multimedia:https://www.prnewswire.com/news-releases/cisco-to-deliver-secure-ai-infrastructure-with-nvidia-302404977.html

SOURCE Cisco Systems, Inc.