Cisco Unveils Plug-and-Play AI Solutions, Accelerating AI Adoption for the Enterprise

2024-10-29

New solutions with NVIDIA are highly adaptive and scalable for the evolving AI demands of enterprises

News Summary:

- Cisco introduces plug-and-play use case and industry-specific AI PODs, making it easier for partners to sell and customers to deploy AI infrastructure
- Company also expands UCS portfolio with new AI server family for NVIDIA accelerated computing
- With an upgraded portfolio, powered by new 800G switches, Cisco highlights strategy to accelerate enterprise AI adoption by offering open, scalable and simplified path to AI infrastructure

LOS ANGELES, Oct. 29, 2024 /PRNewswire/ -- **CISCO PARTNER SUMMIT** -- Today, Cisco (NASDAQ: CSCO) announced new additions to its data center infrastructure portfolio: an AI server family purpose-built for GPU-intensive AI workloads with NVIDIA accelerated computing, and AI PODs to simplify and de-risk AI infrastructure investment. They give organizations an adaptable and scalable path to AI, supported by Cisco's industry-leading networking capabilities.

"Enterprise customers are under pressure to deploy AI workloads, especially as we move toward agentic workflows, and AI begins solving problems on its own," said Jeetu Patel, Chief Product Officer, Cisco. "Cisco innovations like AI PODs and the GPU server strengthen the security, compliance, and processing power of those workloads as customers navigate their AI journeys from inferencing to training."

The exponential growth of AI is transforming data center requirements, driving demand for scalable, sustainable, programmable and secure networks. According to McKinsey, generative AI will add $2.6T to $4.4T per year to global economic output with enterprises at the forefront of value creation. But according to the Cisco AI Readiness Index, 89% of IT professionals plan to deploy AI workloads within the next two years but just 14% of organizations report their infrastructure is ready for AI today.

1

Cisco Breaks Down Barriers to AI Adoption
Cisco's new solutions provide customers with the infrastructure pieces they need to accelerate their AI adoption, no matter their current starting point. These innovations extend customers' existing infrastructure, enabling customers to grow and innovate without adding complexity. The new solutions are managed by Cisco Intersight, which enables centralized control and automation, simplifying everything from configuration to day-to-day operations. New solutions introduced today include:

- **Accelerated Compute for the AI Era:** Cisco is adding to its UCS AI compute portfolio with the new UCS C885A M8 servers purpose-built for GPU-intensive AI workloads. These high-density servers can tackle the most demanding AI training and inference workloads by harnessing the power of the NVIDIA HGX supercomputing platform with NVIDIA H100 and H200 Tensor Core GPUs. Each server includes NVIDIA NICs or SuperNICs to accelerate AI networking performance, as well as NVIDIA BlueField-3 DPUs to accelerate GPU access to data and enable robust, zero-trust security. This is Cisco's first entry into its dedicated AI server portfolio and its first eight-way accelerated computing system built on the NVIDIA HGX platform.
- **Plug-and-Play AI Infrastructure**: Cisco is introducing AI PODs, infrastructure stacks tailored for specific AI use cases and industries. Combining compute, networking, storage, and cloud management, these stacks enable greater scalability and efficiency. Built on the foundation of Cisco Validated Designs (CVDs), the PODs provide customers with an established starting point, easily adaptable to meet their specific needs. The pre-sized and configured bundles of infrastructure eliminate the guesswork from deploying AI inference solutions – from edge inferencing to large-scale clusters with NVIDIA accelerated computing. These solutions include NVIDIA AI Enterprise, an end-to-end, cloud-native software platform that accelerates data science pipelines and streamlines AI development and deployment. This means faster time to value, consistent performance and reduced risk for AI projects.

These new solutions join Cisco's already extensive portfolio of AI and data center infrastructure, including Cisco's recently introduced 800G Nexus switching platforms powered by the Cisco Silicon One G200 chip, and the recently announced Cisco Nexus HyperFabric AI solution with NVIDIA.

Building an AI Ecosystem with Partners
As Cisco's customers and partners continue to navigate an ever-changing and progressing AI market, Cisco is committed to working across the industry to drive innovation and growth. By leaning on the full power of Cisco's networking, security and observability strengths, and honoring customers' unique journeys to AI, Cisco and its partners can deliver true business value to customers.

"A new era of intelligent generative AI applications is driving transformation for enterprises worldwide. With the debut of its dedicated AI server portfolio, AI PODs, and expanding family of systems with full-stack NVIDIA accelerated computing and software, Cisco is advancing the infrastructure enterprises can use as a foundation for growth in the age of AI." - Bob Pette, vice president, Enterprise Platforms at NVIDIA

"AI is set to fundamentally change how we do business, but it requires scaling that only robust infrastructure can provide to support accelerated computing demands. We're excited about Cisco's efforts to make AI easier for our customers to stand up and manage. With these new solutions, Presidio is able to bring customers new solutions that fit within their existing infrastructure, are easy to deploy and manage, and provide the right infrastructure for their AI journeys." - Rob Kim, CTO, Presidio

"Cisco launching these new platforms couldn't come at a better time as many of our customers are working to figure out the role of AI in their business and how they can capitalize on the technology. CDW is excited about Cisco's new AI infrastructure and security platforms to drive business outcomes securely with our joint customers." - Conor Waddell, SVP Integrated Technology Solutions, CDW

"One of the challenges we see for customers is the complexity of the AI infrastructure stack. There are many options and the technology is moving very rapidly. Cisco's solutions like the AI PODs will simplify and accelerate AI adoption for our customers allowing them to focus on the AI outcomes they are seeking for their organizations. At WWT we are excited to work with Cisco to showcase the Cisco AI PODs in our AI Proving Ground Lab and accelerate the AI journey for our mutual customers." - Neil Anderson, VP Cloud, Infrastructure, and AI Solutions, World Wide Technology

Availability

- The Cisco UCS C885A M8 is now orderable and expected to ship to customers by the end of this year
- The Cisco AI PODs will be orderable in November 2024.
- Payment solutions are available for these products to enable partners to lower balance sheet debt, reduce customer credit and payment risk, and improve cash flow and profitability. For more information, please visit cisco.com/go/paymentsolutions.

Additional Resources:

- Executive blog post: Power Your GenAI Ambitions with New Cisco AI-Ready Data Center Infrastructure
- For more information on Cisco Partner Summit, please visit The Newsroom

About Cisco
Cisco (NASDAQ: CSCO) is the worldwide technology leader that securely connects everything to make anything possible. Our purpose is to power an inclusive future for all by helping our customers reimagine their applications, power hybrid work, secure their enterprise, transform their infrastructure, and meet their sustainability goals.

Discover more on The Newsroom and follow us on X (formerly Twitter) at @Cisco.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. A listing of Cisco's trademarks can be found at 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.

Forward-Looking Statements: This release may be deemed to contain forward-looking statements, which are subject to the safe harbor provisions of the Private Securities Litigation Reform Act of 1995. These statements include, among others, statements regarding Cisco's future strategic investment plans to expand and develop secure and reliable AI solutions. Readers are cautioned that these forward-looking statements are only predictions and may differ materially from actual future events or results due to a variety of factors, including, among other things, the ability of Cisco to achieve expected benefits of its investments, business and economic conditions and growth trends in its customer markets and various geographic regions, increased competition, global economic conditions and uncertainties in the geopolitical environment and other risk factors set forth in Cisco's most recent reports on Form 10-K and 10-Q, respectively. Any forward-looking statements in this release are based on limited information currently available to Cisco, which is subject to change, and Cisco will not necessarily update the information.

View original content to download multimedia:https://www.prnewswire.com/news-releases/cisco-unveils-plug-and-play-ai-solutions-accelerating-ai-adoption-for-the-enterprise-302290198.html

SOURCE Cisco Systems, Inc.