



Nutanix Builds a Faster Foundation for Production AI

October 21, 2026

Nutanix Is Expanding Its AI Ecosystem with NVIDIA, Supermicro, and Netris to Help Enterprises and Service Providers Quickly Deploy and Scale Production AI, and Reduce AI Costs

News Summary

- **[Powering Agentic AI with NVIDIA Vera CPU](#)**: Nutanix plans to work with NVIDIA toward Day Zero support for Vera, combining its CPU capabilities with Nutanix software. NVIDIA's high-performance CPU architecture, paired with Nutanix Kubernetes Platform (NKP) on bare metal are expected to provide a predictable, low-latency foundation for running advanced AI agents.
- **[Accelerating AI Deployment and AI Cost Control with Supermicro](#)**: Nutanix and Supermicro are announcing plans for a jointly engineered system combining standardized Supermicro servers with Nutanix's Agentic AI full stack software, aiming to help customers simplify and accelerate deployment of private inference for increased security and lower token costs.
- **[Multitenant AI Infrastructure on Bare Metal and Virtualized Environments with Netris](#)**: The Nutanix Cloud Platform and Netris Network Automation, Abstraction, and Multitenancy (NAAM) platform will simplify multitenant AI infrastructure with network automation, hard tenant isolation, and self-service network operations across bare-metal and virtualized environments. The integrated solution will enable AI cloud providers to deliver GPU infrastructure as a service through a more unified, cloud-like experience.

Why it Matters

As demand for production AI grows, organizations are under pressure to deploy increasingly complex infrastructure faster and turn AI investments into real business value. Nutanix is working across a broad ecosystem of infrastructure and silicon partners to simplify that path, planning to help enterprises bring production AI and agentic workloads online more quickly while enabling neoclouds and service providers to scale the multitenant AI services they deliver to customers.

Together, these advancements are intended to combine Nutanix software with leading infrastructure technologies to help reduce deployment complexity, accelerate time to value, reduce AI costs, and provide a consistent foundation for operating AI at scale.

BERLIN, Oct. 21, 2026 (GLOBE NEWSWIRE) -- [NVIDIA GTC](#) -- [Nutanix](#) (NASDAQ: [NTNX](#)), a hybrid cloud leader and AI platform company, today announced it is collaborating with NVIDIA, Supermicro, and Netris to deliver new AI solutions that aim to simplify the infrastructure required for production AI and help organizations bring new capabilities online faster. By combining Nutanix software with leading hardware, networking, and silicon technologies, Nutanix aims to help enterprises deploy and operate production AI environments more efficiently at scale.

As organizations push more AI workloads into production, capacity constraints and operational complexity can make it harder to move quickly. Nutanix and its partners are working to simplify that path to make it easier to bring AI into production faster and accelerate time to value.

"Organizations need more than just advanced infrastructure to move agentic AI into production – we believe they need an AI cloud that simply connects intelligence, agents, enterprise applications, and data across CPUs, GPUs, virtual machines, containers, on-premises environments, and cloud," said **Thomas Cornely, Executive Vice President, Product Management at Nutanix**. "Through our collaboration with NVIDIA, Supermicro, and Netris, we plan to help customers accelerate deployments of agents and AI services on a unified platform with built-in security, performance, and governance."

Extending Agentic AI to the CPU Layer with NVIDIA Vera

As enterprises adopt agentic AI at scale, they are anticipated to require infrastructure that can support increasingly demanding workloads across the full AI environment, not just GPU-intensive tasks. Agentic applications combine inference with tool calls, code execution, retrieval, orchestration, and outcome processing. As these workflows scale, CPU performance is expected to become essential for completing the steps between GPU operations, reducing delays, and maximizing overall AI-factory efficiency. To help ensure this critical compute path does not become a bottleneck, Nutanix plans to work towards Day Zero support for NVIDIA Vera CPUs, beginning with NKP Bare Metal, to serve highly performance-sensitive AI factory environments, with planned expansion to virtualized Nutanix environments.

By pairing NVIDIA's purpose-built CPU architecture – engineered for high-speed agent execution and reduced latency – and [NVIDIA OpenShell](#) to create safe, private runtime environments for agents with NKP, the collaboration aims to give organizations a predictable, high-performance foundation for their most advanced AI agents.

This support is expected to give organizations greater flexibility in how they safely deploy agentic AI while simplifying operations with a consistent cloud operating model.

Nutanix and Supermicro Take on Runaway AI Costs with Jointly Engineered System for Private Inference

Nutanix and Supermicro are announcing plans for a jointly engineered AI infrastructure solution built around validated Supermicro server systems. By combining Supermicro servers and NVIDIA GPUs with Nutanix Agentic AI (NAI) and [NVIDIA Spectrum-x™ Ethernet](#) networking, the solution is being

designed to help reduce custom configuration and integration delays, with the goal of helping enterprises get on-premises GPU capacity and take on runaway AI costs with a jointly engineered system powered by NAI for private inference.

As AI requirements grow, the architecture is expected to be able to scale GPU capacity to meet evolving requirements. Nutanix provides an AI cloud control plane and AI storage purpose-built to support workloads with resilient operations and high-throughput data workflows. Customers are anticipated to gain a predictable path to deploy, manage, and scale AI inference workloads without assembling and integrating a complex, multivendor environment.

Nutanix and Netris: Making Multitenant AI Infrastructure Easier to Deploy and Operate

For NVIDIA Cloud Partner reference architectures, multitenancy is designed to extend across AI infrastructure, including the physical network. In dynamic neocloud environments, operators also must be able to automate network access and ensure tenant isolation as customers, workloads, and GPU allocations change. Nutanix and Netris are collaborating to automate this network-layer multitenancy and turn complex GPU infrastructure into secure, cloud-like, multitenant AI services.

Nutanix will provide the foundation for compute across virtual machines, containers, CPUs, and GPUs, along with storage, security, Kubernetes, and Nutanix Agent Gateway, model-as-a-service and AI platform services. Netris will provide network automation, abstraction, and multitenancy across the physical network fabric, with hard isolation (enforced on networking hardware) from compute through the physical network. Together, the integrated solution will provide a cloud-like self-service operating experience across bare-metal and virtualized AI infrastructure, helping providers instantly onboard tenants and deliver AI services more efficiently.

Analyst Commentary

"Nutanix built its enterprise business by simplifying virtualized infrastructure and more recently containerized workloads," said **Don Gentile, Analyst, Data Platforms and Resiliency, HyperFRAME Research**. "Now it's extending that expertise to bare-metal AI systems, agent execution, private inference, and multitenant cloud services. Through innovation and targeted partnerships, Nutanix is advancing a common cloud operating model that enterprises and service providers have long wanted: consistent management, automation, and choice across increasingly diverse infrastructure."

Ecosystem Commentary

"AI agents need fast CPUs to execute code, use tools and process data between model calls," said **Hannah Coutand, Director of Product Marketing for Vera CPU at NVIDIA**. "NVIDIA Vera is purpose-built for this work, and our collaboration with Nutanix will help enterprises bring high-performance agentic AI into production with a consistent cloud operating model."

Customer Commentary

"Digital and AI transformation is a fundamental pillar supporting the Kingdom's ambitions to build cutting-edge, resilient, agile, and data-driven public services. Through our collaboration with Nutanix, we look forward to exploring how Nutanix's platform for hybrid cloud computing across multiple environments can support the adoption of AI technologies and help develop national capabilities in advanced cloud and AI technologies."

– Ministry of Environment, Water and Agriculture (MEWA) - Kingdom of Saudi Arabia

About Nutanix

Nutanix is a hybrid cloud leader and AI platform company, offering organizations a unified infrastructure software platform to safely run applications, data, and AI anywhere. Trusted by customers worldwide, Nutanix empowers more than 50% of the Global 2000 to innovate faster with AI, while modernizing infrastructure, simplifying operations, and controlling costs. Learn more at www.nutanix.com or follow us on social media.

©2026 Nutanix, Inc. All rights reserved. Nutanix, the Nutanix logo and all Nutanix product and service names mentioned are registered trademarks or trademarks of Nutanix, Inc. in the United States and other countries. Kubernetes is a registered trademark of The Linux Foundation in the United States and other countries. All other brand names mentioned are for identification purposes only and may be the trademarks of their respective holder(s).

This content may contain express and implied forward-looking statements, including but not limited to statements regarding our plans and expectations relating to new product features and technology under development, the capabilities of such product features and technology, and our plans to release product features and technology. Such statements are not historical facts and are instead based on our current expectations, estimates and beliefs, including statements about our plans to collaborate with NVIDIA, Supermicro, and Netris to deliver new AI solutions and the anticipated capabilities of these offerings, anticipated integrations, collaboration activities, customer benefits, and implementation of contemplated arrangements relating to such offerings. The accuracy of such statements involves risks and uncertainties and depends upon future events, including those that may be beyond our control, and actual results may differ materially and adversely from those anticipated or implied by such statements, including, among others: failure to develop, or unexpected difficulties, delays or disruptions in developing, releasing or distributing, new products, services, product features or technology in a timely or cost-effective basis, failure to complete anticipated integrations or collaboration activities, failure to successfully implement contemplated arrangements relating to such offerings, or failure to achieve the anticipated benefits of such offerings. Any forward-looking statements included speak only as of the date hereof and, except as required by law, we assume no obligation to update or otherwise revise any such forward-looking statements to reflect subsequent events or circumstances. Certain products and features or functionalities described herein remain in varying stages of development and will be offered on a when-and-if-available basis. The development, release, and timing of any such products, features or functionalities are subject to change. Nutanix will not have any liability arising from reliance on this content for any failure to deliver or delay in the delivery of any such products, features or functionalities. Any future product or product feature information is intended to outline general product directions, and is not a commitment, promise, or legal obligation for Nutanix to deliver any functionality. This information should not be used when making a purchasing decision. Certain collaborations, offerings, and related activities described herein remain subject to ongoing development efforts and, where applicable, related commercial arrangements among the parties.

