



NEWS RELEASE

NetApp and Lenovo Offer Converged Infrastructure Solution Optimized for GenAI

2024-05-14

SAN JOSE, Calif. — May 14, 2024 – NetApp® (NASDAQ: NTAP), the intelligent data infrastructure company, today announced NetApp AI Pod with Lenovo ThinkSystem servers for [NVIDIA OVX](#), a converged infrastructure optimized for the generative AI era. NetApp and Lenovo have had a strong partnership for over five years and are now taking the partnership to the next level with the introduction of this simple, cost-effective, and powerful solution for enterprise AI. The new converged infrastructure solution will support NVIDIA NIM inference microservices, the NVIDIA AI Enterprise software platform for generative AI development and deployment, and retrieval-augmented generation (RAG) to enable customers to leverage their private, proprietary data for AI without large-scale model training.

Generative AI has created extensive opportunities for organizations, though many may feel as if AI is out of reach due to the cost, integration complexity, and time to deployment. Meanwhile, pre-built large language models (LLMs) often do not have training on the most up-to-date or relevant data organizations need to generate valuable insights or effectively automate routine tasks.

NetApp has led in unstructured data storage for over three decades, and much of the world's enterprise unstructured data is stored in NetApp storage. Unlocking the full value of this data requires going far beyond foundational model training. By enabling inferencing and RAG, the NetApp AI Pod with Lenovo's leading [server reliability](#), quality, and performance makes it seamless, simple, and fast to combine the power of LLMs with precise, up-to-date, and pertinent information on NetApp storage.

"The NetApp AI Pod with Lenovo ThinkSystem servers for NVIDIA OVX transforms enterprise AI by delivering a pre-integrated, high-performance solution that accelerates the deployment and scaling of generative AI workloads," said Sandeep Singh, Senior Vice President & General Manager, Enterprise Storage at NetApp. "Through our collaboration with NVIDIA and Lenovo, NetApp is empowering organizations to harness the capabilities of AI more efficiently, ensuring that they can focus on creating value and gaining insights from their data without the complexity of building and managing AI infrastructure from scratch."

NetApp AI Pod combines NetApp's robust storage systems with hybrid cloud data management, including Lenovo's high-performance [ThinkSystem SR675 V3 servers](#), the leading platforms supporting NVIDIA-certified OVX architecture with [NVIDIA L40S GPUs](#), as well as NVIDIA Spectrum-X Networking, providing a complete and seamless infrastructure solution to streamline AI adoption. Using this solution, customers will be able to conduct

AI RAG and inferencing operations for use cases like chatbots, knowledge management, and object recognition. NetApp, in collaboration with Lenovo and NVIDIA, simplifies AI deployment by seamlessly integrating into ecosystems, simplifying operations, accelerating implementation, and safeguarding AI infrastructure, enabling organizations to unlock AI's full potential.

"Lenovo is expanding our partnership with NetApp and NVIDIA to bring the power of AI to every business as part of our continued commitment to enabling and simplifying hybrid AI for our joint customers," said Kirk Skaugen, President of Lenovo Infrastructure Solutions Group. "As customers deploy AI, they demand business critical availability, ease of management, and infrastructure efficiency. The NetApp AIpod with Lenovo ThinkSystem servers for NVIDIA OVX delivers optimized and validated solutions to make generative AI more accessible for businesses of every size."

"Making AI accessible and scalable for every industry is more important than ever," said Bob Pette, Vice President of Enterprise Platforms at NVIDIA. "The NetApp AIpod with Lenovo for NVIDIA OVX offers an AI-ready platform that's powerful and easy to use, enabling businesses to drive innovation."

NetApp AIpod with Lenovo for NVIDIA OVX will be available by summer 2024.

Additional Resources

- [NetApp AIpod™ with Lenovo for NVIDIA OVX: The Power of Retrieval-Augmented Generation for Every Enterprise](#)
- [NetApp Unveils Unified Data Storage Built for the AI Era](#)
- [AIpod: Delivering the Promise](#)

About NetApp

NetApp is the intelligent data infrastructure company, combining unified data storage, integrated data services, and CloudOps solutions to turn a world of disruption into opportunity for every customer. NetApp creates silo-free infrastructure, harnessing observability, and AI to enable the industry's best data management. As the only enterprise-grade storage service natively embedded in the world's biggest clouds, our data storage delivers seamless flexibility. In addition, our data services create a data advantage through superior cyber resilience, governance, and application agility. Our CloudOps solutions provide continuous optimization of performance and efficiency through observability and AI. No matter the data type, workload, or environment, with NetApp you can transform your data infrastructure to realize your business possibilities. Learn more at www.netapp.com or follow us on [X](#), [LinkedIn](#), [Facebook](#), and [Instagram](#).

NETAPP, the NETAPP logo, and the marks listed at www.netapp.com/TM are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.