



Nutanix and NVIDIA Collaborate to Accelerate Enterprise AI Adoption

May 21, 2024

Integration will bring together Nutanix's automated and secure enterprise AI platform with NVIDIA NIM inference microservices for accelerated inference of state-of-the-art AI models

BARCELONA, Spain--(BUSINESS WIRE)--May 21, 2024-- **.NEXT Conference** – [Nutanix](#) (NASDAQ: [NTNX](#)), a leader in hybrid multicloud computing, today announced a collaboration with NVIDIA aimed at helping enterprises more easily adopt generative AI (GenAI). Through the integration of [NVIDIA NIM](#) inference microservices with Nutanix GPT-in-a-Box 2.0, customers will be able to build scalable, secure, high-performance GenAI applications across the enterprise and at the edge.

Today, most AI innovation is centered on the public cloud due to access to infrastructure and tooling able to support the needs of AI applications. Additionally, only the largest enterprises with teams of data scientists have made progress in GenAI adoption. However, most enterprises are looking to invest in supporting their AI strategy, including boosting their investment at the edge, according to the [State of Enterprise AI](#). What's missing is a fast-track for organizations to mainstream GenAI beyond the public cloud, across the enterprise, and at the edge.

Nutanix's integration of NVIDIA NIM microservices will enable its customers to leverage Nutanix GPT-in-a-Box 2.0, built on top of the company's rich data services and compute platform, and use it to simplify AI model deployment and more effectively and efficiently run enterprise AI/ML applications. This will expand access to the growing catalog of NVIDIA NIM microservices from across the enterprise and at the edge, helping to fast-track GenAI initiatives without requiring a team of data scientists.

Nutanix's collaboration with NVIDIA helps simplify the experience, which many enterprises find challenging today, of making all the decisions required to stand up AI solutions. These include choosing among hundreds of thousands of models, serving engines, and supporting infrastructure, while lacking the new skill sets needed to deliver GenAI solutions to their customers.

Nutanix GPT-in-a-Box simplifies building an AI-ready stack, integrated with Nutanix Objects and Nutanix Files for model and data storage, enabling customers to maintain control over their data. New features delivered in GPT-in-a-Box 2.0 will also automate deploying and running inference endpoints for a wide range of AI models and secure access to the model using fine-grained access control and auditing.

Running on top of the Nutanix Cloud Platform, NIM microservices will enable seamless AI inferencing on a wide range of models, including open-source community models, [NVIDIA AI Foundation](#) models, and custom models, leveraging industry-standard application programming interfaces. To support the integration, Nutanix also announced certification for the [NVIDIA AI Enterprise 5.0](#) software platform for streamlining the development and deployment of production-grade AI, including NVIDIA NIM.

"Enterprises are looking to simplify GenAI adoption, and Nutanix enables customers to move to production more easily while maintaining control, privacy, and cost," said Tarkan Maner, Chief Commercial Officer at Nutanix. "This collaboration will add to this value by making it even easier for customers to leverage NVIDIA's latest innovation with NIM."

"Across every industry, enterprises are working to efficiently integrate AI into the cloud and data platforms that power their operations," said Manuvir Das, Vice President of Enterprise Computing at NVIDIA. "The integration of NVIDIA NIM into Nutanix GPT-in-a-Box gives enterprises an AI-ready solution for rapidly deploying optimized models in production."

Nutanix GPT-in-a-Box 2.0 is expected to be available in the second half of 2024. More information can be found [here](#).

About Nutanix

Nutanix is a global leader in cloud software, offering organizations a single platform for running apps and data across clouds. With Nutanix, companies can reduce complexity and simplify operations, freeing them to focus on their business outcomes. Building on its legacy as the pioneer of hyperconverged infrastructure, Nutanix is trusted by companies worldwide to power hybrid multicloud environments consistently, simply, and cost-effectively. Learn more at www.nutanix.com or follow us on social media @nutanix.

© 2024 Nutanix, Inc. All rights reserved. Nutanix, the Nutanix logo, and all Nutanix product and service names mentioned herein are registered trademarks or unregistered trademarks of Nutanix, Inc. ("Nutanix") in the United States and other countries. Other brand names or marks mentioned herein are for identification purposes only and may be the trademarks of their respective holder(s). This press release is for informational purposes only and nothing herein constitutes a warranty or other binding commitment by Nutanix. This release may contain express and implied forward-looking statements, which are not historical facts and are instead based on Nutanix's current expectations, estimates and beliefs, including statements about: the benefits and capabilities of our platform, products, services and technology; and our plans and expectations, including the timing of any product releases or upgrades or announcements, regarding new products, services, product features and technology that are under development or in process, including with respect to GPT-in-a-Box 2.0 and integrations with third-party products and services. The accuracy of such statements involves risks and uncertainties and depends upon future events, including those that may be beyond Nutanix's 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. Any forward-looking statements included herein speak only as of the date hereof and, except as required by law, Nutanix assumes no obligation to update or otherwise revise any of such forward-looking statements to reflect subsequent events or circumstances. Certain products and features or functionalities described herein, including with respect to GPT-in-a-Box 2.0 and integrations with third-party products and services, 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 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.

View source version on [businesswire.com](https://www.businesswire.com/news/home/20240521516262/en/): <https://www.businesswire.com/news/home/20240521516262/en/>

Gabrielle Moynan
pr@nutanix.com

Source: Nutanix