NVIDIA and Google Cloud to Create Industry's First AI-on-5G Lab to Speed Development of AI Everywhere

Joint Innovation Lab to Support Vendors in Conceiving, Testing, Optimizing, and Validating 5G and AI Applications Running On-Prem and Across Google Cloud

Mobile World Congress—NVIDIA today announced that it is partnering with Google Cloud to establish the industry's first AI-on-5G Innovation Lab, enabling network infrastructure players and AI software partners to develop, test and adopt solutions that will help accelerate the creation of smart cities, smart factories and other advanced 5G and AI applications.

The lab will provide enterprises with access to Google Cloud's Anthos platform and NVIDIA accelerated computing hardware and software platforms that let them harness data and AI to drive business performance, improve operational efficiency and optimize safety and reliability. The companies will begin development in the second half of the year.

"The global telecommunications industry is undergoing a sweeping transformation as the speed and low latency of 5G create unprecedented opportunities to deliver AI-on-5G at the enterprise edge," said Ronnie Vasishta, senior vice president of Telecom at NVIDIA. "Our collaboration with Google Cloud will help network operators and infrastructure players conceive and create new profit centers built on AI and machine learning."

Growing AI-on-5G Ecosystem
NVIDIA announced in April it is teaming with Google Cloud, Fujitsu, Mavenir, Radisys and Wind River to develop solutions for NVIDIA's AI-on-5G platform. Google Cloud is extending the Anthos application platform to the network edge, allowing telecommunications service providers and enterprises to rapidly deliver new services and applications at the 5G edge.

"Google Cloud, with our Anthos application platform, is enabling service providers and enterprises across many vertical markets to connect seamlessly over 5G at the network edge," said Shailesh Shukla, vice president and general manager for Networking at Google Cloud. "We're excited to extend our collaboration with NVIDIA to deliver a joint innovation lab that provides the infrastructure required for enterprises, communications service providers and partners to develop, test and deploy their edge AI applications."

Anthos offers a consistent platform for application deployments, with a service-centric view of each environment. The platform enables customers to build and deploy enterprise-grade, containerized applications faster with managed Kubernetes in the cloud, on premises and at the network edge.

Anthos supports NVIDIA GPU-accelerated servers, enabling a consistent deployment and operational experience across deployments, while reducing expensive overhead and improving developer productivity. The platform also protects applications and software supply chains and provides an outcome-focused approach to managing policies for applications across environments — a critical feature for network operators and enterprises delivering AI on 5G.

NVIDIA Hardware and Software Development
NVIDIA AI-on-5G provides enterprises, mobile network operators and cloud service providers a single, converged platform for handling both 5G and edge AI computing. The platform uses the NVIDIA Aerial™ software development kit with the NVIDIA BlueField®-2 A100 — a converged card that combines GPUs and DPUs as well as NVIDIA's "5T for 5G" solution. This allows for the creation of high-performance 5G RAN and AI applications to manage precision manufacturing robots, automated guided vehicles, drones, wireless cameras, self-checkout aisles and hundreds of other transformational projects.

The NVIDIA Aerial SDK, in combination with the AI frameworks NVIDIA Metropolis for smart cities, NVIDIA Isaac™ for autonomous machines and NVIDIA Clara™ for healthcare, is an integral part of the AI-on-5G ecosystem and can be deployed on a single NVIDIA-Certified System™.

In related news today, NVIDIA announced its next-generation Aerial A100 AI-on-5G computing platform will incorporate 16 Arm-based CPU cores into the NVIDIA BlueField-3 A100. This results in a self-contained, converged card that delivers enterprise edge AI applications over cloud-native 5G vRAN with improved performance per watt and faster time to deployment.

These solutions will be available for testing and validation in the AI-on-5G lab.

NVIDIA AI-on-5G Data Centers — From Edge to the Cloud
Software-defined RANs are critical for building a modern 5G infrastructure that is capable of running a range of applications on a common platform. NVIDIA Aerial enables the best possible utilization by providing elasticity as network traffic changes throughout the day and the flexibility to offer services based on dynamic customer needs.

The NVIDIA EGX™ platform offered by OEMs through NVIDIA-Certified Systems brings AI computing capabilities to the edge, where data gets created. Expanding on the platform, server makers also can pair NVIDIA GPUs and DPUs to build hyperconverged edge data centers.

About NVIDIA
NVIDIA's (NASDAQ: NVDA) invention of the GPU in 1999 sparked the growth of the PC gaming market and has redefined modern computer graphics, high performance computing and artificial intelligence. The company's pioneering work in accelerated computing and AI is reshaping trillion-dollar industries, such as transportation, healthcare and manufacturing, and fueling the growth of many others. More information at https://nvidianews.nvidia.com/.

Certain statements in this press release including, but not limited to, statements as to: the benefits, impact, performance, features, and availability of our products and services; NVIDIA partnering with Google Cloud to establish the industry's first AI-on-5G Innovation Lab; the benefits, impact, performance, features, and availability of our products and services; the benefits, impact, features and timing for development of the AI-on-5G Innovation Lab; NVIDIA teaming with Google Cloud, Fujitsu, Mavenir, Radisys and Wind River to develop solutions for NVIDIA's AI-on-5G platform; Google Cloud extending the Anthos application platform to the network edge; and NVIDIA solutions being available for testing and validation in the lab are forward-looking statements that are subject to risks and uncertainties that could cause results to be materially different than expectations. Important factors that could cause actual results to differ materially include: global economic conditions; our reliance on third parties to manufacture, assemble, package and test our products; the impact of technological development and competition; development of new products and technologies or enhancements to our existing product and technologies; market acceptance of our products or our partners' products; design, manufacturing or software defects; changes in consumer preferences or demands; changes in industry standards and interfaces; unexpected loss of performance of our products or technologies when integrated into systems; as well as other factors detailed from time to time in the most recent reports NVIDIA files with the Securities and Exchange Commission, or SEC, including, but not limited to, its annual report on Form 10-K and quarterly reports on Form 10-Q. Copies of reports filed with the SEC are posted on the company's website and are available from NVIDIA without charge. These forward-looking statements are not guarantees of future performance and speak only as of the date hereof, and, except as required by law, NVIDIA disclaims any obligation to update these forward-looking statements to reflect future events or circumstances.

Cliff Edwards
NVIDIA Corporation
+1-415-699-2755
cliffe@nvidia.com