NVIDIA AI-on-5G Computing Platform Adopted by Leading Service and Network Infrastructure Providers

Fujitsu, Google Cloud, Mavenir, Radisys and Wind River to Deliver Solutions for Smart Hospitals, Factories, Warehouses and Stores

GTC -- NVIDIA today announced it is teaming with Fujitsu, Google Cloud, Mavenir, Radisys and Wind River to develop solutions for NVIDIA's AI-on-5G platform, which will speed the creation of smart cities and factories, advanced hospitals and intelligent stores.

Enterprises, mobile network operators and cloud service providers that deploy the platform will be able to handle both 5G and edge AI computing in a single, converged platform. The AI-on-5G platform leverages the NVIDIA Aerial™ software development kit with the NVIDIA BlueField®-2 A100 — a converged card that combines GPUs and DPUs including NVIDIA's "5T for 5G" solution. This creates high-performance 5G RAN and AI applications in an optimal platform to manage precision manufacturing robots, automated guided vehicles, drones, wireless cameras, self-checkout aisles and hundreds of other transformational projects.

"In this era of continuous, accelerated computing, network operators are looking to take advantage of the security, low latency and ability to connect hundreds of devices on one node to deliver the widest range of services in ways that are cost-effective, flexible and efficient," said Ronnie Vasishta, senior vice president of Telecom at NVIDIA. "With the support of key players in the 5G industry, we're delivering on the promise of AI everywhere."

NVIDIA and several collaborators in this new AI-on-5G ecosystem are members of the O-RAN Alliance, which is developing standards for more intelligent, open, virtualized and fully interoperable mobile networks. Such collaboration allows operators to use the same computing infrastructure required for 5G networking to provide AI services in enterprise, industrial and consumer and residential settings.

AI-on-5G will benefit multiple industry verticals. IHS Markit expects this 5G-enabled value chain to generate $13.1 trillion in increased productivity by 2035. The NVIDIA Aerial SDK, in combination with NVIDIA Metropolis, NVIDIA Isaac™ and NVIDIA Clara™, is an integral part of the AI-on-5G ecosystem and can be deployed on a single NVIDIA-Certified System™ using NVIDIA GPUs and DPUs on a single card.

Solutions for Public and Private Networks
NVIDIA's collaboration with others working to enable AI-on-5G includes:

Mavenir, which is building two 5G vRAN systems based on the Aerial SDK, and will target the network operator segment for public 5G and for enterprise AI with private 5G. Mavenir and NVIDIA have created a hyperconverged enterprise 5G solution to enable enterprises to implement AI-on-5G applications in a seamless and easy-to-use solution.

"The partnership between Mavenir and NVIDIA allows both organizations the opportunity to offer innovative AI solutions that truly deliver on the potential of 5G. The joint hyperconverged enterprise AI solutions from Mavenir and NVIDIA will enable industries from healthcare to retail to Industry 4.0 to take advantage of the endless possibilities 5G has to offer," said Kuntal Chowdhury, senior vice president and general manager of AI and Analytics at Mavenir.

Fujitsu later this year plans to deliver a 5G Open RAN system for verification starting with Sub6 band. Upon the system's completion, Fujitsu and NVIDIA will be helping NTT DOCOMO and global operators with their evolution toward Open RAN in 5G, and beyond. Aerial software-defined implementation reduces time to market, speeds innovation and helps deliver AI applications to enterprises.

"Fujitsu is committed to developing Open RAN solutions for global operators based on NVIDIA's Aerial AI-on-5G and its accelerated GPU and DPU solutions," said Masaki Taniguchi, senior vice president and head of the Mobile System Business Unit at Fujitsu. "Network operators, including NTT DOCOMO, are asking for hyperconverged and software-defined 5G vRAN systems to deliver innovative solutions to their customers."

Enterprise Deployments
Seeking to solve the challenge of deploying AI in enterprises, Radisys and Wind River plan to deliver NVIDIA Aerial AI-on-5G solutions for enterprise 5G and industrial 5G networks. Over 6 million 5G cells will be deployed by 2027 to smart factories, fulfillment centers and other enterprise, industrial and public zones to provide localized connectivity solutions, according to ABI Research.

"Radisys is pleased to partner with NVIDIA on our 5G vRAN and edge computing portfolio, providing end-to-end system integration and helping industries bring AI to life," said Ganesh Shenbagaraman, head of Integrated Products and

Ecosystems at Radisys.

Google Cloud is extending the Anthos application platform to the network edge, allowing telecommunications service providers and enterprises to enable the rapid delivery of new services and applications at the 5G edge. Enterprises can turn to Google Cloud's managed services and NVIDIA for their Internet of Things economy, and to harness data and AI to drive business performance, improve operational efficiency and optimize safety and reliability.

"AI-on-5G is transformative. Google Cloud's industry-specific AI solutions meet scalable vertical needs," said Shailesh Shukla, vice president and general manager for Networking at Google Cloud. "With the power of 5G, Google's AI offering increases exponentially. We are excited to expand our work with NVIDIA to deliver AI and 5G computing at the edge with Anthos and NVIDIA's accelerated edge technologies."

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 solution 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 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 brings AI computing capabilities to the edge where data gets created. The NVIDIA EGX stack, which is compatible with all commercially available Kubernetes infrastructures, is an ideal platform for Aerial enabling low-power, always-on and high-performance devices, reshaping the telecom industry. 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 impact and benefits of the NVIDIA AI-on-5G GPU-accelerated platform in collaboration with network infrastructure providers, the NVIDIA Aerial SDK with BlueField-2 A100 and the NVIDIA EGX platform; NVIDIA delivering on the promise of AI everywhere; AI-on-5G benefiting multiple industry verticals; the amount of increased productivity the 5G-enabled value chain is expected to generate by 2035; the joint hyperconverged enterprise AI solutions from Mavenir enabling industries to take advantage of the possibilities 5G has to offer; Fujitsu's plans to deliver a 5G Open RAN system for verification; the number of 5G cells that will be deployed by 2027; the impact of Google Cloud extending the Anthos application platform to the network edge; Google's AI offering increasing exponentially with 5G; and the impact and benefits of Anthos' platform for application deployments 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