NVIDIA Sets Path for Future of Edge AI and Autonomous Machines With New Jetson AGX Orin Robotics Computer

NVIDIA Ampere Architecture Comes to Jetson, Delivering up to 200 TOPS and 6x Performance Gain

GTC—NVIDIA today introduced NVIDIA Jetson AGX Orin™, the world's smallest, most powerful and energy-efficient AI supercomputer for robotics, autonomous machines, medical devices and other forms of embedded computing at the edge.

Built on the NVIDIA Ampere architecture, Jetson AGX Orin provides 6x the processing power and maintains form factor and pin compatibility with its predecessor, Jetson AGX Xavier™. It delivers 200 trillion operations per second, similar to that of a GPU-enabled server but in a size that fits in the palm of your hand.

The new Jetson computer accelerates the full NVIDIA AI software stack, allowing developers to deploy the largest, most complex models needed to solve edge AI and robotics challenges in natural language understanding, 3D perception, multisensor fusion and more.

"As robotics and embedded computing transform manufacturing, healthcare, retail, transportation, smart cities and other essential sectors of the economy, the demand for processing continues to surge," said Deepu Talla, vice president and general manager of embedded and edge computing at NVIDIA. "Jetson AGX Orin addresses this need, enabling the 850,000 Jetson developers and over 6,000 companies building commercial products on it to create and deploy autonomous machines and edge AI applications that once seemed impossible."

Jetson AGX Orin features an NVIDIA Ampere architecture GPU and Arm Cortex-A78AE CPUs, along with next-generation deep learning and vision accelerators. High-speed interfaces, faster memory bandwidth and multimodal sensor support provide the ability to feed multiple concurrent AI application pipelines.

Comprehensive Software and Ecosystem Support
Customers using the Jetson AGX Orin can leverage the NVIDIA CUDA-X™ accelerated computing stack, NVIDIA JetPack™ SDK and the latest NVIDIA tools for application development and optimization, including cloud-native development workflows. Pretrained models from the NVIDIA NGC™ catalog are optimized and ready for fine-tuning with the NVIDIA TAO toolkit and customer datasets. This reduces time and cost for production-quality AI deployments, while cloud-native technologies allow seamless updates throughout a product's lifetime.

The Jetson embedded computing partner ecosystem encompasses a broad range of services and products, including cameras and other multimodal sensors, carrier boards, hardware design services, AI and system software, developer tools and custom software development.

"Driven by the advancements in AI, the robotics market is growing exponentially," said Jim McGregor, principal analyst at TIRIAS Research. "NVIDIA is the recognized leader in AI and continues to leverage this expertise to advance robotics through a robust ecosystem and complete end-to-end solutions, including a range of hardware platforms that leverage common tools and neural network models. The new Jetson platform brings the performance and versatility of the NVIDIA Ampere architecture to enable even further advancements in autonomous mobile robots for a wide range of applications ranging from agriculture and manufacturing to healthcare and smart cities."

For specific use cases, software frameworks include NVIDIA Isaac Sim™ on Omniverse for robotics, NVIDIA Clara Holoscan™ SDK for healthcare, and NVIDIA DRIVE™ for autonomous driving. The latest Isaac release includes significant support for the Robot Operating System (ROS) developer community. NVIDIA has also released the new Omniverse Replicator for synthetic data generation and Isaac GEMs, hardware-accelerated software packages that make it easier for ROS developers to build high-performance AI-enabled robots on the Jetson platform.

Computing Platform for Real-Time Sensing Medical Instruments, Devices
NVIDIA AGX Orin powers NVIDIA Clara Holoscan, a new computing platform for the healthcare industry that allows developers to build software-defined medical devices which run low-latency streaming applications on the edge. This processing power is needed for devices such as robotic surgery, endoscopy and diagnostic imaging systems to process physics, imaging and visualization and enable real-time AI-decision support for doctors. Clara Holoscan gives developers a flexible platform to create AI microservices that run low-latency streaming applications on devices while passing more complex tasks to the data center.

New Era of Transportation Built on NVIDIA DRIVE Orin
DRIVE AGX Orin, also powered by the NVIDIA Ampere architecture like Jetson AGX Orin, is the platform of choice for the transportation industry. It is the advanced processor behind the newly announced NVIDIA DRIVE Concierge and DRIVE

Chauffeur — two AI platforms dedicated to redefining the passenger experience inside the car through AI assistants, and to powering safe, autonomous driving, respectively. Numerous global auto and truck makers, new energy vehicle startups and shuttle companies are using its high-performance AI compute for their next generation of intelligent, software-defined mobility solutions.

Jetson AGX Orin Availability
The NVIDIA Jetson AGX Orin module and developer kit will be available in the first quarter of 2022. Register to be notified about availability, access detailed specifications and download the white paper.

Register for free to learn more about Jetson AGX Orin during NVIDIA GTC, taking place online through Nov. 11. Watch NVIDIA founder and CEO Jensen Huang's GTC keynote address streaming on Nov. 9 and in replay.

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, features, performance and availability of our products and services, including NVIDIA Jetson AGX Orin, NVIDIA CUDA-X, NVIDIA JetPack SDK, NVIDIA NGC, NVIDIA TAO toolkit, NVIDIA Isaac Sim, NVIDIA Clara Holoscan SDK, NVIDIA DRIVE, Omniverse Replicator, DRIVE AGX Orin, DRIVE Concierge and DRIVE Chauffeur; NVIDIA setting the path for the future of edge AI and autonomous machines; robotics and embedded computing transforming manufacturing, healthcare, retail, transportation, smart cities and other essential sectors of the economy; the demand for processing continuing to surge; the robotics market growing exponentially; NVIDIA continuing to leverage AI expertise to advance robotics through a robust ecosystem and complete end-to-end solutions; the new Jetson platform bringing the performance and versatility of the Ampere architecture to enable even further advancements in autonomous mobile robots for a wide range of applications; and the processing power needed for devices, such as robotic surgery, endoscopy and diagnostic imaging systems to process physics, imaging and visualization and enable real-time AI-decision support 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.

David Pinto
+1-408-566-6950
dpinto@nvidia.com