NVIDIA and Arm Partner to Bring Deep Learning to Billions of IoT Devices

NVIDIA Deep Learning Accelerator IP to be Integrated into Arm Project Trillium Platform, Easing Building of Deep Learning IoT Chips

GPU Technology Conference -- NVIDIA and Arm today announced that they are partnering to bring deep learning inferencing to the billions of mobile, consumer electronics and Internet of Things devices that will enter the global marketplace.

Under this partnership, NVIDIA and Arm will integrate the open-source NVIDIA Deep Learning Accelerator (NVDLA) architecture into Arm's Project Trillium platform for machine learning. The collaboration will make it simple for IoT chip companies to integrate AI into their designs and help put intelligent, affordable products into the hands of billions of consumers worldwide.

"Inferencing will become a core capability of every IoT device in the future," said Deepu Talla, vice president and general manager of Autonomous Machines at NVIDIA. "Our partnership with Arm will help drive this wave of adoption by making it easy for hundreds of chip companies to incorporate deep learning technology."

"Accelerating AI at the edge is critical in enabling Arm's vision of connecting a trillion IoT devices," said Rene Haas, executive vice president, and president of the IP Group, at Arm. "Today we are one step closer to that vision by incorporating NVDLA into the Arm Project Trillium platform, as our entire ecosystem will immediately benefit from the expertise and capabilities our two companies bring in AI and IoT."

Based on NVIDIA® Xavier™, the world's most powerful autonomous machine system on a chip, NVDLA is a free, open architecture to promote a standard way to design deep learning inference accelerators. NVDLA's modular architecture is scalable, highly configurable and designed to simplify integration and portability.

NVDLA brings a host of benefits that speed the adoption of deep learning inference. It is supported by NVIDIA's suite of powerful developer tools, including upcoming versions of TensorRT, a programmable deep learning accelerator. The open-source design allows for cutting-edge features to be added regularly, including contributions from the research community.

The integration of NVDLA with Project Trillium will give deep learning developers the highest levels of performance as they leverage Arm's flexibility and scalability across the wide range of IoT devices.

"This is a win/win for IoT, mobile and embedded chip companies looking to design accelerated AI inferencing solutions," said Karl Freund, lead analyst for deep learning at Moor Insights & Strategy. "NVIDIA is the clear leader in ML training and Arm is the leader in IoT end points, so it makes a lot of sense for them to partner on IP."

About Arm

Arm technology is at the heart of a computing and connectivity revolution that is transforming the way people live and businesses operate. Our advanced, energy-efficient processor designs are enabling the intelligence in more than 125 billion silicon chips and securely powering products from the sensor to the smartphone to the supercomputer. With more than 1,000 technology partners, including the world's largest consumer brands, we are driving Arm innovation into all areas compute is happening inside the chip, the network and the cloud.

About NVIDIA

NVIDIA's (NASDAQ:NVDA) invention of the GPU in 1999 sparked the growth of the PC gaming market, redefined modern computer graphics and revolutionized parallel computing. More recently, GPU deep learning ignited modern AI — the next era of computing — with the GPU acting as the brain of computers, robots and self-driving cars that can perceive and understand the world. More information at http://nvidianews.nvidia.com/.

Certain statements in this press release including, but not limited to, statements as to: the impact and benefits of the partnership between NVIDIA and Arm; the impact, benefits and abilities of Arm IP offerings incorporating NVDLA; AI inferencing becoming a core capability of IoT devices in the future; the number of IoT devices in use; and the impact, benefits and abilities of NVDLA and how it will be supported 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 reports NVIDIA files with the Securities and Exchange Commission, or SEC, including its Form 10- K for the fiscal period ended January 28, 2018. 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.

Media Contacts

Kristin Uchiyama

+1 408 486 2248

kuchiyama@nvidia.com