NVIDIA and Partners Collaborate on Arm Computing for Cloud, HPC, Edge, PC

NVIDIA GPU + AWS Graviton2-Based Amazon EC2 Instances, HPC Developer Kit with Ampere Computing CPU and Dual GPUs, More Initiatives Help Expand Opportunities for Arm-Based Solutions

GTC -- NVIDIA today announced a series of collaborations that combine NVIDIA GPUs and software with Arm®-based CPUs — extending the benefits of Arm's flexible, energy-efficient architecture to computing workloads from the cloud to the edge.

The initiatives include combining AWS Graviton2-based Amazon EC2 instances with NVIDIA GPUs; supporting the development of scientific and AI applications with a new HPC Developer Kit; boosting video analytics and security features at the edge; and creating a new class of Arm-based PCs with NVIDIA RTX™ GPUs.

"Arm's energy efficiency and IP business model have made it the world's most popular CPU," said Jensen Huang, founder and CEO of NVIDIA. "Arm's ecosystem of technology companies from around the world are ready to take Arm-based products into new markets like cloud, supercomputing, PC and autonomous systems. With the new partnerships announced today, we're taking important steps to expand the Arm ecosystem beyond mobile and embedded."

NVIDIA GPU + AWS Graviton2-based Amazon EC2 Instances
NVIDIA and [AWS announced](#) that they are working together to deploy GPU-accelerated Arm-based instances in the cloud. The new Amazon EC2 instances will bring together AWS Graviton2 processors and NVIDIA GPUs to provide a range of benefits, including lower cost, support for richer game-streaming experiences, and greater performance for Arm-based workloads.

The instances will enable game developers to run Android games natively on AWS, accelerate rendering and encoding with NVIDIA GPUs, and stream games to mobile devices without the need to run emulation software.

Devkit for HPC
NVIDIA unveiled the NVIDIA Arm HPC Developer Kit to support scientific computing amid the growing need for energy-efficient supercomputers and data centers. It includes an Ampere® Altra® CPU, with 80 Arm Neoverse cores running up to 3.3GHz; dual NVIDIA A100 GPUs, each delivering 312 teraflops of FP16 deep learning performance, as well as two NVIDIA BlueField-2® DPUs, which accelerate networking, storage and security.

The devkit runs the [NVIDIA HPC SDK](#), a suite of compilers, libraries and tools that enable engineers, scientists and developers to create and migrate HPC and AI applications to a GPU-accelerated Arm computing system.

Developers and ISV partners can use the devkit to easily migrate and validate their software, and conduct performance analysis. Among initial leading computing centers deploying it are Oak Ridge National Laboratory, Los Alamos National Laboratory and Stony Brook University in the U.S.; the National Center for High Performance Computing, in Taiwan; and the Korean Institute of Science and Technology.

Sean Varley, senior director of Solutions at Ampere Computing, said, "The powerful combination of NVIDIA GPU and DPU accelerators with cloud-native Ampere Altra processors offers our customers leading performance, scalability, instance and core density, all while predictably executing very demanding data center and cloud workloads such as cloud gaming, HPC, scientific compute, artificial intelligence and networking. NVIDIA's Arm HPC Developer Kit is an easily accessible way to test drive this world-class combination for the most challenging workloads."

Developers can apply for [early access to the HPC devkit](#), which will be available in Q3.

Additionally, NVIDIA is working with lab directors, scientists and partners to help them port their software to run on Arm CPUs and NVIDIA GPUs.

Collaborations in Edge, PC
NVIDIA is expanding its collaboration with Marvell to combine OCTEON® DPUs with GPUs to accelerate cloud, enterprise, carrier and edge applications. This combination will speed up AI workloads such as network optimization and security from the edge to cloud, boosting system performance and reducing latency.

"Marvell's Arm-based OCTEON DPU platform has a long, proven track record of delivering industry-leading solutions to address the growing security, networking and storage requirements of cloud data centers," said Matt Murphy, president and CEO of Marvell. "We're thrilled to expand our multi-year relationship with NVIDIA to combine our OCTEON DPUs with

NVIDIA GPUs to accelerate video analytics and cybersecurity solutions for emerging edge to cloud applications."

In PCs, NVIDIA is partnering with MediaTek, one of the world's largest suppliers of Arm-based SoCs, to create a reference platform supporting Chromium, Linux and NVIDIA SDKs. The combination of NVIDIA RTX GPUs with high-performance, energy-efficient Arm Cortex® processors will bring realistic ray-traced graphics and cutting-edge AI to a new class of laptops.

MediaTek CEO Rick Tsai said, "MediaTek is the world's largest supplier of Arm chips, used to power everything from smartphones, Chromebooks and smart TVs. We look forward to using our technology and working with NVIDIA to bring the power of GPUs to the Arm PC platform for gaming, content creation and much more. GPU acceleration will be a huge boost for the entire Arm ecosystem."

Learn more during GTC21, taking place online April 12-16, and watch Huang's keynote address streaming live starting at 8:30 am PT on April 12. Registration is free.

Additional Partner Support
NVIDIA is collaborating on this work with other partners, including Fujitsu and SiPearl.

"Our collaboration with NVIDIA and the Arm ecosystem to accelerate HPC and data science will give researchers a broader development opportunity to build innovative solutions for scientific breakthroughs," said Naoki Shinjo, corporate executive officer, senior vice president and deputy head of the Future Society & Technology Unit at Fujitsu Limited.

"SiPearl is designing and bringing to market the European high-performance, energy-efficient processor, Rhea, that will power Europe's exascale supercomputers," said Philippe Notton, CEO of SiPearl. "NVIDIA's new tools and SDKs will help developers build accelerated HPC apps for Arm, providing supercomputing centers a rich Arm ecosystem."

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.

Certain statements in this press release including, but not limited to, statements as to: the impact and benefits of collaborations combining NVIDIA GPUs and software with Arm-based CPUs, as well as NVIDIA's efforts to accelerate Arm applications; growing interest in Arm-based solutions beyond mobile; growing demand for processing driving an enormous need for Arm's computing capabilities and the need for more investment; the availability of the NVIDIA Arm HPC Developer Kit; and SiPearl's Rhea processor powering Europe's exascale supercomputers 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.

Ken Brown
Corporate Communications
+1-408-486-2626
kebrown@nvidia.com