

NVIDIA GPU Cloud Now Available to Hundreds of Thousands of AI Researchers Using NVIDIA Desktop GPUs

NGC Expands Further, with NVIDIA TensorRT Inference Accelerator, ONNX Compatibility, Immediate Support for MXNet 1.0

LONG BEACH, Calif. -- NVIDIA today announced that hundreds of thousands of AI researchers using desktop GPUs can now tap into the power of [NVIDIA GPU Cloud](#) (NGC) as the company has extended NGC support to [NVIDIA TITAN](#).

NVIDIA also announced expanded NGC capabilities -- adding new software and other key updates to the NGC container registry -- to provide researchers a broader, more powerful set of tools to advance their AI and high performance computing research and development efforts.

Customers using NVIDIA® [Pascal™ architecture](#)-powered TITAN GPUs can [sign up](#) immediately for a no-charge NGC account and gain full access to a comprehensive catalog of GPU-optimized deep learning and HPC software and tools. Other supported computing platforms include [NVIDIA DGX-1™](#), [DGX Station](#) and NVIDIA Volta-enabled instances on Amazon EC2.

Software available through NGC's rapidly expanding container registry includes NVIDIA optimized deep learning frameworks such as TensorFlow and PyTorch, third-party managed HPC applications, NVIDIA HPC visualization tools, and NVIDIA's programmable inference accelerator, [NVIDIA TensorRT™ 3.0](#).

"We built NVIDIA GPU Cloud to give AI developers easy access to the software they need to do groundbreaking work," said Jim McHugh, vice president and general manager of enterprise systems at NVIDIA. "With GPU-optimized software now available to hundreds of thousands of researchers using NVIDIA desktop GPUs, NGC will be a catalyst for AI breakthroughs and a go-to resource for developers worldwide."

An early adopter of NGC is [GE Healthcare](#). The first medical device maker to use NGC, the company is tapping the deep learning software in NGC's container registry to accelerate bringing the most sophisticated AI to its 500,000 imaging devices globally with the goal of improving patient care.

New NGC Containers, Updates and Features

In addition to making NVIDIA TensorRT available on NGC's container registry, NVIDIA announced the following NGC updates:

- [Open Neural Network Exchange](#) (ONNX) support for TensorRT.
- Immediate support and availability for the first release of MXNet 1.0
- Availability of Baidu's PaddlePaddle AI framework

ONNX is an open format originally created by Facebook and Microsoft through which developers can exchange models across different frameworks. In the TensorRT development container, NVIDIA created a converter to deploy ONNX models to the TensorRT inference engine. This makes it easier for application developers to deploy low-latency, high-throughput models to TensorRT.

Together, these additions give developers a one-stop shop for software that supports a full spectrum of AI computing needs -- from research and application development to training and deployment.

[Launched](#) in October, NGC is also available free of charge to users of NVIDIA Volta GPUs on Amazon Web Services and all [NVIDIA DGX-1](#) and [DGX Station](#) customers. NVIDIA will continue to expand the reach of NGC over time.

More information about NGC is available at www.nvidia.com/gpu-cloud.

Keep Current on NVIDIA

Subscribe to the [NVIDIA blog](#), follow us on [Facebook](#), [Google+](#), [Twitter](#), [LinkedIn](#) and [Instagram](#), and view NVIDIA videos on [YouTube](#) and images on [Flickr](#).

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 abilities, benefits, impact and performance of the NVIDIA GPU Cloud container registry, including for AI frameworks, deep learning, high performance computing applications, visualization tools, NVIDIA's programmable inference accelerator and healthcare devices; the continued expansion of the NVIDIA GPU Cloud container registry, including new software, updates and applications; NGC being a catalyst for AI breakthroughs and becoming the go-to resource for developers worldwide; and GE Healthcare's use of the NGC container registry to accelerate bringing the most sophisticated artificial intelligence to GE Healthcare's imaging devices 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-Q for the fiscal period ended October 29, 2017. 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.

© 2017 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, Pascal and TensorRT are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated. Features, pricing, availability and specifications are subject to change without notice.

Media Contacts

Kristin Bryson

+1 203 241 9190

kbryson@nvidia.com