NVIDIA Launches AI Cloud Container Registry to Accelerate Deep Learning, Volta GPUs Debut on Amazon Web Services

Fully Tested, Optimized, Up-to-Date Deep Learning Software Available for Amazon EC2 P3 Instances Now

SANTA CLARA, Calif. -- NVIDIA today announced immediate availability of the NVIDIA® GPU Cloud (NGC) container registry for AI developers worldwide.

In just a few steps, NGC helps developers get started with deep learning development through no-cost access to a comprehensive, easy-to-use, fully optimized deep learning software stack.

The cloud-based service is available immediately to users of the just-announced Amazon Elastic Compute Cloud (Amazon EC2) P3 instances featuring NVIDIA Tesla® V100 GPUs. NVIDIA plans to expand support to other cloud platforms soon.

After signing up for an NGC account, developers can download a containerized software stack that integrates and optimizes a wide range of deep learning frameworks, NVIDIA libraries and CUDA® runtime versions -- which are kept up to date and run seamlessly in the cloud or on NVIDIA DGX™ systems.

“The NVIDIA GPU Cloud democratizes AI for a rapidly expanding global base of users," said Jim McHugh, vice president and general manager of Enterprise Systems at NVIDIA. “NGC frees developers from the complexity of integration, allowing them to move quickly to create sophisticated neural networks that deliver the transformative powers of AI."

Three Steps to Accelerated AI Computing
Developers who want to get started with deep learning right away using the NGC container registry can follow a three-step process:

1. Sign up for a no-cost NGC account at www.nvidia.com/ngcsignup.
2. Run an optimized NVIDIA image on cloud service provider platform.
3. Pull containers from NGC and get started.

Key benefits of the NGC container registry include:

- Instant access to the most widely used GPU-accelerated frameworks: Containerized software includes NVCaffe, Caffe2, Microsoft Cognitive Toolkit (CNTK), DIGITS, MXNet, PyTorch, TensorFlow, Theano and Torch, as well as CUDA for application development.
- Maximum Performance: Tuned, tested and certified by NVIDIA for maximum performance, the NGC container registry enables developers to get optimal performance on NVIDIA GPUs running on clouds.
- Pre-integration: Easy-to-use containers allow users to begin deep learning jobs immediately, eliminating time-consuming and difficult do-it-yourself software integration.
- Up to date: Containers available on the NGC container registry benefit from continuous NVIDIA development, ensuring each deep learning framework is tuned for the fastest training possible on the latest NVIDIA GPUs. NVIDIA engineers continually optimize libraries, drivers and containers, delivering monthly updates.

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 impact, benefits, performance and availability of NVIDIA GPU Cloud container registry; and NVIDIA’s plans to expand support to other cloud platforms 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 July 30, 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.

Media Contacts

Kristin Bryson
+1 203 241 9190

kbryson@nvidia.com