NVIDIA Releases CUDA 5, Making Programming With World's Most Pervasive Parallel Computing Platform Even Easier

Available as Free Download, Latest Version Features New Tools, Libraries and Functionality That Enhance Developer Productivity

SANTA CLARA, CA -- NVIDIA today made available the NVIDIA® CUDA® 5 production release, a powerful new version of the world's most pervasive parallel computing platform and programming model for accelerating scientific and engineering applications on GPUs. It can be downloaded for free from the NVIDIA Developer Zone website.

With more than 1.5 million downloads, supporting more than 180 leading engineering, scientific and commercial applications, the CUDA programming model is the most popular way for developers to take advantage of GPU-accelerated computing.

Building on this success, the new programming features of the CUDA 5 platform make the development of GPU-accelerated applications faster and easier than ever, including support for dynamic parallelism, GPU-callable libraries, NVIDIA GPUDirect™ technology support for RDMA (remote direct memory access) and the NVIDIA Nsight™ Eclipse Edition integrated development environment (IDE).

Developer Accolades for CUDA 5
Developers who evaluated the pre-release version of CUDA 5 have reported often dramatic application acceleration and improved programmability.

The defense and aerospace industries realize the benefits of CUDA GPU acceleration for processing images, video and sensor data, such as radar. According to Dustin Franklin, GPGPU applications engineer at GE Intelligent Platforms in Charlottesville, Va., "CUDA 5 is a significant technology for us. Many of the applications we're using involve streaming sensor data directly into the GPU with low latency, so the GPUDirect support for RDMA on new Kepler GPUs is incredibly important for our customers. We have integrated support for many custom sensors already and are very happy with the results."

Guillaume Belz, a research biochemist at Lyon University Hospital in Lyon, France, has been using dynamic parallelism and GPU-callable libraries for complex signal analysis and data mining. "With GPU acceleration, we can get results in several hours for projects that used to require weeks or even months with CPUs alone. Without GPU acceleration, analysis is not possible at all," said Belz.

Weihua (Wayne) Sun, Ph.D. candidate in imaging science at Rochester Institute of Technology in New York, was impressed with NVIDIA Nsight Eclipse Edition. "When I learned that CUDA 5 included the new Nsight Eclipse Edition IDE, I knew I needed it right away. Having all my programming, debugging and optimization tools in a single integrated development environment is a great productivity boost for me."

New CUDA 5 Features
CUDA 5 enables developers to take full advantage of the performance of NVIDIA GPUs, including GPU accelerators based on the NVIDIA Kepler™ compute architecture -- the fastest, most efficient, highest-performance computing architecture ever built. Key features include:

- **Dynamic Parallelism - Brings GPU acceleration to new algorithms**
  GPU threads can dynamically spawn new threads, allowing the GPU to adapt to the data. By minimizing the back and forth with the CPU, dynamic parallelism greatly simplifies parallel programming. And it enables GPU acceleration of a broader set of popular algorithms, such as those used in adaptive mesh refinement and computational fluid dynamics applications.

- **GPU-Callable Libraries - Enables third-party ecosystem**
  A new CUDA BLAS library allows developers to use dynamic parallelism for their own GPU-callable libraries. They can design plug-in APIs that allow other developers to extend the functionality of their kernels, and allow them to implement callbacks on the GPU to customize the functionality of third-party GPU-callable libraries. The "object linking" capability provides an efficient and familiar process for developing large GPU applications by enabling developers to compile multiple CUDA source files into separate object files, and link them into larger applications and libraries.

- **GPUDirect Support for RDMA - Minimizes system memory bottlenecks**
  GPUDirect technology enables direct communication between GPUs and other PCI-E devices, and supports direct

memory access between network interface cards and the GPU. It also significantly reduces MPISendRecv latency between GPU nodes in a cluster and improves overall application performance.

- **NVIDIA Nsight Eclipse Edition - Generate CUDA code quickly and easily**
  NVIDIA Nsight Eclipse Edition enables programmers to develop, debug and profile GPU applications within the familiar Eclipse-based IDE on Linux and Mac OS X platforms. An integrated CUDA editor and CUDA samples speed the generation of CUDA code, and automatic code refactoring enables easy porting of CPU loops to CUDA kernels. An integrated expert analysis system provides automated performance analysis and step-by-step guidance to fix performance bottlenecks in the code, while syntax highlighting makes it easy to differentiate GPU code from CPU code.

New Online CUDA Resource Center
To help developers maximize the potential of parallel computing with CUDA technology, NVIDIA has launched a free online resource center for CUDA programmers at http://docs.nvidia.com. The site offers the latest information on the CUDA platform and programming model, as well as access to all CUDA developer documentation and technologies, including tools, code samples, libraries, APIs, and tuning and programming guides.

CUDA Registered Developer Program
Parallel programmers are invited to join the free CUDA Registered Developer Program for early access to software releases, tools and resources. Visit http://www.nvidia.com/paralleldeveloper for more information.

About CUDA
CUDA is a parallel computing platform and programming model developed by NVIDIA. It enables dramatic increases in computing performance by harnessing the power of GPUs.

More information about GPU Computing is available at the NVIDIA website. To learn more about CUDA or download the latest version, visit the CUDA website. More NVIDIA news, company and product information, videos, images and other information is available at the NVIDIA newsroom. Follow us on Twitter at @GPUComputing.

About NVIDIA
NVIDIA (NASDAQ: NVDA) awakened the world to computer graphics when it invented the GPU in 1999. Today, its processors power a broad range of products from smartphones to supercomputers. NVIDIA's mobile processors are used in cell phones, tablets and auto infotainment systems. PC gamers rely on GPUs to enjoy spectacularly immersive worlds. Professionals use them to create 3D graphics and visual effects in movies and to design everything from golf clubs to jumbo jets. And researchers utilize GPUs to advance the frontiers of science with high performance computing. The company has more than 5,000 patents issued, allowed or filed, including ones covering ideas essential to modern computing. For more information, see www.nvidia.com.

Certain statements in this press release including, but not limited to, statements as to: the impact, performance, features and benefits of the NVIDIA CUDA parallel computing platform and the effects of the company's patents on modern computing 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 29, 2012. 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.

© 2012 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, CUDA, GPUDirect, Kepler, and Nsight, 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.

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