

New CUDA 4.0 Release Makes Parallel Programming Easier

Unified Virtual Addressing, GPU-to-GPU Communication and Enhanced C++ Template Libraries Enable More Developers to Take Advantage of GPU Computing

SANTA CLARA, CA -- NVIDIA today announced the latest version of the NVIDIA® CUDA® Toolkit for developing parallel applications using NVIDIA GPUs.

The NVIDIA CUDA 4.0 Toolkit was designed to make parallel programming easier, and enable more developers to port their applications to GPUs. This has resulted in three main features:

- [NVIDIA GPUDirect™ 2.0 Technology](#) -- Offers support for peer-to-peer communication among GPUs within a single server or workstation. This enables easier and faster multi-GPU programming and application performance.
- [Unified Virtual Addressing \(UVA\)](#) -- Provides a single merged-memory address space for the main system memory and the GPU memories, enabling quicker and easier parallel programming.
- [Thrust C++ Template Performance Primitives Libraries](#) -- Provides a collection of powerful open source C++ parallel algorithms and data structures that ease programming for C++ developers. With Thrust, routines such as parallel sorting are 5X to 100X faster than with Standard Template Library (STL) and Threading Building Blocks (TBB).

"Unified virtual addressing and faster GPU-to-GPU communication makes it easier for developers to take advantage of the parallel computing capability of GPUs," said John Stone, senior research programmer, University of Illinois, Urbana-Champaign.

"Having access to GPU computing through the standard template interface greatly increases productivity for a wide range of tasks, from simple cashflow generation to complex computations with Libor market models, variable annuities or CVA adjustments," said Peter Decrem, director of Rates Products at Quantifi. "The Thrust C++ library has lowered the barrier of entry significantly by taking care of low-level functionality like memory access and allocation, allowing the financial engineer to focus on algorithm development in a GPU-enhanced environment."

The CUDA 4.0 architecture release includes a number of other key features and capabilities, including:

- [MPI Integration with CUDA Applications](#) -- Modified MPI implementations automatically move data from and to the GPU memory over Infiniband when an application does an MPI send or receive call.
- [Multi-thread Sharing of GPUs](#) -- Multiple CPU host threads can share contexts on a single GPU, making it easier to share a single GPU by multi-threaded applications.
- [Multi-GPU Sharing by Single CPU Thread](#) -- A single CPU host thread can access all GPUs in a system. Developers can easily coordinate work across multiple GPUs for tasks such as "halo" exchange in applications.
- [New NPP Image and Computer Vision Library](#) -- A rich set of image transformation operations that enable rapid development of imaging and computer vision applications.
- [New and Improved Capabilities](#)
 - Auto performance analysis in the Visual Profiler
 - New features in cuda-gdb and added support for MacOS
 - Added support for C++ features like new/delete and virtual functions
 - New GPU binary disassembler

A release candidate of CUDA Toolkit 4.0 will be available free of charge beginning March 4, 2011, by enrolling in the CUDA Registered Developer Program at: www.nvidia.com/paralleldeveloper. The CUDA Registered Developer Program provides a wealth of tools, resources, and information for parallel application developers to maximize the potential of CUDA.

For more information on the features and capabilities of the CUDA Toolkit and on GPGPU applications, please visit: www.nvidia.com/cuda.

About NVIDIA

NVIDIA (NASDAQ: NVDA) awakened the world to the power of computer graphics when it invented the GPU in 1999. Since then, it has consistently set new standards in visual computing with breathtaking, interactive graphics available on devices ranging from tablets and portable media players to notebooks and workstations. NVIDIA's expertise in programmable GPUs has led to breakthroughs in parallel processing which make supercomputing inexpensive and widely accessible. The Company holds more than 1,700 patents worldwide, including ones covering designs and insights that are 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 features, capabilities and impact of

the NVIDIA CUDA 4.0 architecture and NVIDIA GPUs; 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 faster or more efficient technology; 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 31, 2010. 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.

© 2011 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, and CUDA 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