



NVIDIA Tesla GPUs to Communicate Faster Over Mellanox InfiniBand Networks

SC09 -- NVIDIA Corporation (NASDAQ: NVDA) and Mellanox Technologies Ltd. today introduced new software that will increase cluster application performance by as much as 30% by reducing the latency that occurs when communicating over Mellanox InfiniBand to servers equipped with NVIDIA Tesla™ GPUs.

The system architecture of a GPU-CPU server requires the CPU to initiate and manage memory transfers between the GPU and the InfiniBand network. The new software solution will enable Tesla GPUs to transfer data to pinned system memory that a Mellanox InfiniBand solution is able to read and transmit over the network. The result is increased overall system performance and efficiency.

"NVIDIA Tesla GPUs deliver large increases in performance across each node in a cluster, but in our production runs on TSUBAME 1 we have found that network communication becomes a bottleneck when using multiple GPUs," said Prof. Satoshi Matsuoka from Tokyo Institute of Technology. "Reducing the dependency on the CPU by using InfiniBand will deliver a major boost in performance in high performance GPU clusters, thanks to the work of NVIDIA and Mellanox, and will further enhance the architectural advances we will make in TSUBAME2.0."

"In GPU-based clusters, most of the compute intensive processing is running on the GPUs," said Gilad Shainer, director of high performance computing and technical marketing at Mellanox Technologies. "It's a natural evolution of the system architecture to enable GPUs to communicate more intelligently over InfiniBand. This helps create a computing platform that will enable future Exascale computing and dramatically increase performance for a broad spectrum of applications."

"Anyone who cares about performance in their datacenter uses InfiniBand," said Andy Keane, general manager, Tesla business at NVIDIA. "This new feature will further improve application performance on GPU-based clusters by reducing the dependency on the CPU for communicating over InfiniBand."

This software capability will be available in the NVIDIA CUDA™ architecture toolkit beginning in Q2 2010 and will work on existing Tesla S1070 1U computing systems and Tesla M1060 module-based clusters and also with the new Tesla 20-series S2050 and S2070 1U systems.

About NVIDIA

NVIDIA (NASDAQ: NVDA) awakened the world to the power of computer graphics when it invented the graphics processing unit (GPU) in 1999. Since then, it has consistently set new standards in visual computing with breathtaking, interactive graphics available on devices ranging from portable media players to notebooks to workstations. NVIDIA's expertise in programmable GPUs has led to breakthroughs in parallel processing which make supercomputing inexpensive and widely accessible. Fortune magazine has ranked NVIDIA #1 in innovation in the semiconductor industry for two years in a row. For more information, see www.nvidia.com.

Certain statements in this press release including, but not limited to, statements as to: the benefits, features, impact, performance and capabilities of NVIDIA Tesla GPUs 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: development of more efficient or faster technology; design, manufacturing or software defects; the impact of technological development and competition; changes in consumer preferences and demands; customer adoption of different standards or our competitor's products; 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 including its Form 10-Q for the fiscal period ended July 26, 2009. Copies of reports filed with the SEC are posted on our 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.

© 2009 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, Tesla, CUDA, GeForce and Quadro are trademarks 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.

Hector Martinez
Corporate Communications
+1-408-486-3443
hmarinez@nvidia.com