



Xeon® Phi™ Coprocessor



APPLICATION DEVELOPMENT

Intel's Many Integrated Core (MIC) accelerator. Ohio Supercomputer Center engineers began a collaboration with Intel on application development and testing more than a year ago.

Hudak leads group evaluating new processor technologies

The Xeon Phi co-processor is based upon a new technology developed by Intel Corporation to provide greater compute performance through massive parallelism and designed to fit many high-performance computing applications. The first Xeon Phi-based HPC system, Stampede, will be deployed in early 2013 at the Texas Advanced Computing Center as part of the National Science Foundation's Extreme Science and Engineering Discovery Environment (XSEDE) program.

In conjunction with a larger project, David Hudak, Ph.D., program director for cyberinfrastructure and software development at the Ohio Supercomputer Center (OSC), recently led a research team that examined the performance of different HPC algorithms on pre-release Xeon Phi hardware. The group included OSC staff members John Eisenlohr, Ph.D., a systems developer and engineer, and Karen Tomko, Ph.D., a senior researcher in computer science. Kurt O'Hearn, an undergraduate student at Grand Valley State University, worked with the team as part of an XSEDE student engagement project.

"The unique design of the Xeon Phi architecture required the specification of application parallelism for maximum performance," Hudak explained. "It is important for application

developers to experiment and gain experience on the best way to structure applications to exploit the Xeon Phi."

To achieve this end, Hudak's group performed a detailed performance analysis of a well-known communication-avoiding QR factorization algorithm (CAQR), compared CAQR performance between the first- and second-generation Xeon Phi chips and restructured the algorithm for improved performance.

"QR factorization is a linear algebra routine fundamental to solving commonplace problems in the sciences and engineering," said Hudak. "As such, it has been extensively studied on all major HPC architectures: vector, MPP, SMP, cluster, multi-core cluster and GPU cluster."

The parallelization of the CAQR algorithm proved non-trivial due to load imbalance caused by the decreasing amount of matrix to be factored as the algorithm progresses. To better understand the behavior of the algorithm on the Xeon Phi, timing methods were used to measure the total execution time and the four main kernels.

The team concluded that CAQR execution time on the Xeon Phi is sensitive to both matrix shape and tile shape and that the optimal tile shape appears to correlate with cache size (larger tiles perform better with larger caches). They also found that the execution time for two kernels dominates for larger thread counts, indicating a need for improved collective operations.

Project lead: David Hudak, Ohio Supercomputer Center
Research title: Parallel application development for the Intel Xeon Phi platform
Funding source: Intel, Ohio Supercomputer Center
Web site: <http://bit.ly/OSC-RR-Hudak>