Improving energy, performance efficiency for high performance computing

April 8 2015

Shuaiwen Leon Song, a research scientist with PNNL's High Performance Computing group, and Chao Li, a Ph.D. student with North Carolina State University's Department of Electrical and Computer Engineering who spent time as a research intern at PNNL in 2014, are co-authors of, "Locality-Driven Dynamic GPU Cache Bypassing," which recently was accepted by the 29th International Conference on Supercomputing (ICS). The paper, which presents novel cache optimizations for massively parallel, throughput-oriented architectures, such as GPUs, will be presented during the ICS 2015 Conference Program.

"Chao and I are very excited about this work at ICS, which can potentially change current commercial cache designs for major GPU vendors as the dynamic filter approach we propose has not only good performance and energy efficiency improvement, but also very little area and design overhead," Song explained. "For instance, a 16 KB L1 D-cache from our design can outperform a 4 MB baseline cache from current commercial GPU architectures, such as NVIDIA Fermi and Kepler, for many irregular workloads. We believe this contribution is quite important to the HPC field and industry development."

The paper features additional co-authors from NCSU and NVIDIA Research.

"This is the latest excellent publication from Leon, whose work on energy efficiency and resilience in high-performance computing will

also be presented at IPDPS [International Parallel & Distributed Processing Symposium] in May," said Darren Kerbyson, the associate division director of PNNL's HPC group. "It is great to be associated with so many technically excellent researchers at PNNL, who have recently contributed to many top-notch venues over the last year, including Supercomputing, IPDPS, PLDI [Programming Language Design and Implementation], and PPoPP [Principles and Practice of Parallel Programming]."

More information: "Locality-Driven Dynamic GPU Cache Bypassing." In 29th International Conference on Supercomputing (ICS 2015). June 08-11, 2015, Newport Beach, California. www.cs.ucr.edu/~ics15/

Provided by Pacific Northwest National Laboratory