HCC-Memcached

Technology #2018-057

GPUs have evolved from fixed function 3D accelerators to fully programmable units that are starting to be in use for high-performance computing (HPC), machine-learning and data-analytics. GPUs increasingly support programmability features such as shared virtual memory and hardware cache coherence enabling them to run a wider variety of programs. However, the ability to invoke system calls, an important aspect of CPU programming continues to be absent for GPUs.

Description:

Researches at Rutgers University have created the GENESYS framework for invoking generic system calls on GPUs. **HCC-Memcached** is a version of the memcached software utility (memcached.org) that has been created to demonstrate the utility of GPU invoked system call to perform network I/O using the GENESYS framework.

HCC-Memcached is a server application that uses a binary form of the memcached protocol over UDP. By implementing this standard protocol, HCC-Memcached is interoperable with existing memcached clients.

The focus of HCC-Memcached is on the SET and GET commands; SET stores a key-value pair, and GET retrieves a value associated with a key if it is present.

The software supports a fixed-size hash table as back-end storage. CPU and GPU share the hash table, with a configurable bucket size and count. HCC-Memcached enables concurrent operations. CPUs handle SETs and GETs, while the GPU supports only GETs. The GPU implementation parallelizes the hash computation, bucket lookup, and data copy and uses the sendto and recvfrom system calls for UDP network access. Invoking these system calls at work-group granularity with blocking and weak ordering perform best.

Results:

- Performance of a CPU version shows that the GPUs accelerate HCC-Memcached when parallelizing lookups on buckets with more elements.
- HCC-Memcached with GENESYS achieves 30-40% latency and throughput benefits over not just CPU versions, but also GPU versions implemented without direct system calls.
- HCC-Memcached without using GENESYS for system calls results in GPU performance that lags behind CPU performance.
The development of HCC-Memcached was for demonstrating an application making system calls with GENESYS. It is no longer under development or supported. The source code is available on GitHub.

Inventors

Jan Vesely

Jan Vesely, received his B.S. and M.S. degrees in Computer Science from Charles University in Prague and is currently a PhD student at Rutgers University – Computer Science department. Jan’s research involves under the guidance of Dr. Abhishek Bhattacharjee.

Dr. Abhishek Bhattacharjee, PhD

Dr. Abhishek Bhattacharjee, PhD., is an Associate Professor, Department of Computer Science at Rutgers University. His current research interests include designing systems to advance the brain sciences. Dr. Abhishek is designing embedded chips for brain-machine implants that directly interface with brain tissue, and developing compilation techniques to realize efficient models of the brain for computational neuroscience and psychiatry.