Using Hybrid Model OpenSHMEM + CUDA to Implement the SHOC Benchmark Suite

Achievement: Ported several kernels from the Scalable HeterOgeneous Computing (SHOC) benchmark suite from the hybrid MPI + CUDA implementation to an OpenSHMEM + CUDA implementation and demonstrated OpenSHMEM as an alternative communication layer at moderate core counts.

Significance and Impact: This work demonstrates that OpenSHMEM can be effectively used as an alternative to MPI for message passing in inter-device communication in accelerated systems. Specifically, in hybrid models combining OpenSHMEM with CUDA or OpenCL.

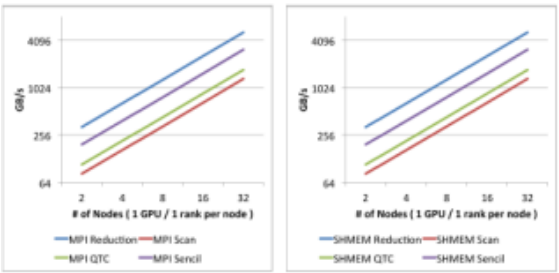

Figure: Truly parallel benchmarks with MPI (left), truly parallel benchmarks with SHMEM (right) indicate similar runtime performance

Research details:

- SHOC benchmarks are written in CUDA and OpenCL, grouped in 3 levels: raw speed tests, fundamental algorithms, application inspired kernels, and use MPI for communication. The research focused on the 4 kernels with truly parallel (TP) implementations as there is no communication for the other (embarrassingly parallel) kernels). Note that the TP benchmarks have a small amount of communication time compared to kernel computation.
- Porting these benchmarks requires both "process teams" and "MPI-style synchronization collectives" supporting teams and was done in a 2-phase approach: first porting on Cray using the Cray teams extensions (Cray Message Passing Toolkit implementation of SHMEM), followed by a port to OpenSHMEM.
- Performance tests run on Cray XK7/Titan run up to 32 nodes

Sponsor/Facility: Work was supported by the DoD, the Extreme Scale Systems Center at ORNL, and the Oak Ridge Leadership Computing Facility

PI and affiliation: M. Grodowitz (ORNL)

Team: M. Grodowitz, E. D'Azevedo, S. Powers (team lead) and N. Imam (project lead)

Publication: M. Grodowitz, E. D'Azevedo, S. Powers and N. Imam, "*Using Hybrid Model OpenSHMEM + CUDA to Implement the SHOC Benchmark Suite*," in Proceedings of Third OpenSHMEM Workshop, Baltimore, MD, USA, August 2016.

Overview:

This work describes the process of porting the Scalable HeterOgeneous Computing (SHOC) benchmark suite from the hybrid MPI + CUDA implementation to an OpenSHMEM + CUDA implementation. SHOC includes a wide variety of benchmark kernels used to measure accelerator performance in both single node and cluster configurations. The hybrid model implementation attempts to place all major computation on accelerator devices, and uses MPI to synchronize and aggregate results. In some cases, MPI Groups are used to gradually reduce the number of accelerators used for computation as the problem size drops. Porting this behavior to OpenSHMEM required implementing several synchronizing collective operations, and using SHMEM teams to replace MPI Group functionality. Benchmark results on a Cray XK7 system with one GPU per compute node show that SHMEM performance is equal to MPI performance in these hybrid tasks. These results and porting experience show that using OpenSHMEM for accelerator devices benefits from adding functionality for synchronization and teams, and would further benefit from adding support for communication within accelerator kernels.