Accelerated Big Data Microbenchmarks

- Companies like Map-D are at the forefront of accelerated data analytics but their solutions are closed-source
- Hadoop, graph analytics are reasonably well represented with current CPU-focused benchmarks, but...
- **There are limited opportunities to compare different accelerated architectures for data analytics**

Data Analytics for the SHOC Suite

- Scalable HeterOgeneous Computing (SHOC): Accelerator-based benchmark suite that provides benchmarks written in multiple languages [1]
- Designed as a tool to compare algorithms across software platforms but also to compare hardware systems
- OpenCL, CUDA, Phi (OpenMP), and OpenACC variants include “speeds and feeds” benchmarks as well as parallel benchmarks
- Currently there is a focus to add more “Big Data” benchmarks to represent non-scientific workloads
- TPC-H [2] primitives and queries are a good candidate along with ML and graph algorithms

Analytics Primitives

- Basic design – 1) partition 2) compute 3) gather
- OpenCL implementations of select, reduce, join, etc. and microbenchmarks that incorporate common patterns (A, B, C).

Evaluation and Results

- Multiple platforms tested and discussed in [3]. Used OpenCL runtimes from NVIDIA, Intel, AMD, and Beignet [4] with common code. Code is available as a branch of SHOC [5]

  ![Project (Compute)](image1)

  ![Select (Compute)](image2)

  ![Select (Compute and Data)](image3)

- Project is highly parallel; .22 gigaops/sec (GOPS) up to 7.54 GOPS on K20m; Xeon Phi is penalized by lack of vectorization opportunities (up to 2.17 GOPS)

- In terms of compute, high-end GPUs provide best overall performance (3.02-4.02 GOPs) while Trinity suffers due to low thread count and clocks (2.16 GOPs for 256 MB)

- PCIe performance brings added penalties to discrete parts like the K20 and Phi while zero-copy semantics benefit Haswell GPU (up to .127 GOPs).

References