Contribution ID: **297** Type: **Oral**

Simulating Diverse HEP Workflows on Heterogeneous Architectures

Wednesday 13 March 2019 15:30 (20 minutes)

The next generation of HPC and HTC facilities, such as Oak Ridge's Summit, Lawrence Livermore's Sierra, and NERSC's Perlmutter, show an increasing use of GPGPUs and other accelerators in order to achieve their high FLOP counts. This trend will only grow with exascale facilities such as A21. In general, High Energy Physics computing workflows have made little use of GPUs due to the relatively small fraction of kernels that run efficiently on GPUs, and the expense of rewriting code for rapidly evolving GPU hardware. However, the computing requirements for high-luminosity LHC are enormous, and it will become essential to be able to make use of supercomputing facilities that rely heavily on GPUs and other accelerator technologies.

ATLAS has already developed an extension to AthenaMT, its multithreaded event processing framework, that enables the non-intrusive offloading of computations to external accelerator resources, and has begun investigating strategies to schedule the offloading efficiently. The same applies to LHCb, which, while sharing the same underlying framework as ATLAS (Gaudi), has considerably different workflow. CMS's framework, CMSSW, also has the ability to efficiently offload tasks to external accelerators. But before investing heavily in writing many kernels for specific offloading architectures, we need to better understand the performance metrics and throughput bounds of the workflows with various accelerator configurations. This can be done by simulating a diverse set of workflows, using real metrics for task interdependencies and timing, as we vary fractions of offloaded tasks, latencies, data conversion speeds, memory bandwidths, and accelerator offloading parameters such as CPU/GPU ratios and speeds.

We present the results of these studies performed on multiple workflows from ATLAS, LHCb and CMS, which will be instrumental in directing effort to make HEP framework, kernels and workflows run efficiently on exascale facilities.

Authors: Dr LEGGETT, Charles (Lawrence Berkeley National Lab (US)); SHAPOVAL, Illya (Lawrence Berkeley National Laboratory); CLEMENCIC, Marco (CERN); JONES, Christopher (Fermi National Accelerator Lab. (US))

Presenter: Dr LEGGETT, Charles (Lawrence Berkeley National Lab (US))

Session Classification: Track 1: Computing Technology for Physics Research

Track Classification: Track 1: Computing Technology for Physics Research