



Contribution ID: 58

Type: Oral

Next generation task scheduler for ATLAS software framework

Monday, 24 October 2022 15:10 (20 minutes)

Experiments at the CERN High-Luminosity Large Hadron Collider (HL-LHC) will produce hundreds of Petabytes of data per year. Efficient processing of this dataset represents a significant human resource and technical challenge. Today, ATLAS data processing applications run in multi-threaded mode, using Intel TBB for thread management, which allows efficient utilization of all available CPU cores on the computing resources. However, modern HPC systems and high-end computing clusters are increasingly based on heterogeneous architectures, usually a combination of CPU and accelerators (e.g., GPU, FPGA). To run ATLAS software on these machines efficiently, we started developing a distributed, fine-grained, vertically integrated task scheduling software system. A first simplified implementation of such a system called Raythena was developed in late 2019. It is based on Ray - a high-performance distributed execution platform developed by Riselab at UC Berkeley. Raythena leverages the ATLAS event-service architecture for efficient utilization of CPU resources on HPC systems by dynamically assigning fine-grained workloads (individual events or event ranges) to ATLAS data-processing applications running simultaneously on multiple HPC compute nodes.

The main purpose of the Raythena project was to gain the experience of developing real-life applications with the Ray platform. However, in order to achieve our main objective, we need to design a new system capable of utilizing heterogeneous computing resources in a distributed environment. To accomplish this, we have started to evaluate HPX as an alternative to TBB/Ray. HPX is a C++ library for concurrency and parallelism developed by the Stellar group, which exposes a uniform, standards-oriented API for programming parallel, distributed, and heterogeneous applications.

This presentation will describe the preliminary results of the evaluation of HPX for implementation of the task scheduler for ATLAS data-processing applications aimed to enable cross-node scheduling in heterogeneous systems that offer a mixture of CPU and GPU architectures. We present the prototype applications implemented using HPX and the preliminary results of performance studies of these applications.

Significance

This presentation describes design ideas and first simple prototype implementations of the distributed and heterogeneous task scheduling system for the ATLAS experiment. Given the increased data volumes expected to be recorded in the era of HL LHC, it becomes critical for the experiments to efficiently utilize all available computing resources, including the new generation of supercomputers, most of which will be based on heterogeneous architectures.

References

Experiment context, if any

ATLAS

Primary authors: STANISLAUS, Beojan (Lawrence Berkeley National Lab. (US)); Dr LEGGETT, Charles (Lawrence Berkeley National Lab (US)); ESSEIVA, Julien (Lawrence Berkeley National Lab. (US)); CALAFIURA, Paolo (Lawrence Berkeley National Lab. (US)); TSULAIA, Vakho (Lawrence Berkeley National Lab. (US)); JU, Xiangyang (Lawrence Berkeley National Lab. (US))

Presenter: STANISLAUS, Beojan (Lawrence Berkeley National Lab. (US))

Session Classification: Track 1: Computing Technology for Physics Research

Track Classification: Track 1: Computing Technology for Physics Research