Contribution ID: **62**   Type: **Oral**

Bringing heterogeneity to the CMS software framework

Monday, 4 November 2019 14:30 (15 minutes)

The advent of computing resources with co-processors, for example Graphics Processing Units (GPU) or Field-Programmable Gate Arrays (FPGA), for use cases like the CMS High-Level Trigger (HLT) or data processing at leadership-class supercomputers imposes challenges for the current data processing frameworks. These challenges include developing a model for algorithms to offload their computations on the co-processors as well as keeping the traditional CPU busy doing other work. The CMS data processing framework, CMSSW, implements multithreading using the Intel's Threading Building Blocks (TBB) library, that utilizes tasks as concurrent units of work. In this talk we will discuss a generic mechanism to interact effectively with non-CPU resources that has been implemented in CMSSW. In addition, configuring such a heterogeneous system is challenging. In CMSSW an application is configured with a configuration file written in the Python language. The algorithm types are part of the configuration. The challenge therefore is to unify the CPU and co-processor settings while allowing their implementations to be separate. We will explain how we solved these challenges while minimizing the necessary changes to the CMSSW framework. We will also discuss on a concrete example how algorithms would offload work to NVIDIA GPUs using directly the CUDA API.

Consider for promotion

Yes

Authors: CMS COLLABORATION; GUTSCHE, Oliver (Fermi National Accelerator Lab. (US)); JONES, Christopher (Fermi National Accelerator Lab. (US))

Presenter: GUTSCHE, Oliver (Fermi National Accelerator Lab. (US))

Session Classification: Track 5 –Software Development

Track Classification: Track 5 –Software Development