

Evolution of the ATLAS Distributed Computing system during the LHC Long Shutdown

Monday, 14 October 2013 14:14 (22 minutes)

The ATLAS Distributed Computing project (ADC) was established in 2007 to develop and operate a framework, following the ATLAS computing model, to enable data storage, processing and bookkeeping on top of the WLCG distributed infrastructure. ADC development has always been driven by operations and this contributed to its success. The system has fulfilled the demanding requirements of ATLAS, daily consolidating worldwide up to 1PB of data and running more than 1.5 million payloads distributed globally, supporting almost one thousand concurrent distributed analysis users. Comprehensive automation and monitoring minimized the operational manpower required. The flexibility of the system to adjust to operational needs has been important to the success of the ATLAS physics program.

The LHC shutdown in 2013-2015 affords an opportunity to improve the system in light of operational experience and scale it to cope with the demanding requirements of 2015 and beyond, most notably a much higher trigger rate and event pileup. We will describe the evolution of the ADC software foreseen during this period. This includes consolidating the existing Production and Distributed Analysis framework (PanDA) and ATLAS Grid Information System (AGIS), together with the development and commissioning of next generation systems for distributed data management (DDM/Rucio) and production (PRODSYS2). We will explain how new technologies such as Cloud Computing and NoSQL databases, which ATLAS investigated as R&D projects in past years, will be integrated in production. Finally, we will describe more fundamental developments such as breaking job-to-data locality by exploiting storage federations and caches, and event level (rather than file or dataset level) workload engines.

Summary

Primary author: CAMPANA, Simone (CERN)

Presenter: CAMPANA, Simone (CERN)

Session Classification: Distributed Processing and Data Handling B: Experiment Data Processing, Data Handling and Computing Models

Track Classification: Distributed Processing and Data Handling B: Experiment Data Processing, Data Handling and Computing Models