Contribution ID: **154** Type: **presentation**

Towards an Event Streaming Service for ATLAS data processing

Wednesday, 11 July 2018 11:30 (15 minutes)

he ATLAS experiment is gradually transitioning from the traditional file-based processing model to dynamic workflow management at the event level with the ATLAS Event Service (AES). The AES assigns fine-grained processing jobs to workers and streams out the data in quasi-real time, ensuring fully efficient utilization of all resources, including the most volatile. The next major step in this evolution is the possibility to intelligently stream the input data itself to workers. The Event Streaming Service (ESS) is now in development to asynchronously deliver only the input data required for processing when it is needed, protecting the application payload from WAN latency without creating expensive long-term replicas. In the current prototype implementation, ESS processes run on compute nodes in parallel to the payload, reading the input event ranges remotely over the network, and replicating them in small input files that are passed to the application. In this contribution, we present the performance of the ESS prototype for different types of workflows in comparison to tasks accessing remote data directly. Based on the experience gained with the current prototype, we are now moving to the development of a server-side component of the ESS. The service can evolve progressively into a powerful CDN-like capability for data streaming, ultimately enabling the delivery of 'virtual data' generated on demand.

Primary authors: MAGINI, Nicolo (INFN e Universita Genova (IT)); TSULAIA, Vakho (Lawrence Berkeley National Lab. (US)); WENAUS, Torre (Brookhaven National Laboratory (US)); GUAN, Wen (University of Wisconsin (US)); WALKER, Rodney (Ludwig Maximilians Universitat (DE)); NILSSON, Paul (Brookhaven National Laboratory (US)); MAENO, Tadashi (Brookhaven National Laboratory (US)); LASSNIG, Mario (CERN); DI GIROLAMO, Alessandro (CERN)

Presenter: MAGINI, Nicolo (INFN e Universita Genova (IT))

Session Classification: T4 - Data handling

Track Classification: Track 4 - Data Handling