Contribution ID: **10**   Type: **not specified**

ATLAS ADC analytics

Tuesday 30 May 2017 09:20 (20 minutes)

The ATLAS Analytics effort is focused on creating systems which provide ATLAS Distributed Computing (ADC) with new capabilities for understanding distributed systems and overall operational performance. These capabilities include to correlate information from multiple systems (PanDA, Rucio, FTS, Dashboards, Tier0, PilotFactory, …), predictive analytics to execute arbitrary data mining or machine learning algorithms over raw and aggregated data, the ability to host new third party analytics services on a scalable compute platform, to satisfy a variety of use cases for different user roles for ad-hoc analytics, and to provide an open platform with documented collections and tools. ADC Analytics is hosted on two backends: ElasticSearch and HDFS. We use Jupyter and Zeppelin as web-based notebooks for advanced analytics, (dist-)Keras+Tensorflow for machine learning, and Pig+Spark for HDFS batch computation. With this talk we will detail the usage numbers and our projections and expectations for the future.

Authors: LASSNIG, Mario (CERN); VUKOTIC, Ilija (University of Chicago (US))

Presenters: LASSNIG, Mario (CERN); VUKOTIC, Ilija (University of Chicago (US))

Session Classification: Going beyond relational

Track Classification: Going beyond relational