Contribution ID: **131** Type: **Oral**

Using machine learning algorithms to forecast network and system load metrics for ATLAS Distributed Computing

Thursday 13 October 2016 12:00 (15 minutes)

The increasing volume of physics data is posing a critical challenge to the ATLAS experiment. In anticipation of high luminosity
physics, automation of everyday data management tasks has become necessary. Previously many of these tasks required human
decision-making and operation. Recent advances in hardware and software have made it possible to entrust more complicated duties to
automated systems using models trained by machine learning algorithms.
In this contribution we show results from three ongoing automation efforts. First, we describe our framework for Machine Learning
as a Service. This service is built atop the ATLAS Open Analytics Platform and can automatically extract and aggregate data, train
models with various machine learning algorithms, and eventually score the resulting models and parameters. Second, we use these
models to forecast metrics relevant for network-aware job scheduling and data brokering. We show the characteristics of the data
and evaluate the forecasting accuracy of our models. Third, we describe the automation of data management operations tasks. The
service is able to classify and cluster run-time metrics based on operational needs. The operator is notified upon a significant
event, and potential resolutions are proposed. The framework learns the decisions of the operator through reinforcement algorithms
over time, yielding better classification of events and proposals for notification or automated resolution.

Tertiary Keyword (Optional)

Network systems and solutions

Secondary Keyword (Optional)

Distributed data handling

Primary Keyword (Mandatory)

Artificial intelligence/Machine learning

Author: LASSNIG, Mario (CERN)

Presenter: LASSNIG, Mario (CERN)

Session Classification: Track 4: Data Handling

Track Classification: Track 4: Data Handling