



Contribution ID: 79

Type: **Oral**

Operational Intelligence

Tuesday, November 5, 2019 2:00 PM (15 minutes)

In the near future, large scientific collaborations will face unprecedented computing challenges. Processing and storing exabyte datasets require a federated infrastructure of distributed computing resources. The current systems have proven to be mature and capable of meeting the experiment goals, by allowing timely delivery of scientific results. However, a substantial amount of interventions from software developers, shifters and operational teams is needed to efficiently manage such heterogeneous infrastructures. For instance, every year thousands of tickets are submitted to ATLAS and CMS issue tracking systems, hence further processed by the experiment operators. On the other hand, logging information from computing services and systems is being archived on ElasticSearch, Hadoop, and NoSQL data stores. Such a wealth of information can be exploited to increase the level of automation in computing operations by using adequate techniques, such as machine learning (ML), tailored to solve specific problems. ML models applied to the prediction of intelligent data placements and access patterns can help to increase the efficiency of resource exploitation and the overall throughput of the experiments distributed computing infrastructures. Time-series analyses may allow for the estimation of the time needed to complete certain tasks, such as processing a certain number of events or transferring a certain amount of data. Anomaly detection techniques can be employed to predict system failures, leading for example to network congestion. Recording and analyzing shifter actions can be used to automate tasks such as submitting tickets to support centers, or to suggest possible solutions to repeating issues. The Operational Intelligence project is a joint effort from various WLCG communities aimed at increasing the level of automation in computing operations. We discuss how state-of-the-art technologies can be used to build general solutions to common problems and to reduce the operational cost of the experiment computing infrastructure.

Consider for promotion

Yes

Primary author: ATLAS AND CMS COLLABORATIONS

Presenter: DI GIROLAMO, Alessandro (CERN)

Session Classification: Track 3 – Middleware and Distributed Computing

Track Classification: Track 3 – Middleware and Distributed Computing