Contribution ID: **12** Type: **not specified**

Anomaly Detection in the CERN Cloud Infrastructure

Thursday, 18 March 2021 16:25 (25 minutes)

Anomaly Detection in the CERN Openstack Cloud is a challenging task due to the large scale of the computing infrastructure and the large volume of data to monitor.

The current solution to spot anomalous server machines in the cloud infrastructure relies on a threshold-based alarming system carefully set by the system managers on the performance metrics of each infrastructure component. The goal of this work is to explore fully automated and unsupervised machine learning solutions in the Anomaly Detection field. We exploit the current state-of-the-art solutions, including both traditional Anomaly Detection and Deep Anomaly Detection approaches.

This contribution will firstly describe the end-to-end data analytics pipeline that has been implemented to digest the large amount of monitoring data and expose anomalies to the system managers. The pipeline uses open source tools and frameworks, such as Spark, Apache Airflow, Kubernetes, Grafana, Elasticsearch. In addition, the performance of the aforementioned Anomaly Detection algorithms will be discussed.

Speaker release

Yes

Desired slot length

Primary author: GIORDANO, Domenico (CERN)

Presenter: GIORDANO, Domenico (CERN)

Session Classification: Grid, Cloud & Virtualisation

Track Classification: Grid, Cloud & Virtualisation