Contribution ID: **166**   Type: **Short Talk**

The ESCAPE Data Lake: The machinery behind testing, monitoring and supporting a unified federated storage infrastructure of the exabyte-scale

Thursday, 20 May 2021 10:50 (13 minutes)

The EU-funded ESCAPE project aims at enabling a prototype federated storage infrastructure, a Data Lake, that would handle data on the exabyte-scale, address the FAIR data management principles and provide science projects a unified scalable data management solution for accessing and analyzing large volumes of scientific data. In this respect, data transfer and management technologies such as Rucio, FTS and GFAL are employed along with monitoring enabling solutions such as Grafana, Elasticsearch and perfSONAR. This paper presents and describes the technical details behind the machinery of testing and monitoring of the Data Lake –this includes continuous automated functional testing, network monitoring and development of insightful visualizations that reflect the current state of the system. Topics that are also addressed include the integration with the CRIC information system as well as the initial support for token based authentication / authorization by using OpenID Connect. The current architecture of these components is provided and future enhancements are discussed.

Primary author: DONA, Rizart (CERN)

Co-author: DI MARIA, Riccardo (CERN)

Presenter: DONA, Rizart (CERN)

Session Classification: Monitoring

Track Classification: Distributed Computing, Data Management and Facilities