Contribution ID: **154**   Type: **Short Talk**

ESCAPE Data Lake: Next-generation management of cross-discipline Exabyte-scale scientific data

Wednesday, 19 May 2021 11:03 (13 minutes)

The European-funded ESCAPE project (Horizon 2020) aims to address computing challenges in the context of the European Open Science Cloud. The project targets Particle Physics and Astronomy facilities and research infrastructures, focusing on the development of solutions to handle Exabyte-scale datasets. The science projects in ESCAPE are in different phases of evolution and count a variety of specific use cases and challenges to be addressed. This contribution describes the shared-ecosystem architecture of services, the Data Lake, fulfilling the needs in terms of data organisation, management, and access of the ESCAPE community. The Pilot Data Lake consists of several storage services operated by the partner institutes and connected through reliable networks, and it adopts Rucio to orchestrate data management and organisation. The results of a 24-hour Full Dress Rehearsal are also presented, highlighting the achievements of the Data Lake model and of the ESCAPE sciences.

Primary author: Dr DI MARIA, Riccardo (CERN)

Co-author: DONA, Rizart (CERN)

Presenter: Dr DI MARIA, Riccardo (CERN)

Session Classification: Storage

Track Classification: Distributed Computing, Data Management and Facilities