



Contribution ID: 204

Type: **Talk**

Advancing Large-Scale Scientific Collaborations with Rucio: A Data Management Story

Tuesday 22 October 2024 16:51 (18 minutes)

Managing the data deluge generated by large-scale scientific collaborations is a challenge. The Rucio Data Management platform is an open-source framework engineered to orchestrate the storage, distribution, and management of massive data volumes across a globally distributed computing infrastructure. Rucio meets the requirements of high-energy physics, astrophysics, genomics, and beyond, pioneering new ways to facilitate research at the exabyte-scale.

This presentation introduces Rucio, highlighting its key features and strategic roadmap that underscore its flexibility towards diverse scientific domains, deep diving into concrete operational experience from various EU projects (ESCAPE, DaFab, InterTwin).

A special emphasis will be placed on the contributions of the CERN IT department, whose active engagement with the Rucio project has increased recently and catalysed significant contributions to the core software. This collaboration has not only enhanced Rucio's capabilities but also solidified its role in LHC experiments such as ATLAS and CMS, and provided a path forward for SMEs (Small and Medium experiments) to benefit from a converged data management platform.

Primary author: GONZALEZ LABRADOR, Hugo (CERN)

Co-author: BARISITS, Martin (CERN)

Presenter: GONZALEZ LABRADOR, Hugo (CERN)

Session Classification: Parallel (Track 1)

Track Classification: Track 1 - Data and Metadata Organization, Management and Access