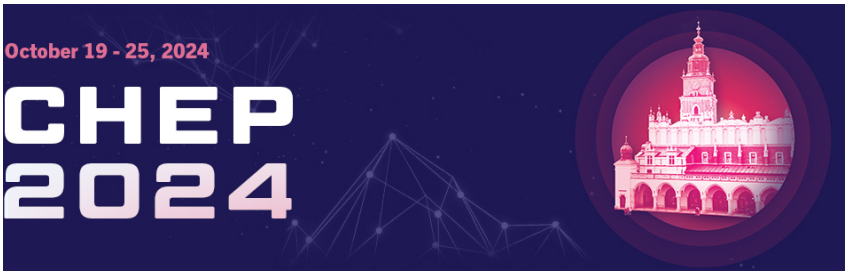

Contribution ID: **373** Type: **Talk**

DUNE Rucio development and monitoring

Wednesday 23 October 2024 13:48 (18 minutes)

The Deep Underground Neutrino Experiment (DUNE) is scheduled to start running in 2029, expected to record 30 PB/year of raw data. To handle this large-scale data, DUNE has adopted and deployed Rucio, the next-generation Data Replica service originally designed by the ATLAS collaboration, as an essential component of its Distributed Data Management system.

DUNE's use of Rucio has demanded the addition of various features to the Rucio code base, both specific functionality for DUNE alone, and more general functionality that is crucial for DUNE whilst being potentially useful for other experiments. As part of our development work, we have introduced a "policy package" system allowing experiment-specific code to be maintained separately from the core Rucio code, as well as creating a DUNE policy package containing algorithms such as logical to physical filename translation, and special permission checks. We have also developed other features such as improved object store support, and customisable replica sorting. A DUNE-specific test suite that will run on GitHub Actions is currently under development.

Recently, DUNE has deployed new internal monitoring to Rucio, enabling us to extract more useful information from core Rucio servers, and daemons such as transmogrifier, reaper, etc. Additionally, DUNE has implemented monitoring for Rucio transfer and deletion activities which are sent to a Message Queue via Rucio Hermes daemon. Information such as data location, accounting, and storage summary is extracted from the Rucio internal database and dumped into Elasticsearch for visualisation. The visualisation platforms utilised are based at Fermilab and Edinburgh. This monitoring is crucial for the ongoing DUNE data transfers and management development.

Primary authors: WHITE, Brandon (FNAL); PERRY, James; YUAN, Wenlong (The University of Edinburgh (GB))

Presenter: YUAN, Wenlong (The University of Edinburgh (GB))

Session Classification: Parallel (Track 1)

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