Data Management Development

Steven Timm / Doug Benjamin / Wenlong Yuan

6 Jan 2025

Main areas of emphasis for 2025

- Development

  - Getting Rucio to work with Tokens. (James Perry dev., Doug Integration)

  - Consistency Checking (Yuyi Guo)

  - SAM Retirement (Steve + developers to be named)

- Operations

  - NP02 raw data running

  - Data Challenge 2025. (DUNE only)

🔷 Fermilab DUNE

Rucio with CILogon tokens

- James Perry collaborating with Anil Panta from JLAB which is also using CILogon tokens

- 5 major authentication "flows"

  - Admin (3rd party copy and delete, credentials used by Rucio daemons to talk to storage and FTS3)

  - User Auth (how users authenticate to the Rucio server)

  - Rucio download

  - Rucio upload

  - Tape staging

- First 2 already work in Rucio with IAM tokens and some progress made in CILogon

- Second 3 yet to be implemented by Rucio devs.

- More updates tomorrow at DM meeting and general Rucio meeting on Jan 9

🟦 **Fermilab** DUNE

Consistency Checking

- 2 Major consistencies:

  - Rucio vs. individual storages

  - Metacat vs. Rucio

- For Rucio vs. individual storages, Yuyi is evaluating a package that Igor Mandrichenko wrote for CMS

- Metacat vs. Rucio consistency checking pending the entry of the remaining legacy SAM locations into Rucio.

Fermilab DUNE

SAM retirement

- 3 Major things that need to be done:

  - Developing an alternative to Sam4Users

  - Surveying current SAM usage and finding users that are running on legacy datasets and making sure those datasets are all available in the new system

  - Bringing forward all legacy SAM location information into Rucio

- Plus need an unambiguous designation in Metacat for files that are virtual files, i.e. files that were only declared to show parentage but the file was never stored with a location.

Fermilab DUNE

DC25

- Want to get this done before the NP02 beam run

- Was originally thought we would do it in February but will probably push to April due to all the data challengers also being involved in the DOE Review.

- Goals to simulate a data pipeline and remote keep-up processing on ProtoDUNE-VD data.

🔷 Fermilab   DUNE