

Building an improved user experience

ACAT 2022, Bari

Integration of a user-dedicated CLI capable of opportunistic data registration and user file upload and registration.

Abstract

With a growing experience on managing CMS data via RUCIO, we started to work on how to make it more user-friendly. The priority was on two main use cases in mind, but the idea is to iterate over users feedback and new use cases.

Context

Temp RSEs are non-deterministic RSEs for that are used as special areas where Users and Rucio have W and RW permissions respectively, in contrast to `cms_type:real` Rses, where Users and Rucio have R and RW permissions, as per the CMS Permission model. [R-Read, RW-Read and Write]

As a result:
Users cannot directly upload data to `real` RSEs directly by a simple rucio upload.
Thus creating the need for a two step data injection process where we use the Temp RSEs as Buffer.

Methodology



Use of *_Temp RSEs as the temporary staging area

- Use user proxy to upload files to Temp RSEs



Ask rucio to move it to desired RSE

- Create rules to desired RSE where user has quota

Local data Upload and Registration

Uploading a local file to RSE

Provide the commands with parameters for new rule creation and location of file to upload.

Two flavours:

- Single File Upload
- Dataset Upload (Multiple files grouped under a dataset)

Remote Data Registration

Moving Tau Embedded Samples to global storage

These are produced on opportunistic HPC resources by users.

After validation, provided to entire CMS collaboration for usage.

- > We use group accounts in Rucio to park these samples under
- > Run script to:

- register data in Rucio
- publish in CMS Data Book Keeping System (DBS) prod/global

CLI Usage

```
[rchauhan@lxplus774 ~]$ cmsrucio-import --help
Usage: cmsrucio-import [OPTIONS] COMMAND [ARGS]...
```

Options:

- install-completion [bash|zsh|fish|powershell|pwsh] Install completion for the specified shell.
- show-completion [bash|zsh|fish|powershell|pwsh] Show completion for the specified shell, to copy it or customize the installation.
- help Show this message and exit.

Commands:

- generate-yaml-template
- upload-dataset Upload a files specified in the folder and...
- upload-dataset-yaml Same as upload dataset Arguments specified in...
- upload-file Upload a single file and create a rule for rse
- upload-file-yaml Same as upload file Arguments specified in yaml...

Summary

The new interface makes it possible for users to upload data from their local systems.

It also allows registration of opportunistically produced datasets from computing resources outside experiment specific infrastructure.

Enabling to move toward a model where user data are owned and managed by RUCIO.

- A Possible shift in quota management from physical at each site, to a centrally managed logical quota

Future Plans

Incorporate optional DBS (A CMS Bookkeeping Service) registration in opportunistic data registration and transfer.

Use FTS for a Third Party Transfer of files for initial registration in Rucio.

Upstream these changes in Rucio to understand if other communities can benefit from similar injection models.

[Diego Ciangottini](#) (INFN, Perugia)
[Eric Wayne Vaandering](#) (FNAL)
[Kevin Patrick Lannon](#) (Notre Dame, IN)
[Rahul Chauhan](#) (CERN)

For the CMS Collaboration