

Theme / Date

Overall goals (then we set an ETA...)

ACTIVITY 1

Data Injection, Policies and Replication

Data Processing (Large scale/ data preparation)

Data Processing (Analysis: platforms, users, citizen science)

Leads: Yan, Riccardo, Frederic(?)

DAC21: More than three Analysis facilities/ Science platforms integration with the Data Lake

STATUS OR INPROGRESS RISK

DAC21: Data Management: define and test data acquisition, injection, replication, lifecycles per experiment

DAC21: Data registration and/or buffering from remote sources (non-RUCIO, unidirectional RSEs, etc.)

STATUS RISK

DAC21: Data Processing: Production (= experiment "data preparation"). Use cases: reprocessing, reconstruction, calibration, etc. requiring input and output from/to the Data Lake.

DAC21: Heterogeneous resources integration Clouds and HPCs (caches, ephemeral RSEs,...)

DAC21: Software repos, image registry, calibration files=>WP3

DAC21: Mock exercise (DR) per experiment

DAC21: Use case workflow scripts ready to test (august)

10) Real-life QoS classes, including tape

Containers and Open Data

External resources, data lake integration with clouds: cpu and storage,

DLaaS->http solution

DIRAC/WP5 for large scale processing

HW at the sites (QoS mappings)

CS3MESH4EOSC

Overview of use case needs (aka "the Matrix") (ASAPI!)

Concrete demonstration deliverables, and requirements per UC (May/June)

Link between ESAP and Rucio (find data in data lake) -> depend on Tokens, but crude version is in a branch (june)

Rucio access from Jupyter notebook launched from ESAP, sharing the token (June/July)

Leveraging site caches

ACTIVITY 2

Rucio technology deployment , Multiple Rucio instances

Continuous testing and Monitoring

Leads: Rosie, Agustin(?), Rizart, DepOps team

1) Use of metadata

6) Rucio subscriptions

15) Rucio Collaboration strengthen

8) Rucio development - rules with date-specified start and end dates

9) Data corruption testing

Rucio K8s cluster monitoring

3) Testing (monitoring) improvements

2) File deletion - understand behaviours

4b) Experiment-specific policies

Continued NW BW tests

12) Complete Experiment Onboarding

5) QoS policy mapping to QoS classes

17) Rucio Documentation Improvements

Data centric dashboard view

ACTIVITY 3

AAI / Token Integration/ ACLS

Networks

Leads: Andrea, Paul, Marek(?)

Streaming behaviour study

long haul Rucio behaviour tests- small and large scale files / data sets

FTS alternative - long haul

7) Token-based authentication and authorization

4a) Support for embedded data, including multi-VO Rucio

Continued PerSNAAR monitoring and reporting

1) HA IAM Service

Investigation / Statement of support for multiple small communities - c.f. EOSC