

Use Case 1 Long-haul Transfer

Test Number	Description	Total Volume	File Size	Number of files	Configuration
Test 1	Automatic detection and transfers from La Palma to PIC. Then automatic deletion	300GB	1.5GB	190	LP(GridFTP)2PIC Disk2Disk
Test 2	Test1: with XrootD	300GB	1.5GB	190	LP(XrootD)2PIC Disk2Disk
Test 3	Tests 1: More data	10TB	1.5GB	8000	LP(GridFTP)2PIC Disk2Disk
Test 4	Test 3: Even more Data	40TB	1.5GB	32000	LP(GridFTP)2PIC Disk2Disk
Test 5	Test 3: with Tape Storage	10TB	1.5GB	8000	LP(GridFTP)2PIC Disk2Tape
Test 6	Replication at CNAF	10TB	1.5GB	8000	LP(GridFTP)2PIC > PIC2CNAF Disk2Disk

Six tests defined 4-5 completed successfully (see table below).

Amount of data transferred in <1day is comparable to the volume required by CTA.

The inclusion of Tape storage meets an important goal in the CTA bulk data management plan.

Technical difficulties in deletion at the source and replication in CNAF.

More tests planned next year, including some new tests regarding file size and priority data



Use Case 1 Long-haul Transfer

Test name	Protocol	Estimated #Files, data volume	STORAGE	Replicas/RSEs	Results	Observations
Test 1	GSIFTP	300 GB	Disk	CTA-RUCIO: non-deterministic and deterministic RSEs	Completed successfully.	
Test 2	XROOTD	300 GB	Disk		Completed successfully.	
Test 2.1	XROOTD	10 TB	Disk		Completed successfully.	Initially not planned, is a mix of Tests 2 and 3
Test 3	GSIFTP	10 TB	Disk		Completed successfully.	
Test 4	GSIFTP	40 TB	Disk		Failed. Completed in more time than stated	Problems with the Rucio server side transferring large number of files.
Test 5	GSIFTP	10 TB	Tape		Completed successfully.	
Tes 6	GSIFTP	10 TB	Disk		Failed. Completed internally.	We couldn't run the test on production environment since we are experiencing problems to connect from PIC to CNAF due to an update of Grid certificates in both sides, changes in the CA



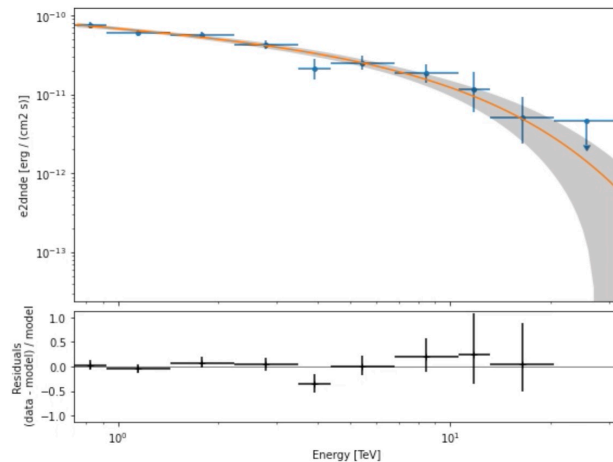
Use Case 2: Reprocessing (DIRAC-RUCIO)

- Raw (DL0) data is identified on tape (obsid or time range) via **metadata** e.g. using getMetaData method
- Data volume is calculated
- Data is staged from tape (COLD) storage to temporary disk (HOT)
- Data is reprocessed using CTA pipeline software
- Final data products (DL1) are verified
- Cache and temporary data is cleared
- Ingest the resulting new DL1 data into the datalake
- Delayed due to initial tech difficulties and scheduling.
- The number of methods available in DIRAC to manage the catalogue is much larger than that of Rucio. It may be necessary to implement new missing methods.
- Understanding the cache in dCache will also be important wrt to accessing the data on Tape.
- Upgrade of Rucio from v1.23 to v.1.26 may solve some issues.
- Work planned to continue in January



Use Case 4 Analysis

- User logs in to the ESAP and is identified
- Search for Data in the datalake
- Search for corresponding IRF (instrument response function) for the data selected
- Search for software via OSSR
- The data can now be analysed in interactive mode using Jupyterhub



- Analysis completed on both ESAP (via “OSSR” & Jupyterlab/mybinder) and DLaaS
- ESAP shopping basket would need to be adapted to accommodate files of this type, replicate to http download service or the workflow should change towards IVAO TAP like services.
- DLaaS requires a preloaded container or container that can install software on the fly.



Summary

- The results of UC1 are already very interesting from a CTAO construction perspective
- Several important requirements have been met. Several more are within reach
- We will continue next year with further tests, which will aid in the construction project
- Within striking reach of a telescope-to-user chain, using ESCAPE tech

Agustin Bruzzese
Jordi Delgado
Frederic Gillardo
Luisa Arrabito (DIRAC)
Nadine Neyroud
Gonzalo Merino
Cedric Serfon
Matthias Füßling
Nektarios Benekos
Karl Kosack

