



Science & Technology
Facilities Council

UK Research
and Innovation

Using and enhancing WLCG data management software to support other communities

Ian Collier

On behalf of Andrew Lister, Ian Johnson, Alastair Dewhurst

Data Management

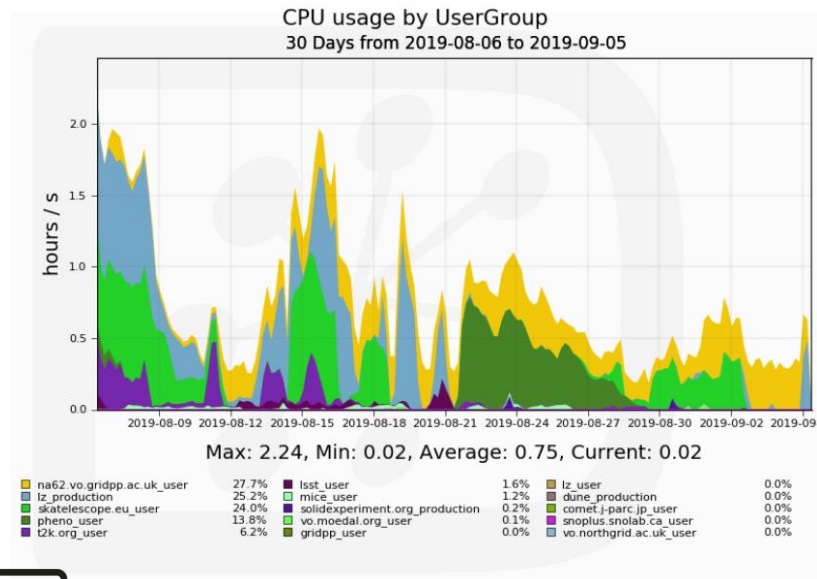
- Creating lots of data is no longer restricted to huge experiments.
 - Can't expect VOs to have dedicated computing teams.
- Experimental data is expected to be stored 'forever'
 - Mistakes/Failures can be extremely costly.
 - Have to assume software will change.

Motivation for Rucio

- Requirements:
 - Must be able to meet the needs of a diverse community.
 - Must be scalable.
 - Must have clear long term support.
 - Must be able to work with a variety of storage backends.
 - Should build on our existing areas of knowledge.
- Rucio is the only service that meets these requirements.

UK support for Non-LHC VOs

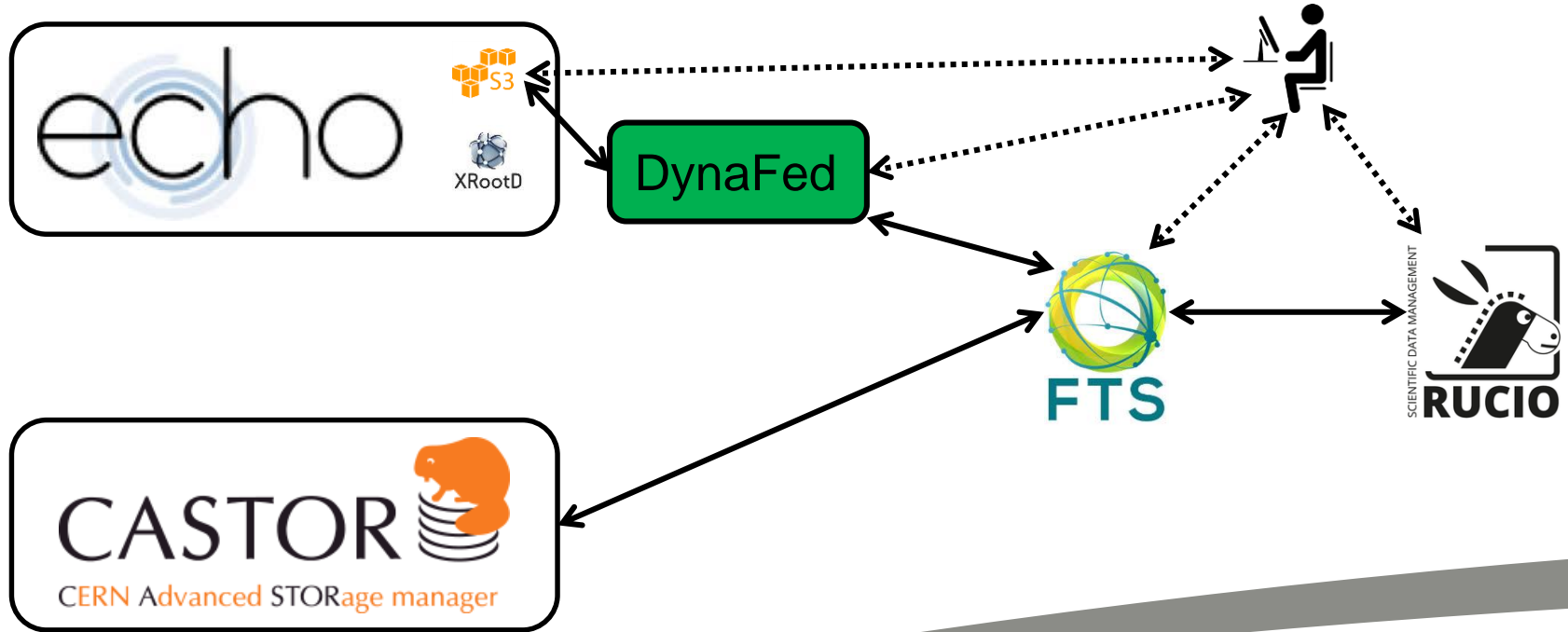
- The UK supports a rapidly growing community of other experiments.
 - HEP, Astronomy and Space.
- The UK runs a multi-VO DIRAC instance to allow non-LHC VOs to utilize UK grid resources.
- Although DIRAC is used by e.g. LHCb for all of their computing needs, development resources are limited.



Continue to use DIRAC for workflow management.
Use Rucio for data management.

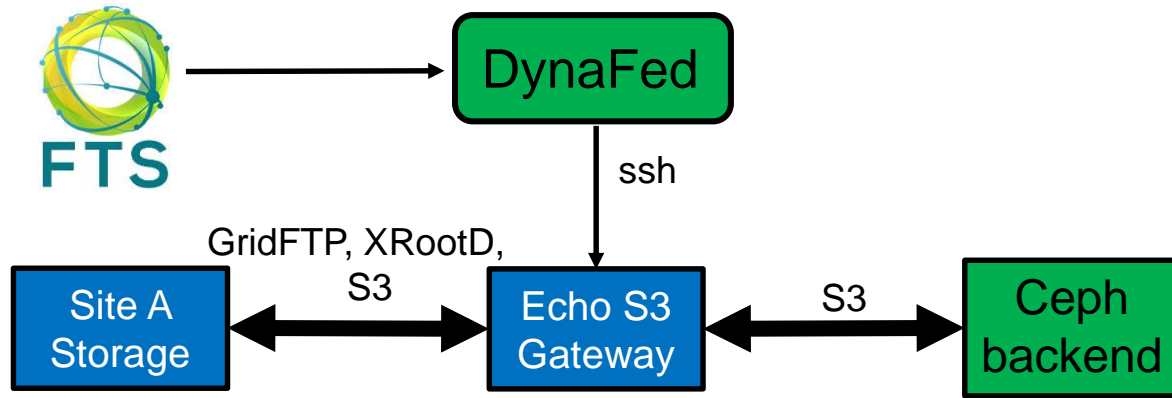


RAL Tier-1



DynaFed and S3

7



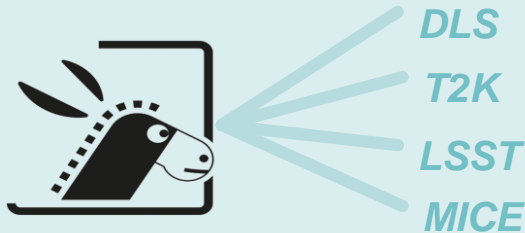
- DynaFed allows Cloud based storage to appear very much like a Grid storage endpoint.
- Can use DynaFed's "4th party transfers" to transfer data between sites

Multi-VO Rucio

Single-VO



Multi-VO



A single instance capable of supporting many VOs. Compared to multiple instances it:

- Saves on maintenance and setup times
- Allows for adding new VOs on the fly – at lower ‘cost of entry’
- Offer as a service to any new experiments

Majority of code added to Rucio 1.20

- Over 8000 lines of code changed.
- Full functionality expected in 1.22.

Usage (User – Current)

An example of using Rucio for a typical (single-VO) user

Authentication

- Add Account, and credentials to “rucio.cfg” config file.

List Files

```
rucio ls user.jdoe:*
```

- Use Rucio CLI

Upload Files

```
rucio upload --pfn ~/a_file.txt --rse endpoint1 ...
```

- Use Rucio CLI

Usage (User – Multi-VO)

An example of using Multi-VO Rucio for a typical user

Authentication

- Add Account, **VO**, and credentials to “rucio.cfg” config file.

List Files

```
rucio ls user.jdoe:*
```

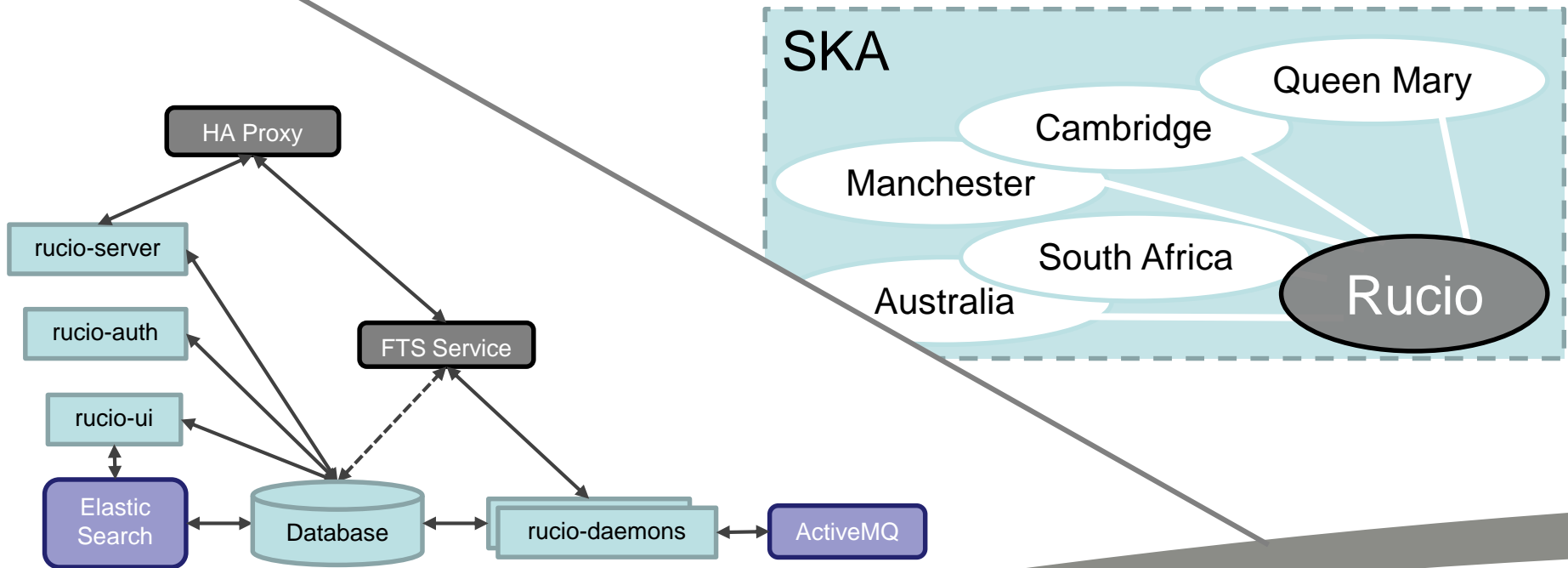
- Use Rucio CLI

Upload Files

```
rucio upload --pfn ~/a_file.txt --rse endpoint1 ...
```

- Use Rucio CLI

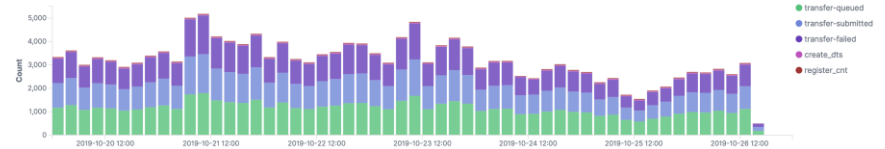
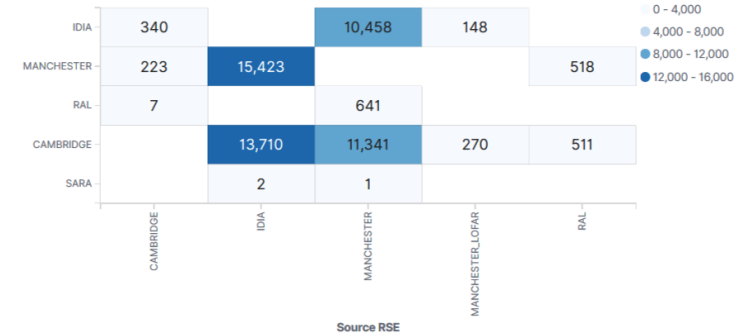
Rucio@RAL



Rucio Status and Plans

- The multi-VO Rucio instance at RAL in final testing.
 - Areas under scrutiny are usability and isolation
- Plans:
 - Adapt two remaining daemons
 - Demo Multi-VO at RAL for integration tests
 - Supported Multi-VO Rucio service running at RAL
 - Iterate permissions scheme as users join

Count Matrix - completed transfers



PaNOSC

- STFC are providing 4PB of storage in Echo (S3).
 - 1PB a year for the 4 years of the project.
 - Effort to support and develop software.
- Initial goal to transfer data from ILL to RAL (and DESY).
 - Preferably via FTS transfers.
- Integrating FTS with EOSC Hub
- Rucio may well be a very good fit for their data management needs.

Summary

- RAL is looking at long term solutions to common data management problems.
 - The UK is incorporating Rucio into its strategy to support many small VOs centrally from RAL Tier 1.
- RAL runs a stable production ready Rucio instance.
 - Development for multi-VO instance of Rucio is virtually complete.

Questions?