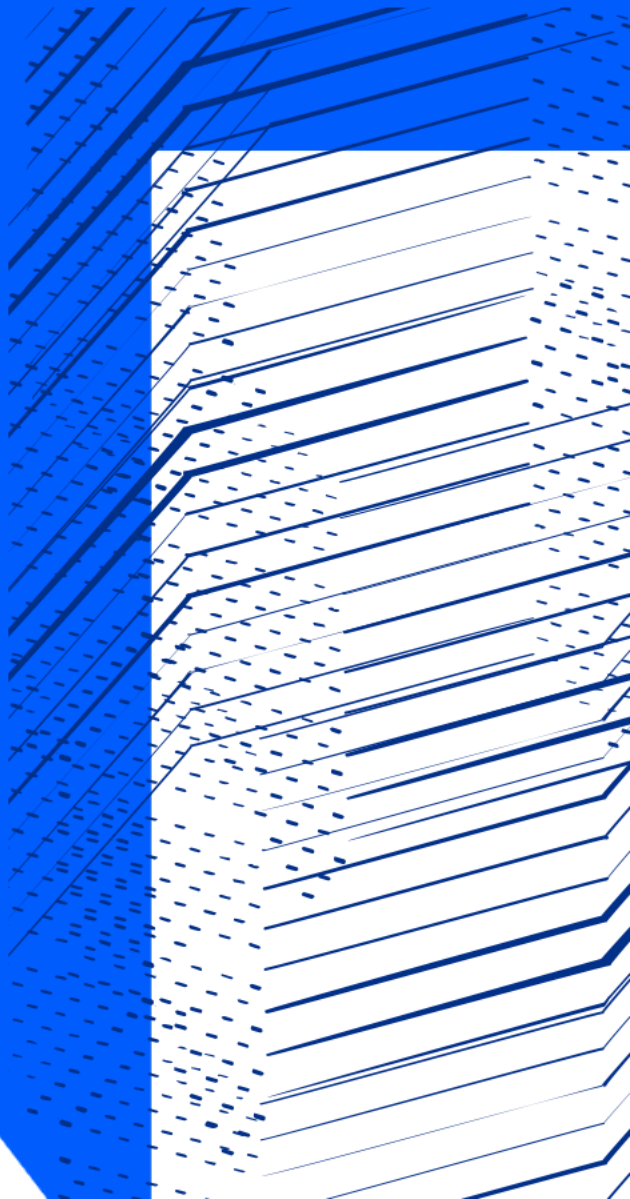
Science and Technology Facilities Council

Multi-VO Rucio and Cloud Storage at RAL

Ian Johnson



Storage strategy

- GridPP6 strategic objective:

  "***To provide broader benefit to STFC and their communities by continuing established initiatives to reduce the operational cost of the infrastructure, whilst increasing support for non-LHC communities and developing common infrastructure and operations.***"

- Grid Storage APIs and Security model are a huge barrier for non-LHC communities to effectively use storage.
- Object storage using Cloud APIs is generating a lot of interest
  - Gives experiments a choice of providers (see they are getting value for money).
  - Plenty of documentation and existing software to build on
- We do not want to end up with two incompatible storage / data management solutions
  - Actively working to ensure that Cloud and Grid Storage can be used together.

Cloud Storage - Security Challenge

- Cloud APIs typically provide an access and secret key
  - This provides full access to their storage.
- Best practice says that these credentials should not be used to access the storage directly they should be used to generate temporary credentials instead [1]
  - This adds complexity so experiments/developers don't like it!
- It is more important for us to insist on a proper security because if an account is compromised :
  - It is easier to disrupt ECHO than Amazon or Google etc.
  - We are likely to be left clearing up the mess after an incident.

[1] https://docs.aws.amazon.com/IAM/latest/UserGuide/best-practices.html#lock-away-credentials

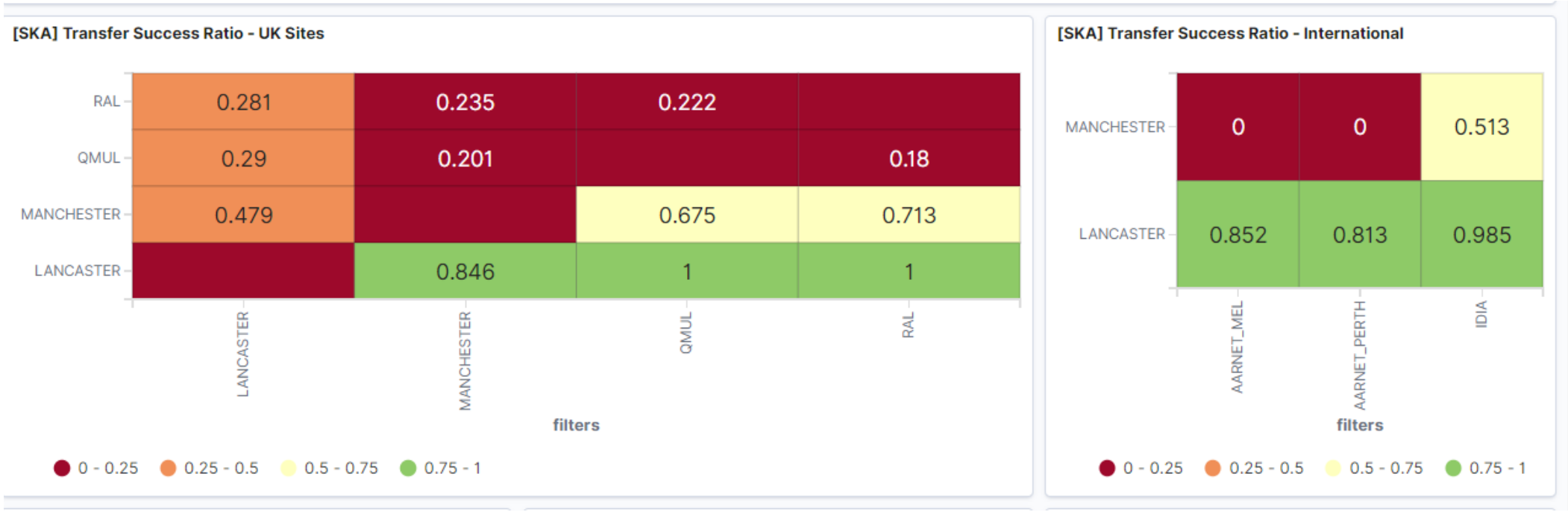
Multi-VO Rucio

Rucio instance at RAL

- Rucio at RAL has been in operation since November 2018, when trial system moved onto SCD's OpenStack cloud.
- Rucio at RAL has been supporting a non-LHC VO since its start – the SKA Organisation uses Rucio to replicate data across a mesh of UK and international storage elements
- A set of functional tests generate replication rules, leading to transfer requests to FTS
- A Kibana dashboard shows transfer activity in the RSE mesh

Monitoring Dashboard

- Dashboard currently linked to transfers for SKA VO
- Separates out international and GridPP transfers
  - Planning on contacting sites to fix RSE configuration errors



Rucio Status

- RAL are now running the multi-VO instance
  - This is an IRIS Digital Asset created by three graduates in SCD working from April 2019 to September 2020 (Andrew Lister, Eli Chadwick, and Patrick Austin)
  - It supports VOs for SKA, DTEAM, and DIRAC (for integration with multi-VO DIRAC)
- VOs are isolated from each other
  - Cannot see each others RSEs, accounts, Rules, etc
  - A user of the multi-VO Rucio instance only needs to specify the name of the VO they are registered with (typically in an environment variable); all Rucio command invocations stay the same as in a single-VO instance

Future Plans

- Recruit additional FTE to work on Rucio
- Integrate solutions developed by others in GridPP:
  - Monitoring – Teng Li and Wenlong Yuan
  - Policy packages for multi-VO Rucio – James Perry
- Fix bugs / missing features that are identified in development with DIRAC integration.
- Improve useability in preparation for becoming a full production service:
  - Documentation
  - WebUI for defining Rucio Rules
  - OIDC integration with the IRIS IAM service produced by Tom Dack

Science and Technology Facilities Council

Cloud Storage Users

Non-LHC use of ECHO Cloud Storage

- STFC are collaborating with ILL and ESRF (both in Grenoble) via the PaNOSC Horizon 2020 project:
  - '**P**hoton and **N**eutron **O**pen **S**cience **C**loud' – a project integrating data and metadata from various research institutions (RIs) with resources from the European Open Science Cloud (EOSC) and other sources
- ILL need a second, offsite location to store their experimental data
  - A 'cold' archive providing data safety, allowing disaster recovery
  - Requirement is for a simple backup-style service
- ECHO provides storage for this archive
  - EGI funds staff effort, IRIS allocates disk resources

ECHO resources for PaNOSC

- Allocation of 1 PB ECHO storage per year of the PaNOSC project, 4 PB in total
- Using ECHO S3 interface, compatible with public cloud storage
- ILL have experimental data from 1973, wish to archive all of this (~ 2 PB)
- Recent cycles produce ~ 80 TB of data, 3 or 4 cycles a year
- ILL have been stalled due to effects of Covid-19, but have recently recruited staff to work as data admins for the archiving task
- STFC are to give hands-on training to ILL data admins Friday

Rclone Workflow

- Rclone was chosen as it is a simple tool that meets requirement.
  - Works like ‘rsync’.
  - Also evaluated Rucio, FTS
- Access/secret keys are in config, which is restricted to (currently) a single host.
- In initial testing, seeing around 200MB/s with multiple file transfers. Further tuning seems possible.

```
[ceph]
type = s3
provider = Ceph
env_auth = false
access_key_id = XXX
secret_access_key = YYY
region =
endpoint = https://ceph.endpoint.example.com
location_constraint =
acl =
server_side_encryption =
storage_class =
```

Other PaNOSC sites considering ECHO

- ESRF are producing ~ 1PB of data a month
  - Too much for the PaNOSC allocation, but they may wish to store a subset of data in ECHO
- The only other site in PaNOSC near readiness is ELI (based in Aachen). They are looking at on-site backup, but may wish to use ECHO after hearing of ILL's experience

Other Communities

- Growing demand and active development in S3 endpoints.
- We are actively collaborating to ensure security concerns are addressed in the experiment's software model.
  - Data export from small number of hosts (e.g. PaNOSC solution just described) is acceptable.
  - Providing (easy) access to users is biggest challenge.

	Allocation (TB)	Usage (TB)	Notes
Diamond	1000	Test Data	S3, Requested 2PB from IRIS. Developing Code.
RFI	TBD	Test Data	S3, Developing Code.
LSST	4250 in 2021	0	S3, Expected to grow to 9PB in 2023

Science and Technology Facilities Council

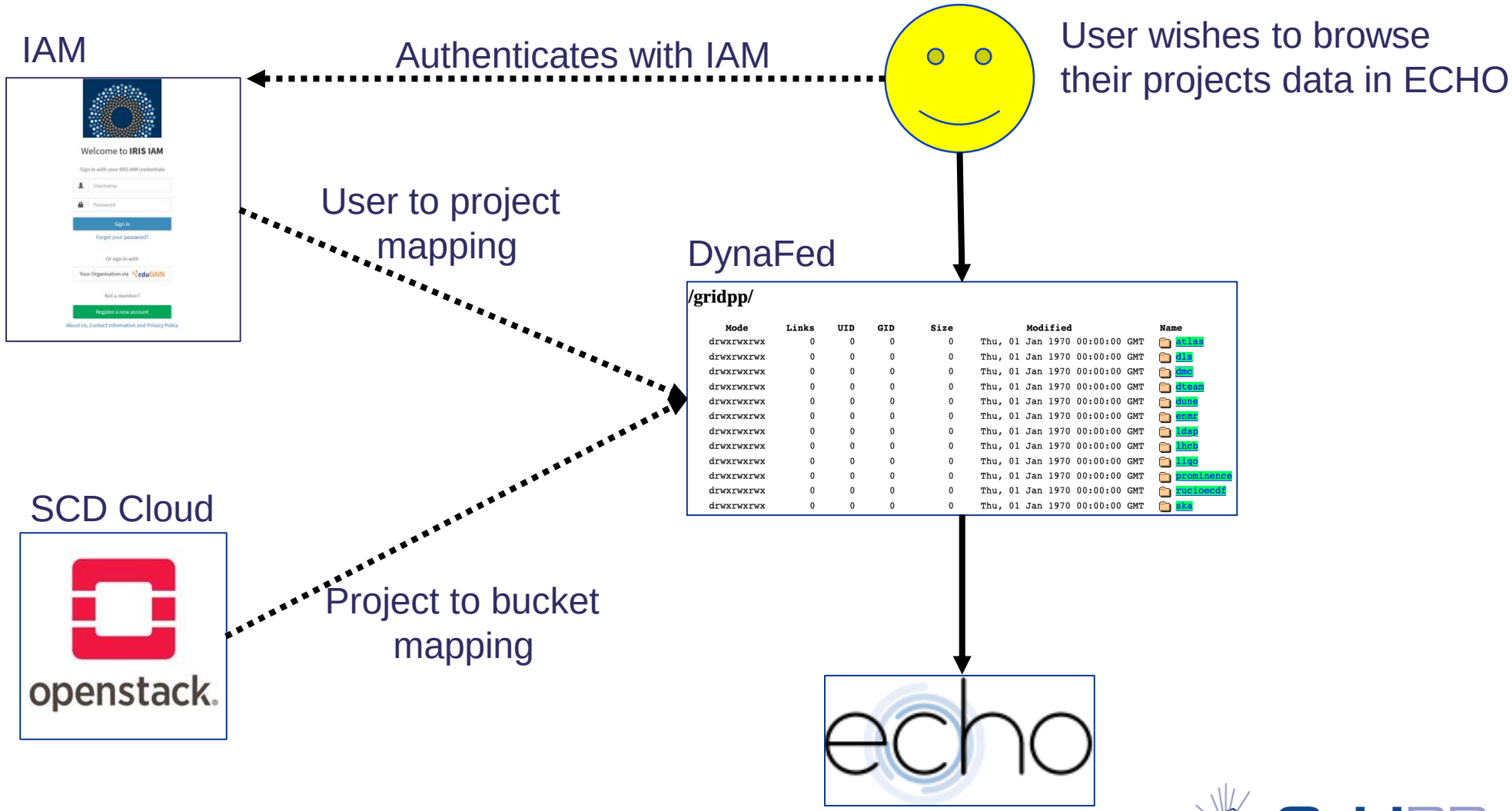
DynaFed development

DynaFed Digital Asset

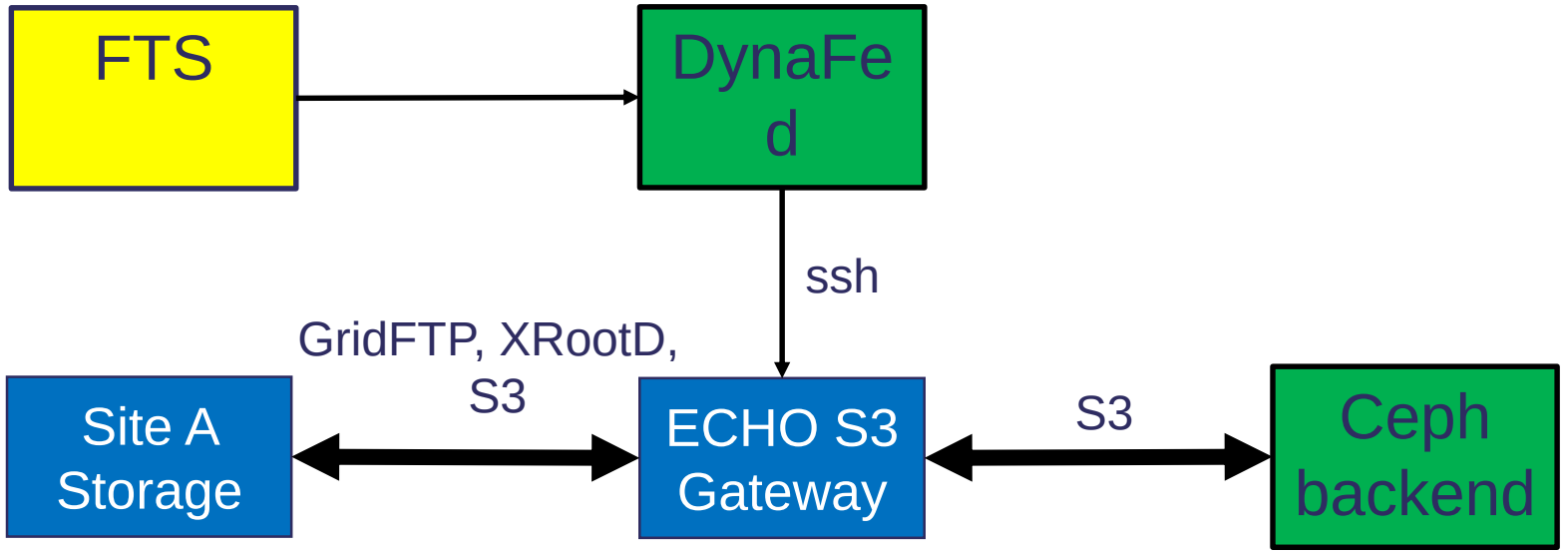
- Sam Glendenning is an SCD graduate working on DynaFed development.
  - Started in September on Task 1.2 as it was seen as highest priority.

Task	Description	When / PM
1.1 Integrate Rucio with DynaFed	Develop scripts that will automatically update DynaFed if configuration changes are made to the Multi-VO Rucio instance at RAL.	February - March 2021 (3 PM)
1.2 Integrate DynaFed with IAM	Enable DynaFed to accept users who have authenticated via the IRIS IAM service.	October - November 2020 (2 PM)
1.3: Enable 4th Party Transfers	Create, deploy, test and validate the necessary scripts to allow 4th party transfers to be performed.	December 2020 - January 2021 (3 PM)

IAM integration



4th Party Copies



- A lot of progress has been made by the DOMA TPC group with WebDAV transfers.
  - But not all sites support WebDAV and not all types of transfers are possible. (e.g. S3 → S3)
- If DynaFed is requested to mediate the transfers it will "stream" it by running a script.
  - We can modify this to ssh into an ECHO gateway and stream the transfer through that.
  - This should offer "full TPC support" for all protocols!

Summary

- Lots of interest from non-LHC communities to use S3 storage endpoints.
- We are continuing to develop and support Rucio.
  - It is vital that it can manage both Grid and Cloud storage service.
- DynaFed development will lower the barrier to using Cloud Storage for all communities.

Science and
Technology
Facilities Council

Questions?