

Multi-VO Rucio

Eli Chadwick, Andrew Lister, Ian Johnson

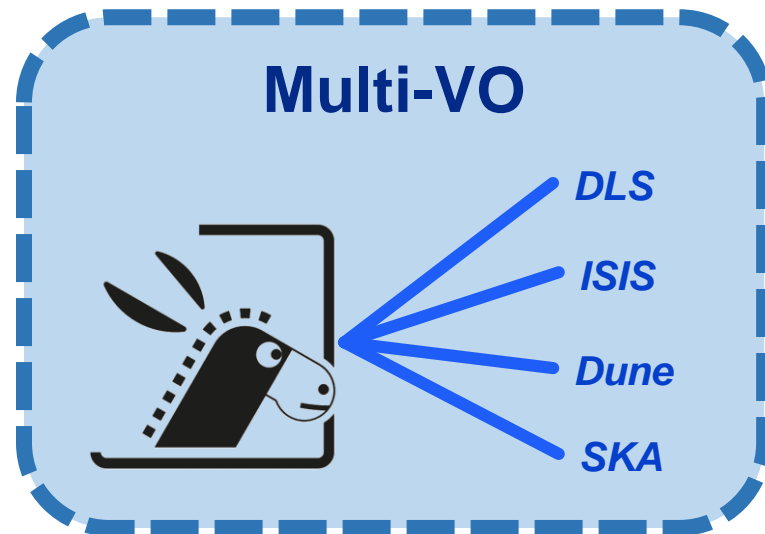
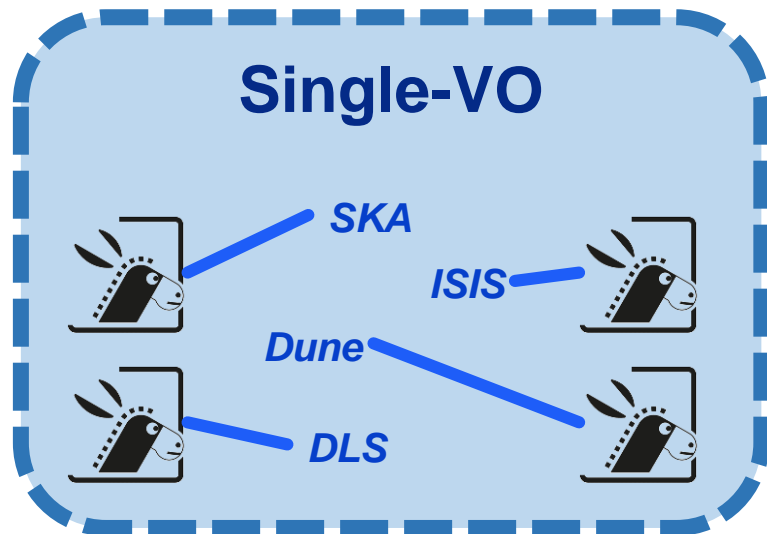
Development Team

- Eli Chadwick – until June 2020
- Patrick Austin – until August 2020
- Ian Johnson
- Andrew Lister – finished September 2019

Alastair Dewhurst (RAL Tier-1 Manager) is seeking funding from multiple bodies to maintain development effort and support long-term maintenance of the RAL Rucio instance

What is multi-VO Rucio?

A single instance capable of supporting smaller VOs



What is the use case?

- STFC supports the HEP, Astronomy and Space science communities
- Many experiments with:
 - Up to 10PB of data
 - Up to 10 million files
 - Very little effort for data management
- RAL provide STFC's only general purpose Archival Tape service
 - Tape is not straightforward to use
 - Migrating from Castor to CTA, so why not teach them to use Rucio?

Current users of RAL who could benefit from Rucio

Experiment	Space Used (TB)	# Files	Comment
DiRAC	2605	140K	Currently evaluating Rucio.
Na62	1017	8M	Tape usage at RAL expected to rise significantly.
T2K	831	5M	Use GridPP DIRAC instance.
ILC	216	3M	Use GridPP DIRAC instance.
SNO+	129	800K	Use GridPP DIRAC instance.
MICE	44	30K	Data taking finished.
MINOS	12	400K	Data taking finished 2012.

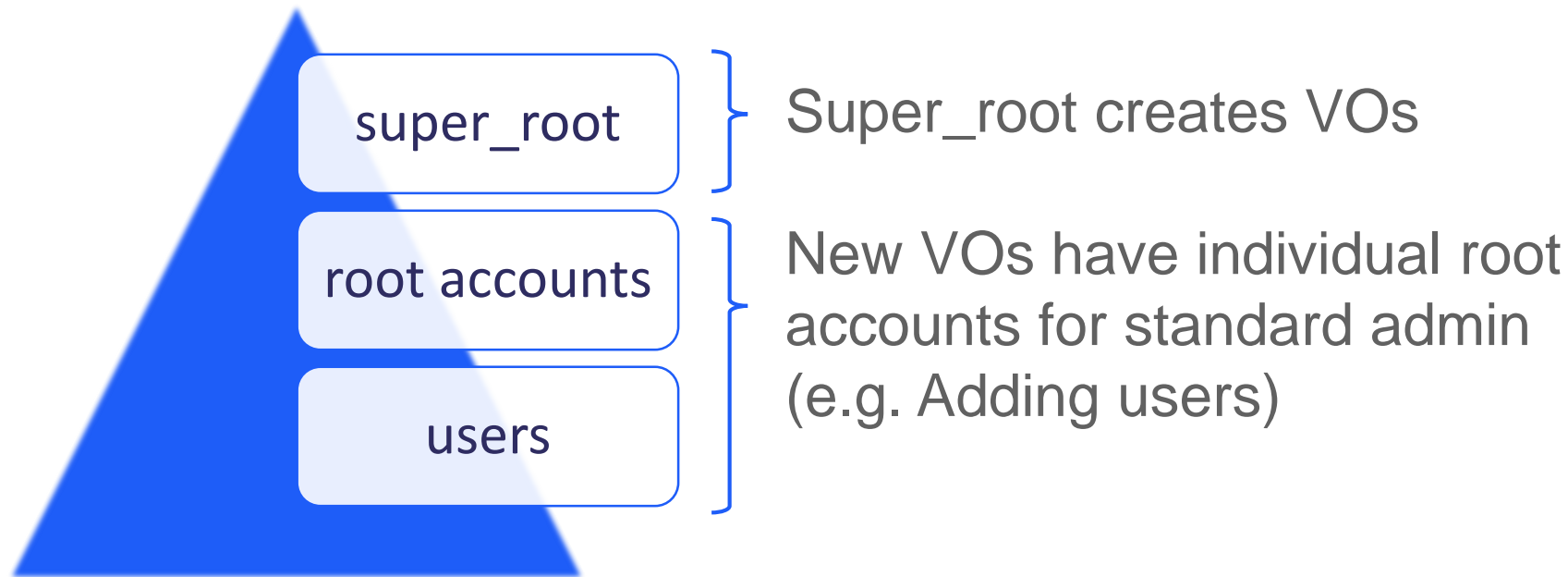
Why use multi-VO?

- Centrally managed service lowers the barrier for new experiments to use Rucio
- Experiments could use Rucio WebUI to interact with their VO
- Intended users are small (compared to ATLAS) so scale should not be an issue
- Easier to provide long-term support for occasional users
- Lower maintenance for host

Why not use multiple Kubernetes pods?

- Looking for an evolution to Rucio at RAL, not a revolution
- Tier-1 at RAL doesn't have enough experience with Kubernetes yet
- It's not a production service on the STFC Cloud yet
- In the future we will look at using Kubernetes to run the multi-VO instance at RAL

How is multi-VO Rucio implemented?



Usage (User – Current)

An example of using Rucio for a typical 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

Development

Split into 4 phases, each ending in a Pull Request:

1. Separate internal and external representations for RSE
2. Internal representations for Account and Scope
3. Database schema updates ← we are here
4. Multi-VO features! ← some progress made

No impact to single-VO instances

Development Phases

1. Separate internal and external representations for RSE
 - Unique internal ID for each RSE
2. Internal representations for Account and Scope
3. Database schema updates
4. Multi-VO features!

Development Phases

1. Separate internal and external representations for RSE
2. Internal representations for Account and Scope
 - account@vo, scope@vo
3. Database schema updates
4. Multi-VO features!

Development Phases

1. Separate internal and external representations for RSE
2. Internal representations for Account and Scope
3. Database schema updates ← we are here
 - Add VO table, add VO column to RSE table
 - Will be in 1.22 release
4. Multi-VO features!

Development Phases

1. Separate internal and external representations for RSE
2. Internal representations for Account and Scope
3. Database schema updates
4. Multi-VO features! ← some progress made
 - Add VO-specific authentication
 - Add VO field to all Rucio functions
 - Daemons

Future Plans

- Finish multi-VO features
- Pre-production by June 2020 with trial users
- Iterate permissions scheme as users join
- Multi-VO monitoring

Thank you