Distributed data management system: Rucio

Farida Fassi
Faculty of Sciences Mohammed V
University in Rabat, Morocco

Computing Training course
I-COOP+ 2016 project: COOPB20247
3-14, July 2017. IFIC-Valencia, Spain

Distributed Data Management system: Rucio Overview

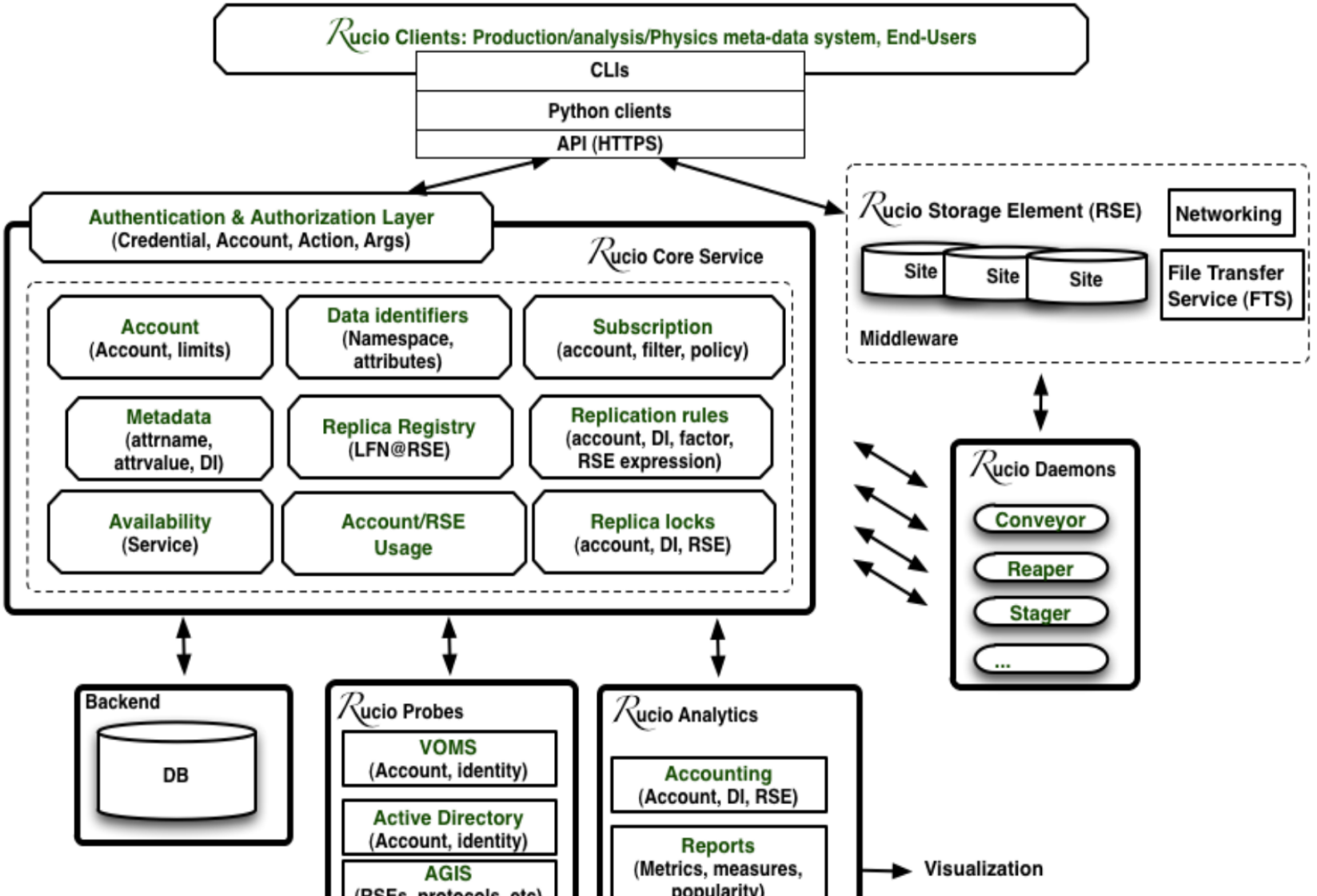
- The ATLAS Distributed Data Management project is charged with managing ATLAS data on the grid
  - **The purpose is helping the collaboration store, manage and process LHC data in a heterogeneous distributed environment**

- **Requirements:**
  - **Catalog data**
  - **Transfer data to/from sites**
  - **Delete data from sites**
  - **Ensure data consistency at sites**
  - **Enforce ATLAS computing model requirements**

Rucio components

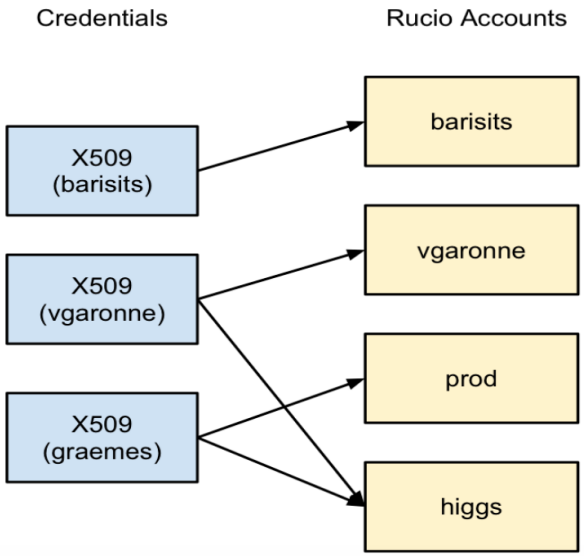
- **The main Rucio components are:**
  - **Catalog of the: datasets, content and location**
  - **Set of daemons to move the data from one site to another**
  - **Manage the data deletion**
    - **all datasets in ATLAS have a lifetime**
  - **Rucio client tools**
  - Command Line Interface and web interface
  - **that allow end-users or other service to interact with Rucio**

Overview of Rucio Architecture



Rucio account concept

- **Rucio account :**
  - **It can represent users, groups (higgs), activities (production).**
    - **Rucio account is the same as the CERN login name**
    - **Quota is associated to one account.**
  - **One can connect to a Rucio account using x509 certificate/proxy**
  - One credential can be used to map to different accounts.



Rucio namespace concept

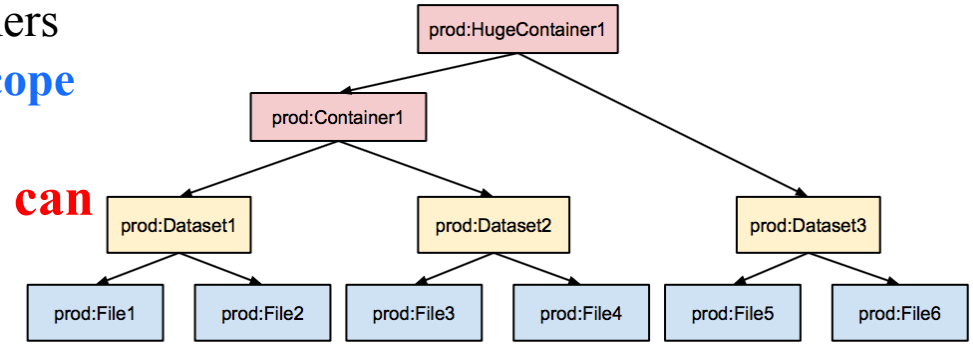
- **In general:**
  - **Each job produces one file.**
  - **Each JEDI task produces one dataset.**
  - **Each task request produces one container**.

➢ **3 types of Data Identifiers (DIDs):**

➢ **Files, Datasets, Containers.**

➢ Allows multiple hierarchy level for containers
All Data Identifiers are identified by a scope and a name.

➢ **A name is unique within a scope but can be used in other scopes**



For convenience containers of containers are produced manually.

Rucio Storage Elements

- **Rucio Storage Elements (RSE):**
  - Abstraction for storage end-point.
  - **Define different type of places to store data.**
    - **Usually look like SITE_TOKEN, e.g. CERN PROD_SCRATCHDISK**
  - **Centrally produced data is hosted on DATADISK RSE.**
  - **When you run jobs on the Grid output goes to local SCRATCHDISK RSEs (temporary space).**
  - **LOCALGROUPDISK RSEs for long-term storage.**

Permission/quotas

- **2 different types of RSEs open to end-users to store output datasets:**
  - **SCRATCHDISK area : Default area for task outputs.**
    - **Double copy: one on RSE where job ran, the other on SCRATCHDISK with free space**
    - **Default lifetime : 2 weeks.**
  - **LOCALGROUPDISK area : For long term storage of user data.**
    - **No default lifetime.**

- **Every user has a quota on all SCRATCHDISK RSEs.**
- **Everyone gets a quota by default on LOCALGROUPDISK RSEs in their country. We do not have this RSE in Morocco**

- **LOCALGROUPDISK RSEs:** Admins of the country can modify the quota, approve replication requests, delete old data

Rules

- **Replication rules:**

- Rules describe how a Data Identifier must be replicated on a list of Rucio Storage Elements.
  - **It is a request to transfer data.**
  - **It Can have a lifetime after which data can be cleaned.**

- **Rules mechanisms allows Rucio:**
- **to create the minimum number of replicas**
- **to optimize storage space, and**
- **To minimize the number of transfers and automate data distribution.**

Summary of concepts in Rucio

- **In Rucio, the files, datasets and containers are defined as:**
- **Data IDentifiers (or DIDs).**

- All DIDs are created within a scope and identified by the concatenation of the scope and the DID name separated by a colon **(e.g. user.jdoe:test.root where user.jdoe is the scope and test.root is the DID name).**

- **Each data identifier name is unique within a scope, but 2 data identifiers can have the same name in different scopes**
  - **(e.g. user.jdoe:myfile.root and user.janedoe:myfile.root)**

- When you query DIDs matching a certain pattern or metadata you must always specify a scope.

Tools for users

- **CLIs - “rucio ” command for end users**
  - Find datasets.
  - Find where they are.
  - Download them.

- **R2D2 – web interface**
  - https://rucio-ui.cern.ch/r2d2/request
  - Request transfer of datasets by making rules.
  - Normally to your LOCALGROUPDISK RSEs.

If you are new to this inte to take the tour. x

If you find any errors or have suggestions for improvements for this interface please report it to Jira. x

Your input will be saved until you submit it. If you want to clear the form please click here.

1. Select Data Identifiers (DIDs)

DID Pattern Search | List of DIDs

Please start by entering a DID or DID wildcard and search for either containers or datasets. Then select the requested DIDs. Please do not use a trailing '/' for containers.

Data pattern | scope:name... | Search | Container | Dataset

2. Select Rucio Storage Elements (RSEs)

3. Options

4. Summary

Data Identifiers and Scope

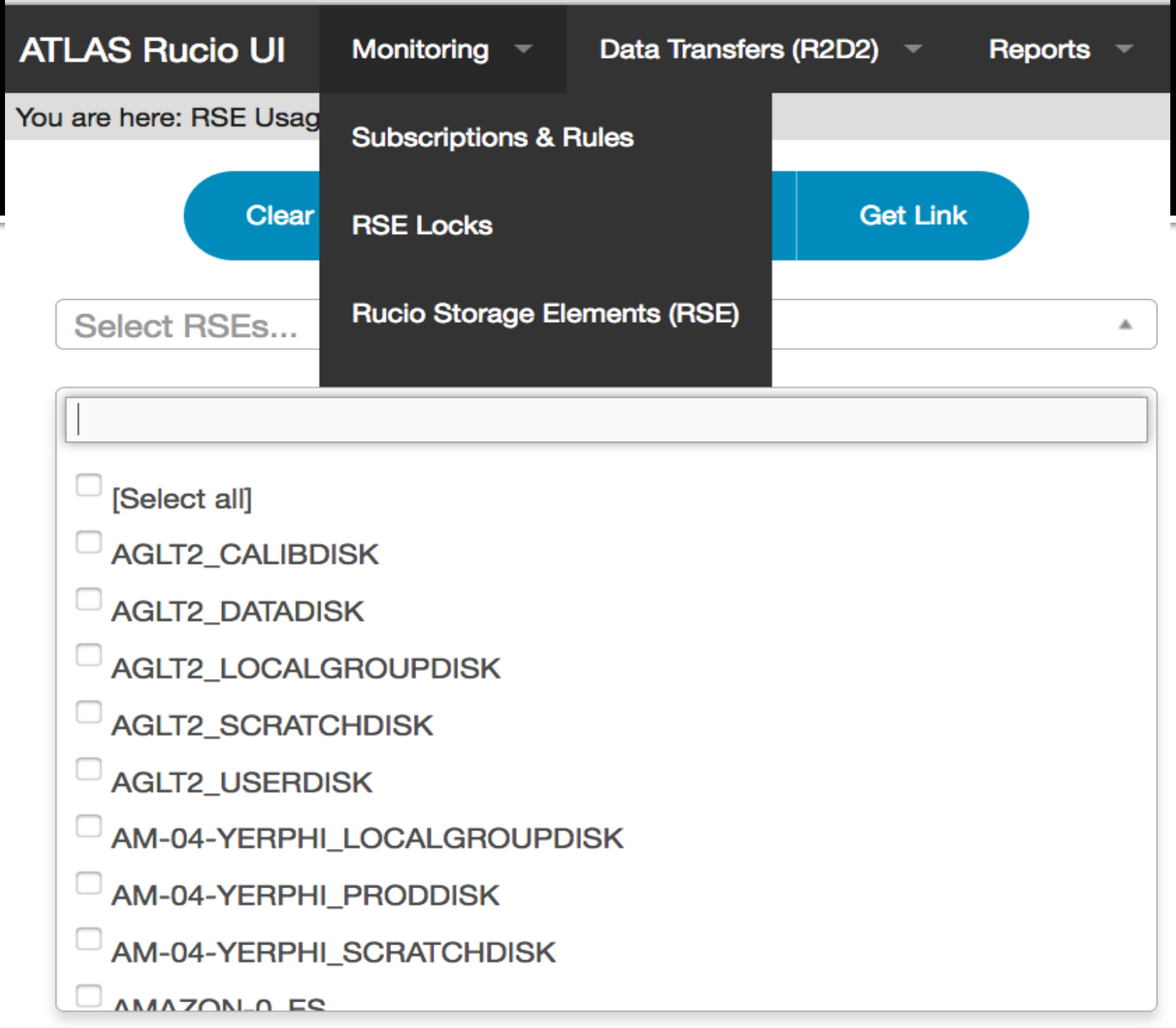
Files, datasets and containers share the same naming convention, which is composed of two strings: the scope and the name, separated by a colon. The combination of scope and name is called a data identifier (DID).

The scope is used to divide the name space into several, separate sub spaces for production and individual users. User scope always start with 'user.' followed by the account name.

By default users can read from all scopes but only write into their own one. Only privileged accounts have the right to write into multiple scopes including production scopes like `mc15_13TeV.`

Examples:

```
Official dataset:
data15_13TeV.00266904.physics_Main.
merge.DAOD_SUSY1.f594_m1435_p2361_tid05608871_00
```


ATLAS Rucio UI
Monitoring
Data Transfers (R2D2)
Reports
You are here: RSE Usag
Subscriptions & Rules
RSE Locks
Rucio Storage Elements (RSE)
Clear
Get Link
Select RSEs...
[Select all]
AGLT2_CALIBDISK
AGLT2_DATADISK
AGLT2_LOCALGROUPDISK
AGLT2_SCRATCHDISK
AGLT2_USERDISK
AM-04-YERPHI_LOCALGROUPDISK
AM-04-YERPHI_PRODDISK
AM-04-YERPHI_SCRATCHDISK

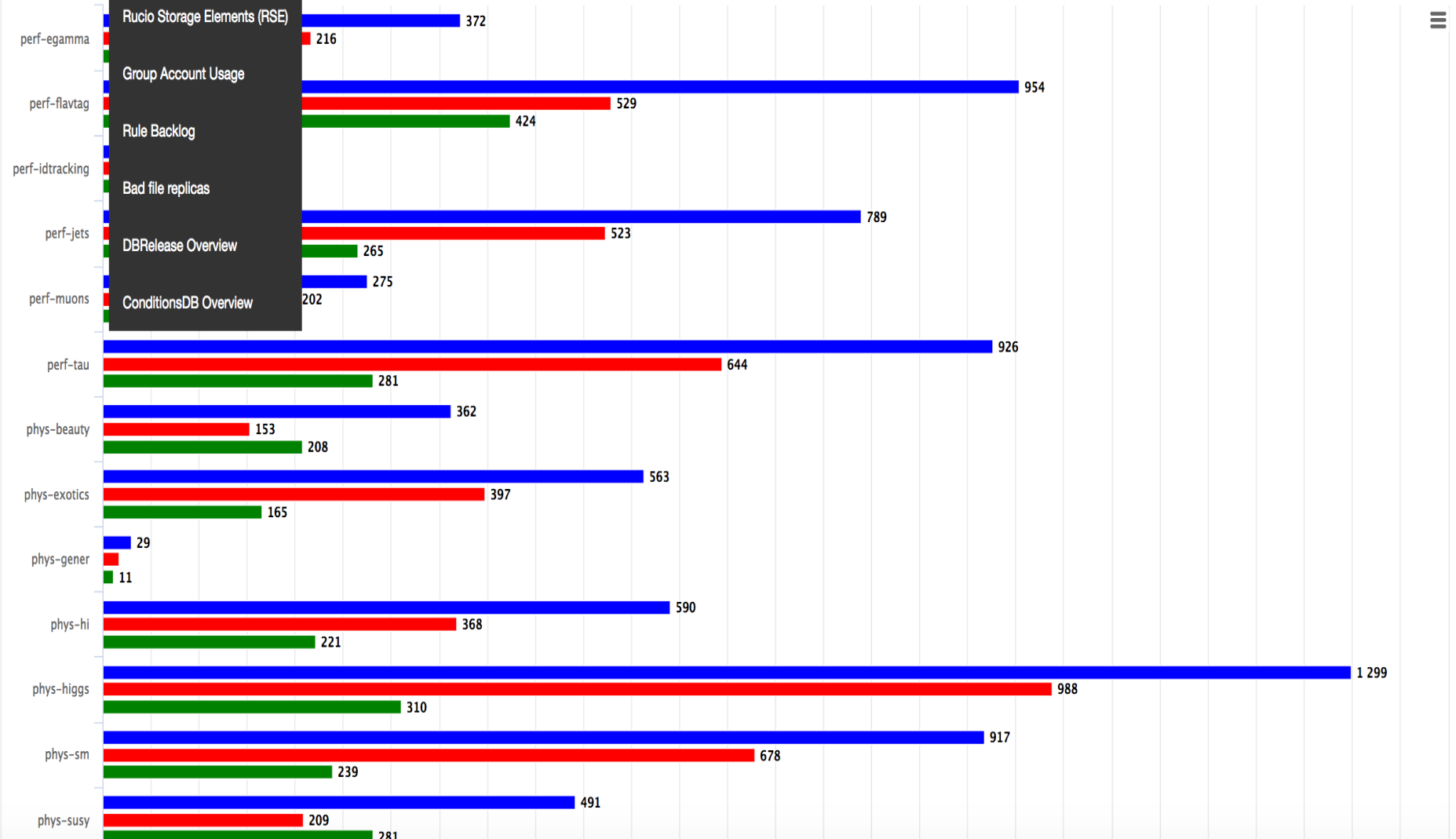
ATLAS Rucio UI | Monitoring | Data Transfers (R2D2) | Reports | pattern OR name OR rule id | Search | Using account: farida | Other Monitoring | Help

You are here: Group Acc

Rucio Version (WebUI / Server): 1.11.4 / 1.11.3

- Subscriptions & Rules
- RSE Locks
- Rucio Storage Elements (RSE)
- Group Account Usage
- Rule Backlog
- Bad file replicas
- DBRelease Overview
- ConditionsDB Overview

Group Acc w (in TB / no TAPE)



Rucio CLI Examples

- `rucio --version`
- `rucio ping`
- `rucio add-replicas`
- `rucio add`
- `rucio del`
- `rucio list`
- `rucio list-files`
- `rucio list-replicas`
- `rucio upload`
- `rucio download`
- `rucio search`
- `rucio get-metadata`
- `rucio set-metadata`
- `rucio del-metadata`
- `rucio list-rse-usage`
- `rucio list-account-usage`

To get the list of all the available RSEs, you can try:

```
$> rucio list-rses
AGLT2_CALIBDISK
AGLT2_DATADISK
AGLT2_LOCALGROUPDISK
AGLT2_PERF-MUONS
AGLT2_PHYS-HIGGS
... output omited ...
ZA-UJ_PRODDISK
ZA-UJ_SCRATCHDISK
ZA-WITS-CORE_LOCALGROUPDISK
ZA-WITS-CORE_PRODDISK
ZA-WITS-CORE_SCRATCHDISK
```

You can filter the type of the endpoint

```
$> rucio list-rses | grep SCRATCHDISK
AGLT2_SCRATCHDISK
AM-04-YERPHI_SCRATCHDISK
ANLASC_SCRATCHDISK
AUSTRALIA-ATLAS_SCRATCHDISK
... output omited ...
WEIZMANN-LCG2_SCRATCHDISK
WUPPERTALPROD_SCRATCHDISK
ZA-UJ_SCRATCHDISK
ZA-WITS-CORE_SCRATCHDISK
```

RucioClientsHowTo

https://twiki.cern.ch/twiki/bin/viewauth/AtlasComputing/RucioClientsHowTo

Thanks!!
Any Questions?