Multi-VO Dirac/Rucio Integration Planning

Janusz Martyniak

GridPP Technical Meeting
11/10/2019

Aim of the project

- Use Dirac for workload management
  - Use existing user/group/VO registry
  - Use existing resources as defined in Dirac configuration (CEs, SEs)
  - Drop Dirac File Catalog for VO willing to use Rucio for data management
- Use Rucio for data management
  - More flexible than Dirac, i.e. allows rule-based automatic replica management
- Work to to start at month 12 (Jan 2020)

Implementation

- Dirac catalog replacement
  - Dirac allows using multiple catalogs (plugins), event multiple catalogs per VO.
  - Dirac own catalog (DFC) is part of the server and it is contacted by clients using RPC
  - Another possibility is to contact a catalog directly, this method is used for the LFC, still used by some Dirac instances (not the GridPP Dirac)
  - The direct route seems to be most suitable for integrating Dirac with Rucio

Rucio File Catalog Plugin for Dirac

- Configure a VO to use a catalog plug-in
- User Dirac data management client API or scripts (e.g. *dirac-dms-add-file*) to contact Dirac
- Dirac selects a correct plugin based on VO information stored in a proxy
- Plugin performs Rucio data handling operations using Rucio APIs or REST iface.
- User identity has to be passed by Dirac to Rucio API (VO in particular)

First Tests

Some testing has been done using a very simple *RucioFileCatalogClient* plugin

- Use Dirac UI + pip installed Rucio client hybrid
- Test Rucio server at Imperial (single VO)
- *dirac-dms-add file* operation successfully uploaded a file to a SE/RSE
- A Rucio plugin code was executed in an attempt to register a replica with Rucio (*ReplicaClient.add_replica(….)*)

Dirac-Rucio mapping

- Dirac *lfn*: */voname/scope/some/other/file.ext*
- second part of the *lfn* is used as Rucio scope
- Remaining part of the *lfn* forms the DID
- Dirac configuration is planned to be a master, so we need to:
  - Map users to Rucio accounts
  - Map SEs to RSEs

Services Required for Development

- Dirac test instance at Imperial
- Rucio server at RAL (even single VO for a start)
  - With a few RSEs defined (names could match DiracSE names for simplicity)
  - With client level access from Imperial development host (account name(s) match Dirac user names)
  - A defined VO – we use gridpp VO for non-Rucio work and MICE for Rucio work on our test Dirac instance