Brookhaven National Laboratory

U.S. DEPARTMENT OF ENERGY

Current status of DIRAC and Rucio cooperation

Cédric Serfon (BNL)

October 17, 2023

Introduction

- DIRAC and Rucio have been very popular into the HEP community and beyond for many years now
- Even though DIRAC has its data management functionalities, some communities have started considering using DIRAC for Workload/Workflow Management and Rucio as Distributed Data Management
  - AFAIK, it was first mentioned in 2nd Rucio workshop and 9th Dirac User Workshop
  - Belle II was the first community to have a first production instance of DIRAC + Rucio
- This talk tries to summarize the current status of DIRAC and Rucio cooperation

First Rucio/DIRAC integration

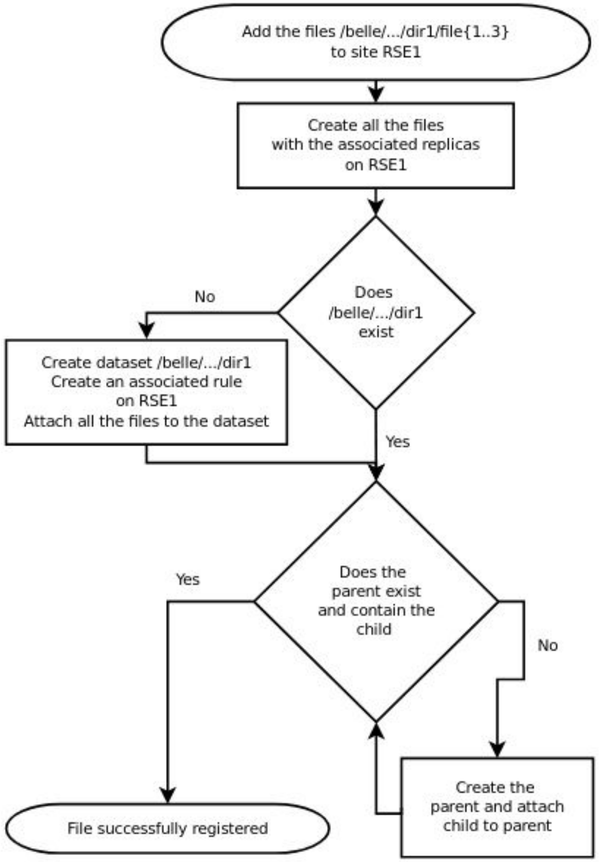
- Belle II uses its own DIRAC extension called BelleDIRAC
- In order to integrate Rucio and DIRAC in Belle II, the initial developments were done as part of BelleDIRAC
- The idea was to create a RucioFileCatalogue (RFC) similar to LFC or DFC :
    - No change for the download/upload still done via the DataManager
    - All the replication policies, 3rd party copy are handled by Rucio subscriptions and rules
    - The synchronization between DIRAC and Rucio is done via different DIRAC agents
- More technical details in backup if you are interested

Rucio specific developments into DIRAC

- BelleDIRAC RFC was merged into Vanilla Dirac and adapted to support multiVO (Janusz)
- In addition 2 agents to synchronize Rucio and Dirac were added :
  - RucioSynchronizerAgent :
    - To synchronize accounts, groups, SEs from Dirac into Rucio
    - Also add to specific RSEs attributes used for Rucio subscriptions (can be very community specific and needs to be overridden into extension)
  - RSSAgent : To synchronize the status of SEs in DIRAC in Rucio
- Rucio now also available in DiracOS(2)

DIRAC specific development into Rucio

- In order to provide methods needed for the RFC, new dirac REST endpoint was created in Rucio (mainly to have an atomic addfile method)
- Method to extract scope from LFN is also available in Rucio or in Rucio policy package
- New metadata methods or new option added to mimic DFC behaviour (see later)


Add the files /belle/.../dir1/file{1..3} to site RSE1
Create all the files with the associated replicas on RSE1
Does /belle/.../dir1 exist
No
Create dataset /belle/.../dir1 Create an associated rule on RSE1 Attach all the files to the dataset
Yes
Does the parent exist and contain the child
Yes
No
Create the parent and attach child to parent
File successfully registered

Flow for addFile method

Other collaborations

- A few other collaborations did some tests using DIRAC + Rucio or plan to do it :
  - CTAO :
    - Some tests already performed ~1 year ago by Frederic and Luisa using a Rucio shared instance at PIC in the context of the ESCAPE project :
      - Managed to upload/download files after tweaking Rucio instance configuration
    - More tests are being done now on a dedicated instance
  - Juno (see Xiaomei report yesterday)
  - GridPP

Brookhaven National Laboratory

New Rucio-(Belle)DIRAC developments

- The developments mentioned previously are all included in Vanilla DIRAC, but new additional developments were done in BelleDIRAC :
    - Report of “traces” to Rucio to determine dataset popularity. Might also be merged into Vanilla DIRAC
    - Support of Rucio metadata :
        - DFC metadata methods were also implemented into RFC. Can be easily merged into Vanilla DIRAC (see next slides)
        - Different developments were done in the Production/Fabrication system and pilot to report metadata to Rucio. Probably too Belle II specific, no plan to merge to Vanilla Dirac
        - Use of Rucio metadata was tested and validated well beyond Belle II scale

Metadata

- A few collaborations seem to be interested in using Rucio metadata
- Current Belle II integration :
  - Uses PostGreSQL json, but easy to use other backend (e.g. NoSQL like MongoDB)
  - Has metadata inheritance like in DFC
  - Doesn’t allow (yet) query by metadata. Would need to define metadata schema + indices
- Status in Belle II :
  - DFC methods in BelleDIRAC for a long time. Can open a PR to integrate them into DIRAC in the coming weeks
  - Use of these methods in Production/Fabrication Service + pilot still not in production yet (hopefully end of the month)

Metadata in Rucio

- Extensive tests done to insert/read json metadata.
  - No bottleneck found (can be scaled horizontally)
  - Insertion rate validated order of magnitude beyond Belle II’s need
- Import of ~70M metadata in Rucio production instance (between 10 and 20 metadata/file) show no degradation of the performance

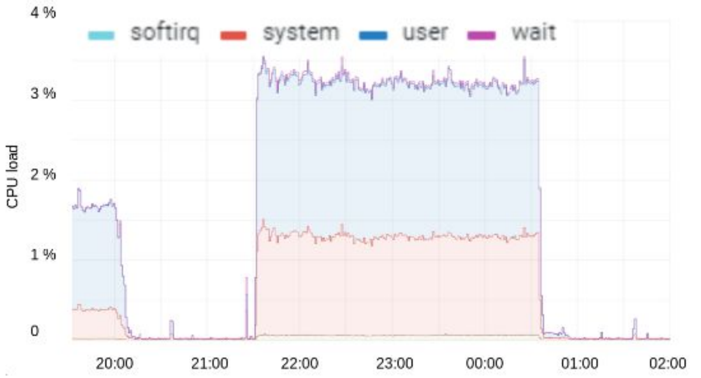
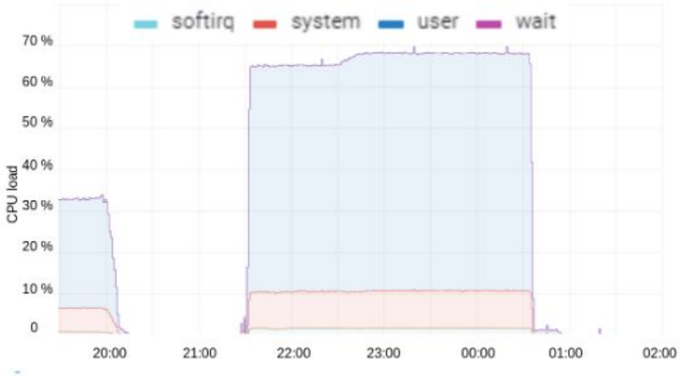
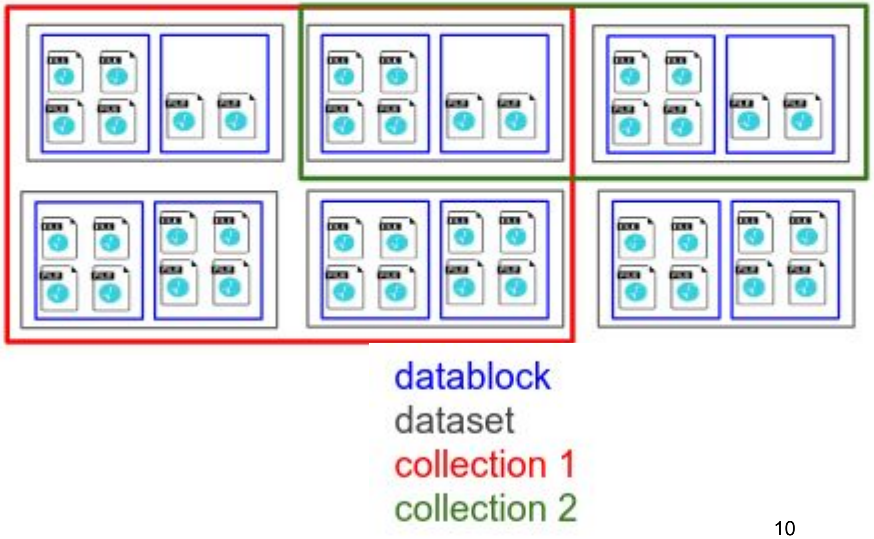


Figure 2. CPU usage of the Database node during the stress test. The highest plateau corresponds to the maximum load of the test (360 parallel jobs) with the load reaching only 3%



New Rucio-(Belle)DIRAC developments

- More BelleDIRAC developments :
  - Different developments affecting end-user tools :
    - Replacement of datamanager with rucio native download client
    - Introduction of asynchronous replication/deletion for the end-users
    - Introduction of “collections” (aggregation of Rucio datasets into containers)
- No plan (yet) to merge them into Vanilla DIRAC (basic client use cases can be fulfilled with the DIRAC DMS client APIs,without using Rucio client API)



New Rucio-(Belle)DIRAC developments

- More BelleDIRAC developments :
  - Staging agent for datasets on TAPE interfaced to the Production System
  - Automatic deletion of intermediate datasets for productions
- Here again, need to evaluate if it can be merged to Vanilla DIRAC (probably doesn’t make sense until we move to DiracX)

Rucio + DiracX

- DiracX is a good opportunity to push forward better Dirac/Rucio integration. Here is a list of topics that can be addressed :
  - Integration of Rucio subscriptions into DIRAC (e.g. possibility to define a subscription in DIRAC CS)
  - Dataset lifetime management
  - Common monitoring (based on ELK stack)
  - Tokens : How Rucio and DiracX can play together ?
  - Introduce some unit-tests
- docker-compose/helmcharts to start a DIRAC-Rucio instance

Rucio + DiracX

- From my point of view, could be useful to have common technical sessions (2-3 times per year) where we can :
  - Discuss common challenges (e.g. tokens)
  - Learn from the other community about (non-exhaustive list) :
    - New tools/technologies (e.g. for monitoring)
    - Development/deployment model
    - Best practices
  - Discuss specific issues for communities using both Dirac(X) and Rucio
- Should we formalize this need for further integration, e.g. by having DiracX SIG in Rucio ?

Conclusion

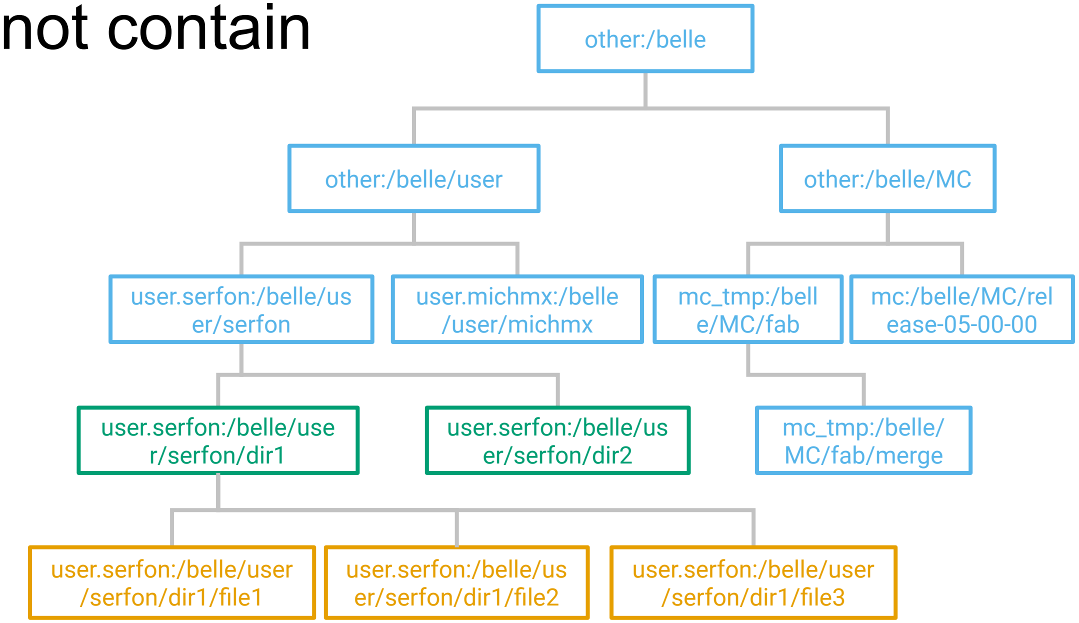
- DIRAC and Rucio can fit nicely together as demonstrated by Belle II
- Will continue working on further integration (e.g. metadata) in current DIRAC and also later on DiracX
- This combined workshop is a good opportunity to learn about each other and share experience. Hopefully, it’s just a starting point

Mapping Rucio concepts to DIRAC ones : Scope

- Scope is a central concept of Rucio but not present in DIRAC → Need to be completely hidden
- Scope can be extracted from LFN using an extract-scope function defined in Rucio :
    - Needed both on the client and server side
    - Can be VO-specific. The Belle II one is rather complicated
    - Expected to be defined in the policy package

Mapping Rucio concepts to DIRAC ones : Hierarchical namespace

- Hierarchical namespace
- Constraint : Containers cannot contain files → directories cannot contain directories AND files



other:/belle

other:/belle/user

other:/belle/MC

user.serfon:/belle/user/serfon

user.michmx:/belle/user/michmx

mc_tmp:/belle/MC/fab

mc:/belle/MC/release-05-00-00

user.serfon:/belle/user/serfon/dir1

user.serfon:/belle/user/serfon/dir2

mc_tmp:/belle/MC/fab/merge

user.serfon:/belle/user/serfon/dir1/file1

user.serfon:/belle/user/serfon/dir1/file2

user.serfon:/belle/user/serfon/dir1/file3

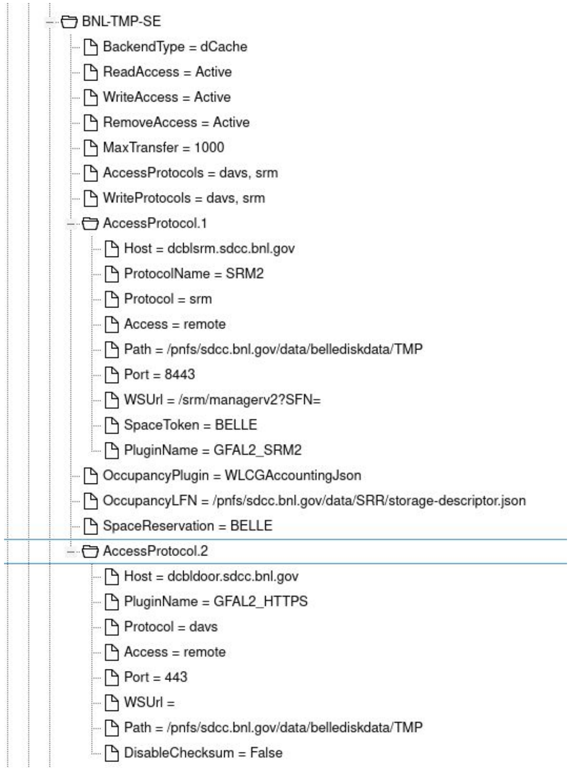
File → File

Directory containing files → Dataset

Directory containing directories → Container

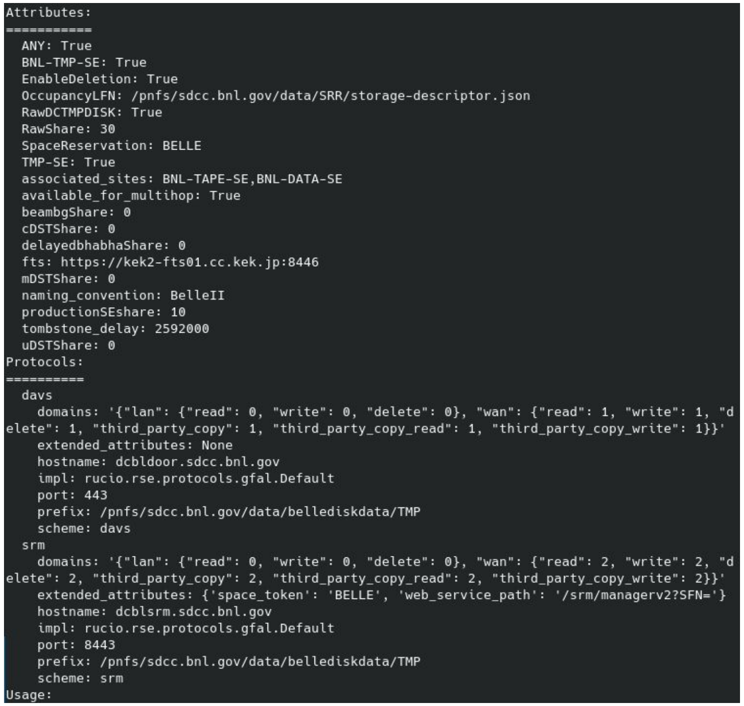
Mapping Rucio concepts to DIRAC ones : SE

- DIRAC SEs and Rucio RSEs map nicely and are automatically synchronized from DIRAC CS



```
BNL-TMP-SE
  BackendType = dCache
  ReadAccess = Active
  WriteAccess = Active
  RemoveAccess = Active
  MaxTransfer = 1000
  AccessProtocols = davs, srm
  WriteProtocols = davs, srm
  AccessProtocol.1
    Host = dcblsrm.sdcc.bnl.gov
    ProtocolName = SRM2
    Protocol = srm
    Access = remote
    Path = /pnfs/sdcc.bnl.gov/data/bellediskdata/TMP
    Port = 8443
    WSUrl = /srm/managerv2?SFN=
    SpaceToken = BELLE
    PluginName = GFAL2_SRM2
  OccupancyPlugin = WLCGAccountingJson
  OccupancyLFN = /pnfs/sdcc.bnl.gov/data/SRR/storage-descriptor.json
  SpaceReservation = BELLE
  AccessProtocol.2
    Host = dcbldoor.sdcc.bnl.gov
    PluginName = GFAL2_HTTPS
    Protocol = davs
    Access = remote
    Port = 443
    WSUrl =
    Path = /pnfs/sdcc.bnl.gov/data/bellediskdata/TMP
    DisableChecksum = False
```

DIRAC CS



```
Attributes:
===========
  ANY: True
  BNL-TMP-SE: True
  EnableDeletion: True
  OccupancyLFN: /pnfs/sdcc.bnl.gov/data/SRR/storage-descriptor.json
  RawDCTMPDISK: True
  RawShare: 30
  SpaceReservation: BELLE
  TMP-SE: True
  associated_sites: BNL-TAPE-SE,BNL-DATA-SE
  available_for_multihop: True
  beambgShare: 0
  cDSTShare: 0
  delayedbhabhaShare: 0
  fts: https://kek2-fts01.cc.kek.jp:8446
  mDSTShare: 0
  naming_convention: BelleII
  productionSEshare: 10
  tombstone_delay: 2592000
  uDSTShare: 0
Protocols:
==========
  davs
    domains: '{"lan": {"read": 0, "write": 0, "delete": 0}, "wan": {"read": 1, "write": 1, "d
elete": 1, "third_party_copy": 1, "third_party_copy_read": 1, "third_party_copy_write": 1}}'
    extended_attributes: None
    hostname: dcbldoor.sdcc.bnl.gov
    impl: rucio.rse.protocols.gfal.Default
    port: 443
    prefix: /pnfs/sdcc.bnl.gov/data/bellediskdata/TMP
    scheme: davs
  srm
    domains: '{"lan": {"read": 0, "write": 0, "delete": 0}, "wan": {"read": 2, "write": 2, "d
elete": 2, "third_party_copy": 2, "third_party_copy_read": 2, "third_party_copy_write": 2}}'
    extended_attributes: {'space_token': 'BELLE', 'web_service_path': '/srm/managerv2?SFN='}
    hostname: dcblsrm.sdcc.bnl.gov
    impl: rucio.rse.protocols.gfal.Default
    port: 8443
    prefix: /pnfs/sdcc.bnl.gov/data/bellediskdata/TMP
    scheme: srm
Usage:
```

Rucio