Introduction talk

Ruslan Mashinistov

Ruslan Mashinistov

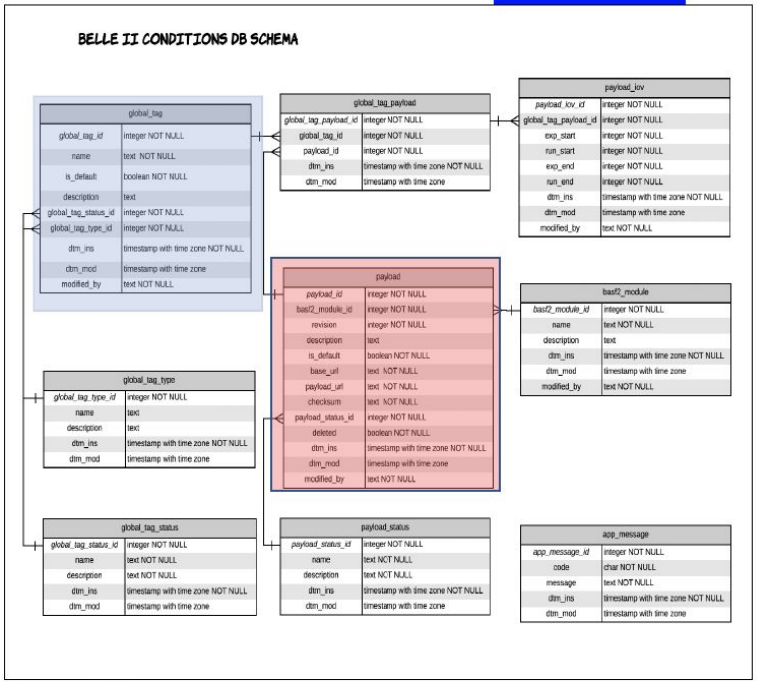
- **I’m the member of the BelleII group.** The Belle II experiment is a leading world class B-physics experiment. BNL hosts a Belle II primary computing center (the largest outside of Japan) and also BNL responsible for the Conditions Database (CDB) and Distributed Data Management system (DDM).
- I’m responsible for support and development of the CDB application code.
  - New functionalities and features
  - Functional tests
- Also I’m sharing responsibilities of support and development of the DDM
  - Currently BelleII uses the DIRAC’s extension
  - Ongoing migration to Rucio
- Data Production expert shifts

Belle II CDB data model relational database

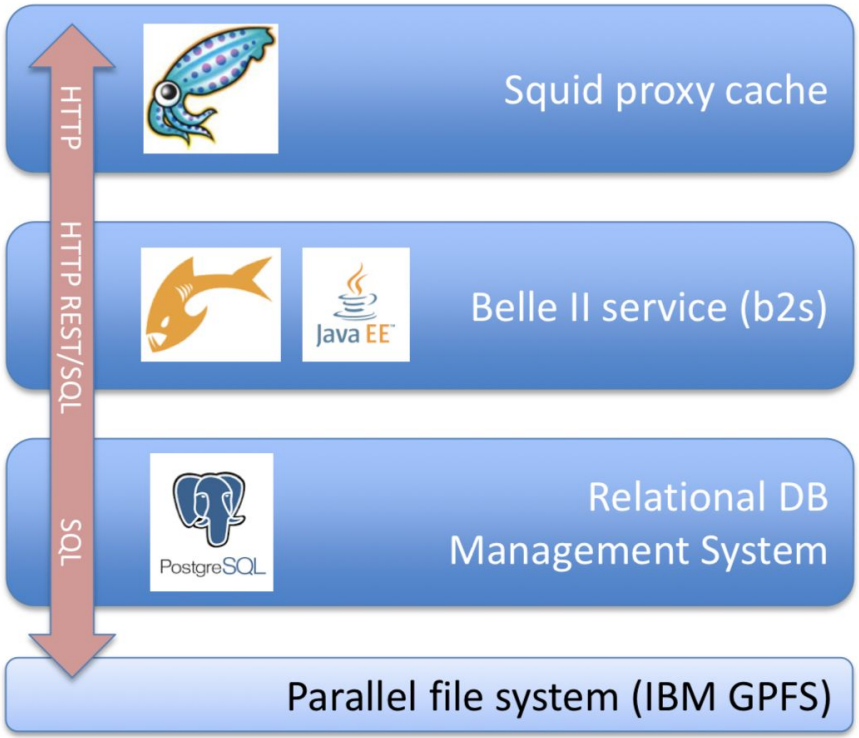
Intervals of Validity (IOV) specify starting and ending experiments and runs for a given payload for that global tag. Can be a fixed run range (closed) or starting at a given run (open)

Global tag (GT) contain list of IOV-payload relationships and are used to select a complete set of conditions for a given reprocessing effort.

Payloads A sample of conditions data (e.g. BeamParameters) stored in a file
File type is agnostic for the CDB server
ROOT file format when restricted at client side
CDB server metadata keeps track of the checksum of the file



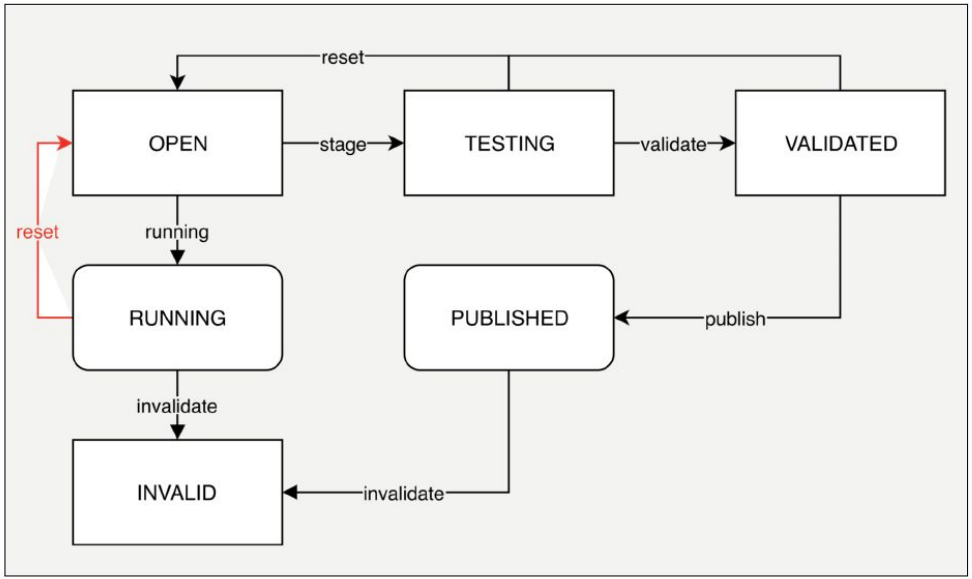
Belle II Conditions Database (CDB)



- Squid HTTP cache
  - Configured as reverse proxy – many clients, few servers
  - Supports multiple requests for the same query
  - Caches the most common global tags in Belle II
- “b2s” Belle II service layer
  - Payara-based (JavaEE) server to translate REST requests into SQL queries
  - REST API built using Swagger tools
- PostgreSQL database

New CDB developments

Global tag state machine implementation into the CDB server



Deployed in July 1st 2019

Previously Global Tag states only supported three states, NEW, PUBLISHED, INVALID

Different “kinds” of Global Tags

Published Global Tag

- new global tags can be modified freely
- completely immutable after publishing

For data reprocessing

- prepare in advance
- stays stable

➪ New GT for each reprocessing

- either start from empty one
- or clone old one into new name and modify

Running Global tag

- add payloads for new data
- only allow changes for new runs

For HLT and prompt reco:

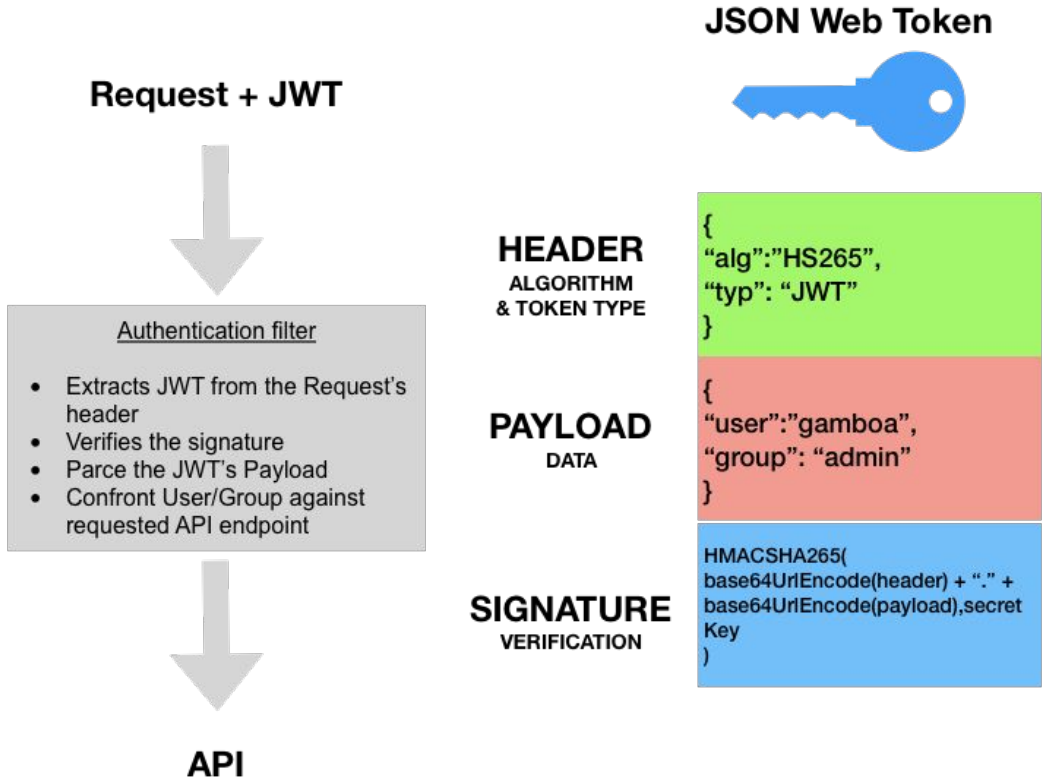
- one GT that grows with data
- keep track of what was used at the time.

➪ Updates for new runs in “staging tag”

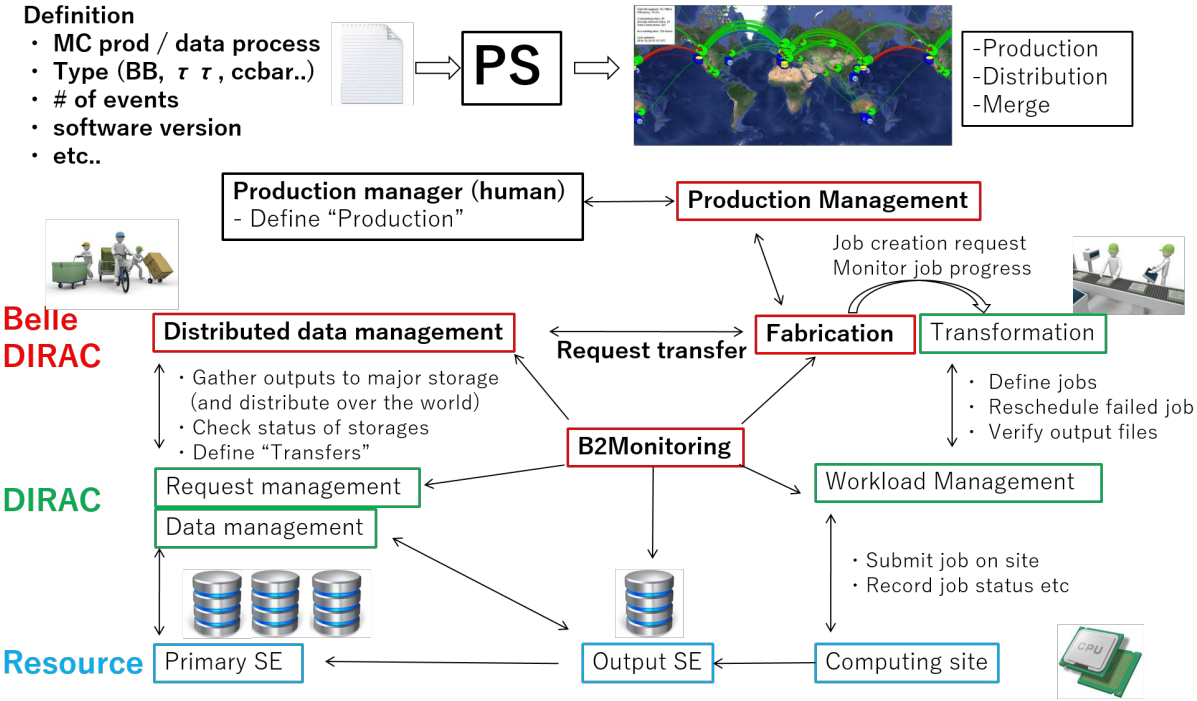
- use separate tag to prepare updates
- test and validate there and then move to running tag

Jason Web Token

- Proof of concept JWT auth implemented
- Basic functionality was tested
  - JWT signature based on shared secret
- Future plans: Users/Groups managing



Belle II Distributed Computing



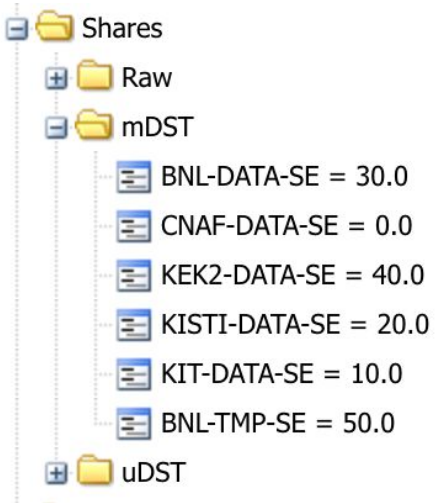
Development & ops:

- Belle DDM is a part of the BelleDIRAC (an extension of the basic DIRAC)
- Development new functionality and features
  - I'm most contributed to the ReplicaPolicy component

DDM DIRAC servers @ BNL

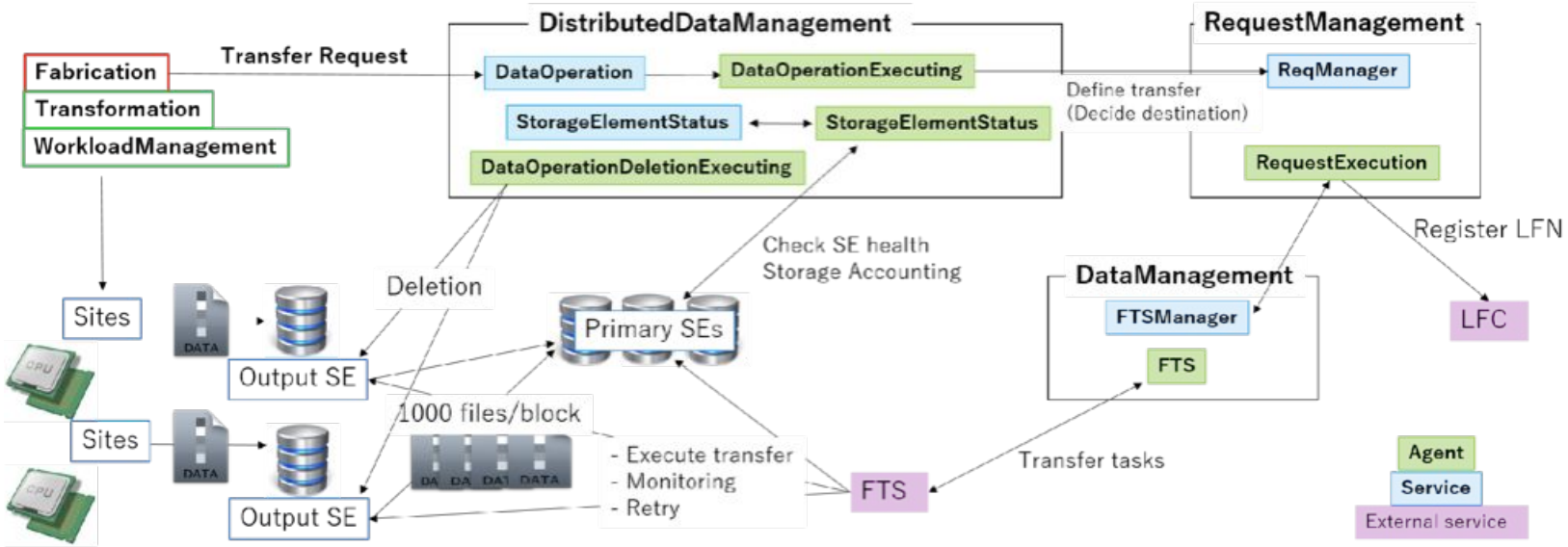
- “Production” — to run Belle II DDM components
- “Certification” — to test new BelleDIRAC codes
- “Migration” — to test upgrade of base DIRAC, upgrade of BelleDIRAC components with big jumps, …
- Development servers

ReplicaPolicy



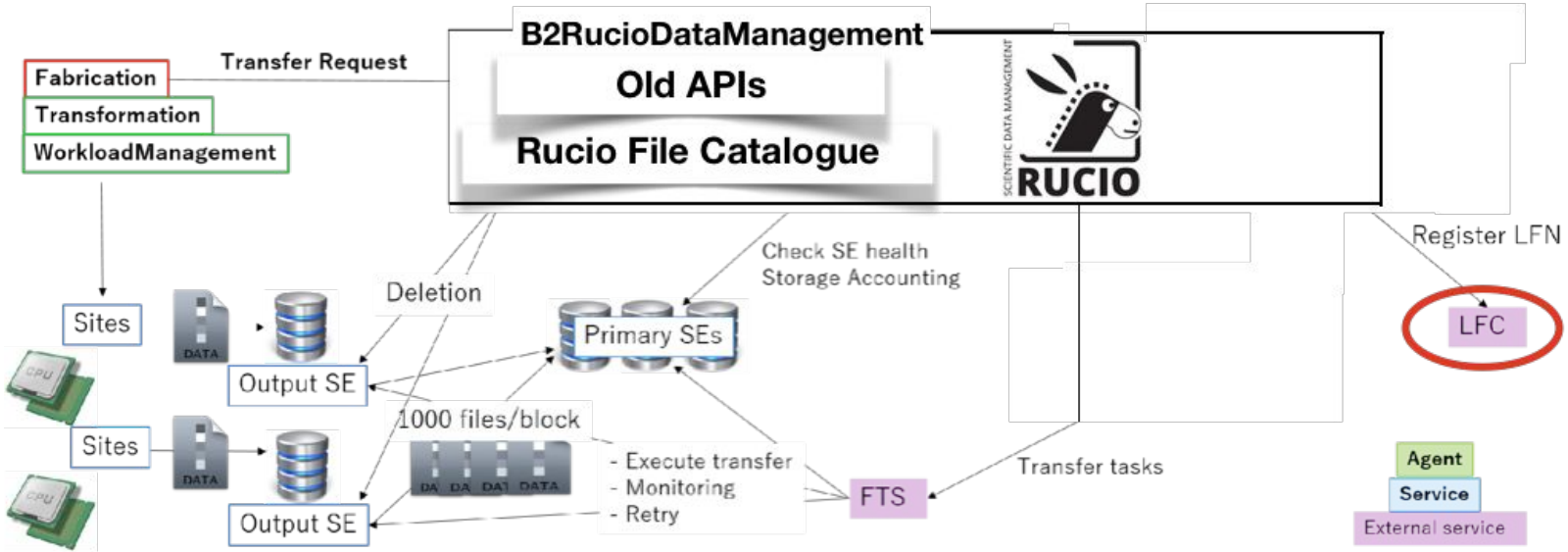
- 2 Data Distribution tools
  - Distribution scripts
    - Bash, Python, gb2-tools, DIRAC-client
  - ReplicaPolicy
    - Created by PNNL and refactored by me
    - DIRAC component, part of DDM
- RP defined with BaseLPN and Number of desired replicas
- Shares
  - List of SEs + Weights
  - Grouped by DataLevel
- Destination SEs are choosing basing on
  - Shares
  - SE health check
  - SE free space
- RP agent creates MigrateAndRegister DataOperation requests

Belle II DDM. Migration to Rucio



- ***Custom* DDM** solution inherited from **PNNL** (not based on Dirac data management), only basic functionality, lots of effort needed to fix implementation, key features either untested or missing, LFC file catalogue may soon be extinct
- ***Data-taking started in March 2019 - could not break anything!***
- ***Strong steer from DOE and Belle II reviewers to use Rucio***

Belle II Rucio Data Management - Stage 1

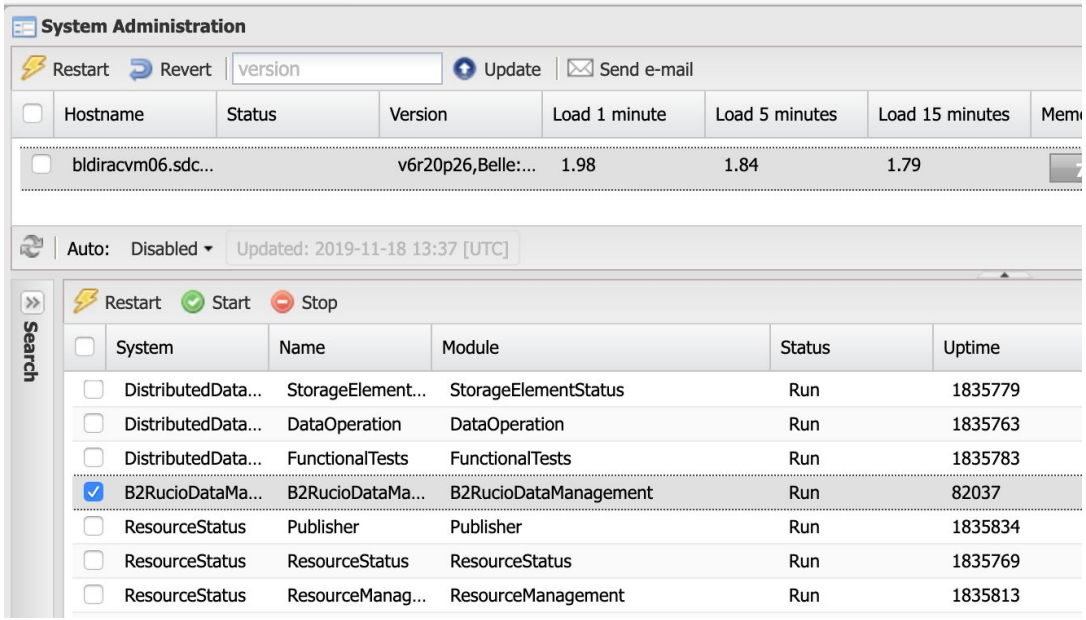


- In first stage migration, ***maintain current API*** to minimize impact
- **DDM** uses **Rucio** behind-the-scenes (only the **DDM** node has *rucio-clients*)
- **Rucio** *file catalogue* is not exposed, **LFC** is still *master file catalogue*
- Fabrication system is ***DDM-type*** aware to allow ***bi-directional migration***

Rucio DDM

- While DIRAC DDM operates the DataOperation Requests and Tasks - Rucio is using the concept of Rules
  - Declarative data management allows you to say what you want, and let Rucio figure out the details how to do it. Manage your data with expressive statements. Examples: Three copies of my file on different continents, and have one backup on tape
- The idea is that the B2RucioDataManagement will translate the external calls from DIRAC (Requests/Tasks) language for Rucio (Rules for Files and Datasets) hiddenly
- Naming convention
  - Rucio uses 4 nested layers of collections of data. 3 of them will be used at Belle II
    - Scope (DataLevel) + Dataset (Datablock) or File = DID(Data Identifier)
  - New Rucio method construct_surl_BelleII
    - Defines relative SURL for Belle II specific replicas

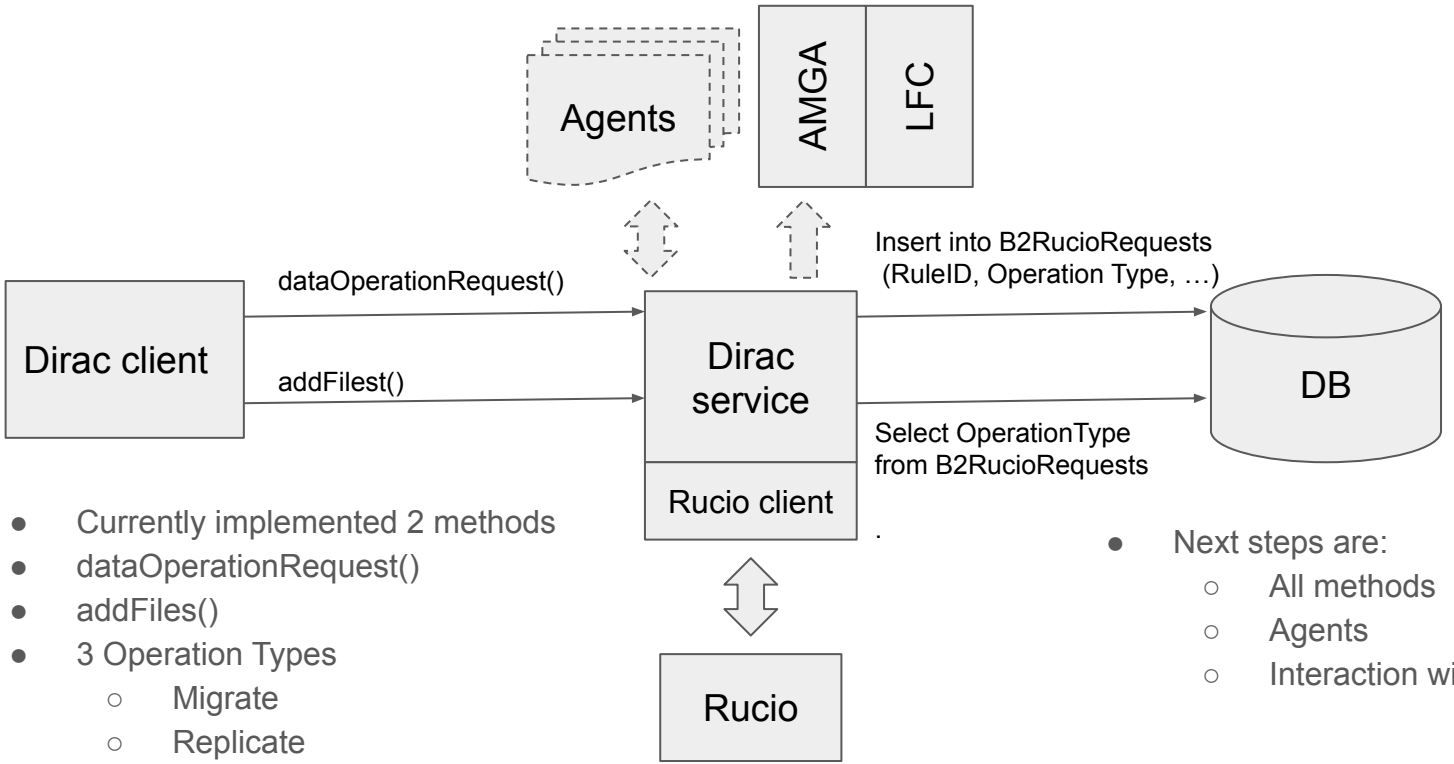
B2RucioDataManagement



- New DIRAC component is deployed at Rucio Dev
  - Service
  - DB

```
mysql> show tables;
+-----------------------+
| Tables_in_B2RucioDDMDB |
+-----------------------+
| B2RucioFiles          |
| B2RucioRequests       |
+-----------------------+
```

Service implementation

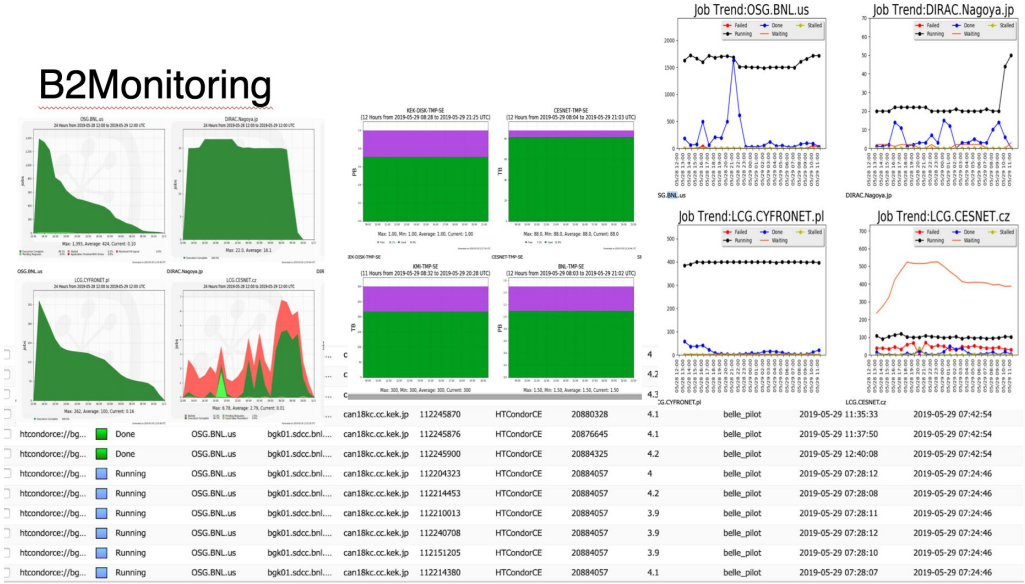


- Currently implemented 2 methods
- dataOperationRequest()
- addFiles()
- 3 Operation Types
  - Migrate
  - Replicate
  - DeepDelete

- Next steps are:
  - All methods
  - Agents
  - Interaction with LFC & AMGA

DP Expert Shifts

- DP Expert shifter is responsible for monitoring of the all components of the DC
  - 1 shift = 7 days / 24 hours
  - Investigate issues
  - DIRAC/Grid commands
  - Plots
  - Logs
  - Tickets
  - Contact with the site admins
  - User support
  - Documentation improvement



Thank you