Rucio

Commissioning, migration & operational experiences

Vincent Garonne (University of Oslo)
On behalf of the ATLAS Collaboration

Outline

- Introduction: Why Rucio ?
- Rucio concepts & Architecture
- Migration strategy
- Status
- New features
- Conclusion

DDM: Don Quijote 2 (DQ2) & Rucio

- The Distributed Data Management project is charged with managing all ATLAS data
- DQ2 was used for 8 years and evolved gradually with time to become a complete framework of Data Management tools
- But DQ2 would not have scaled for Run-2, e.g., 20 Hz of transfers, 25Hz of deletion
- Rucio is a complete redesign of the whole DDM system

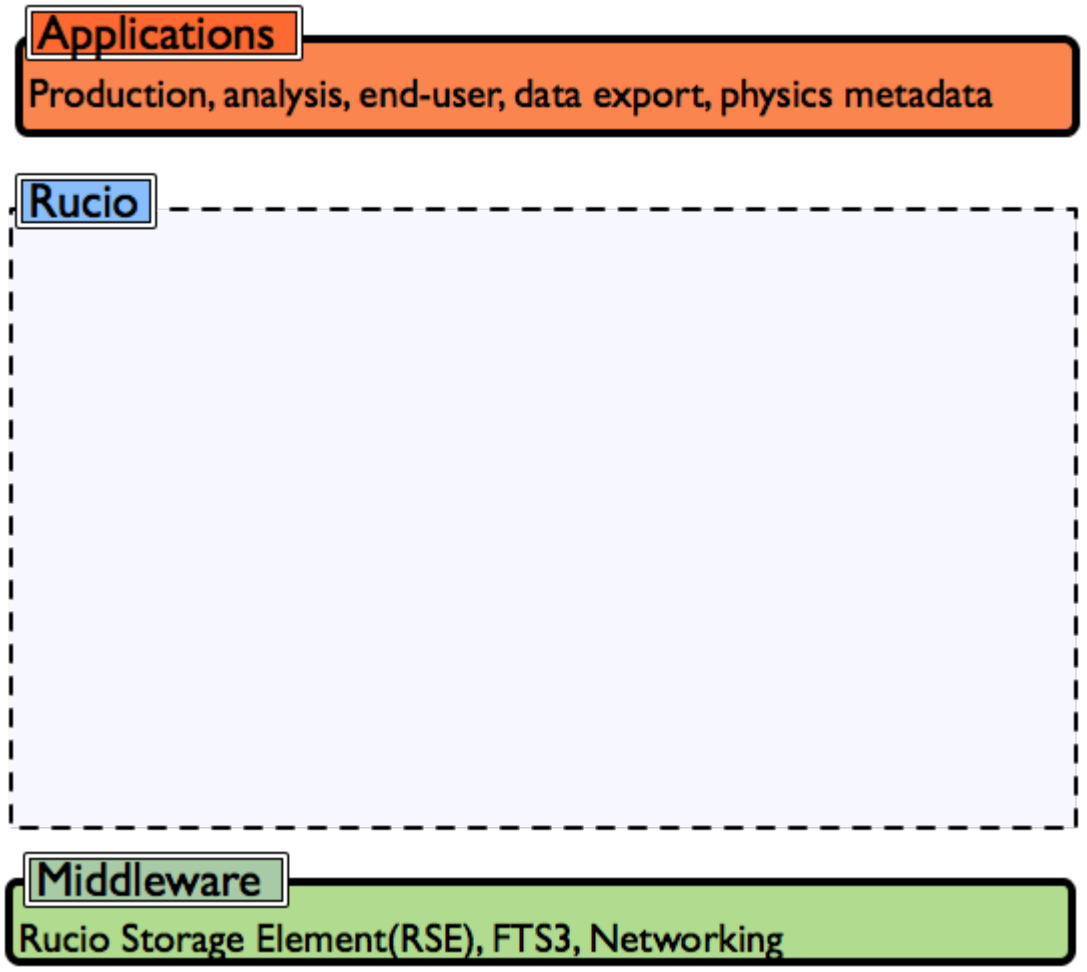
Rucio vs. DQ2 in a Nutshell

- Not only SRM supported, e.g. WebDAV, xrootd, S3, posix, gridftp, ...
- No dependencies on an external file catalog: Physical file names can be obtained from the file name via a deterministic function

  Better handling of users, groups, activities (multiple replicas ownership, quota...)
- Smarter and automated data placement tools

➥See M. Barisits[207]

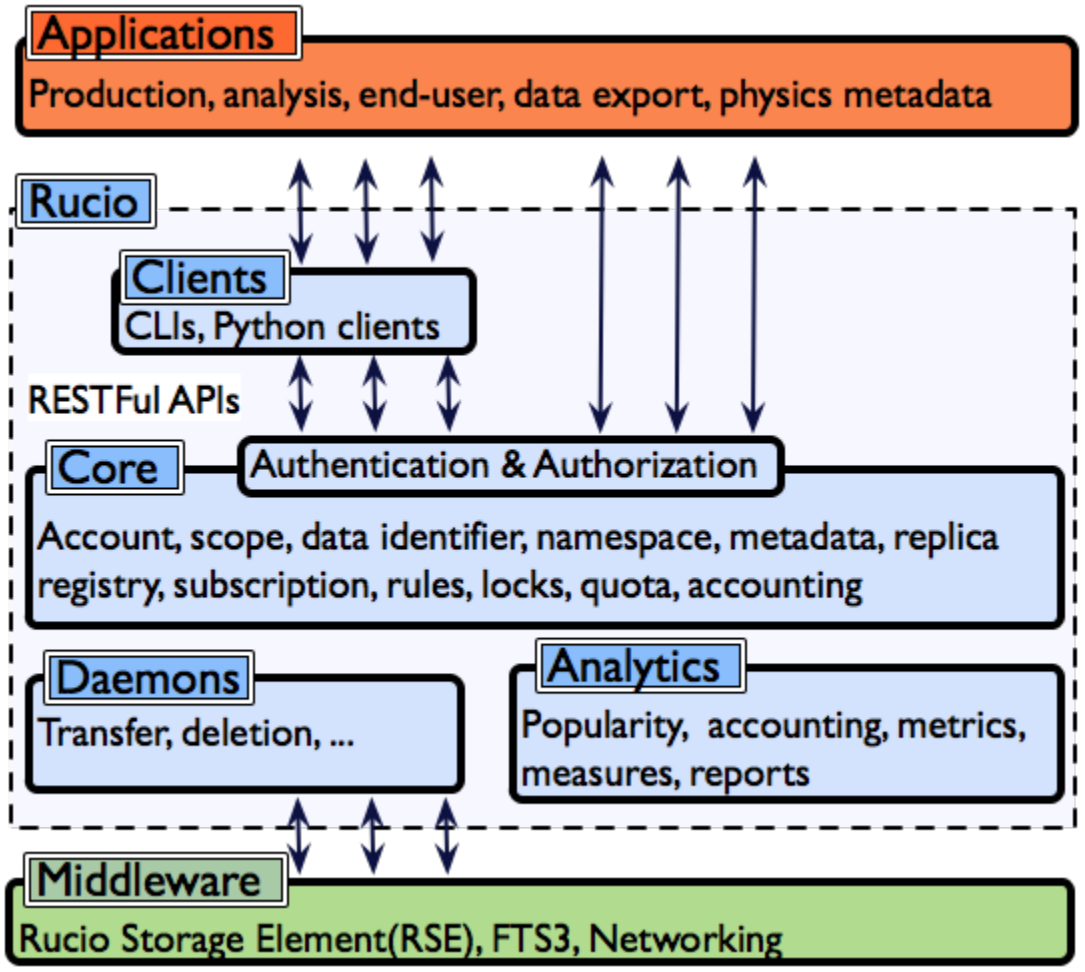
Rucio - SW Stack Overview



Open and standard technologies:

- WSGI server
- Caching
- Token-based authentication
- New middleware capabilities

Rucio - SW Stack Overview

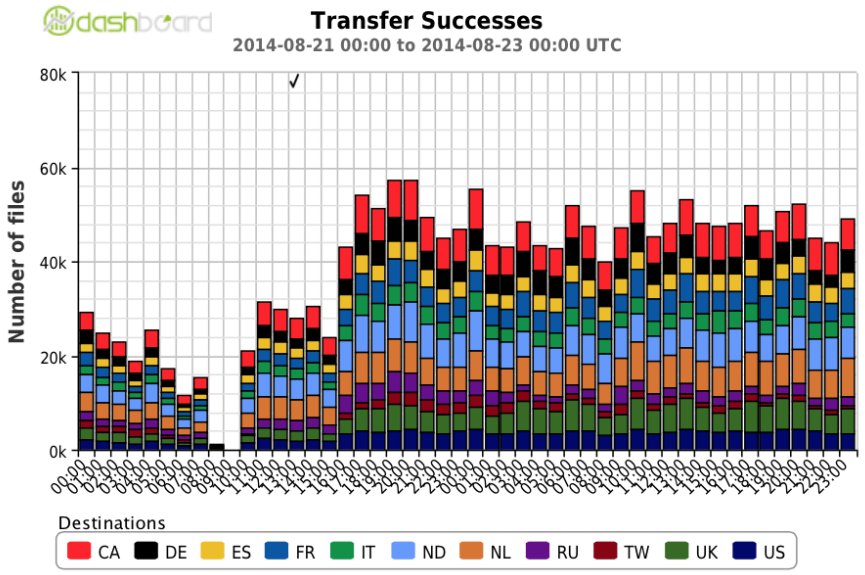


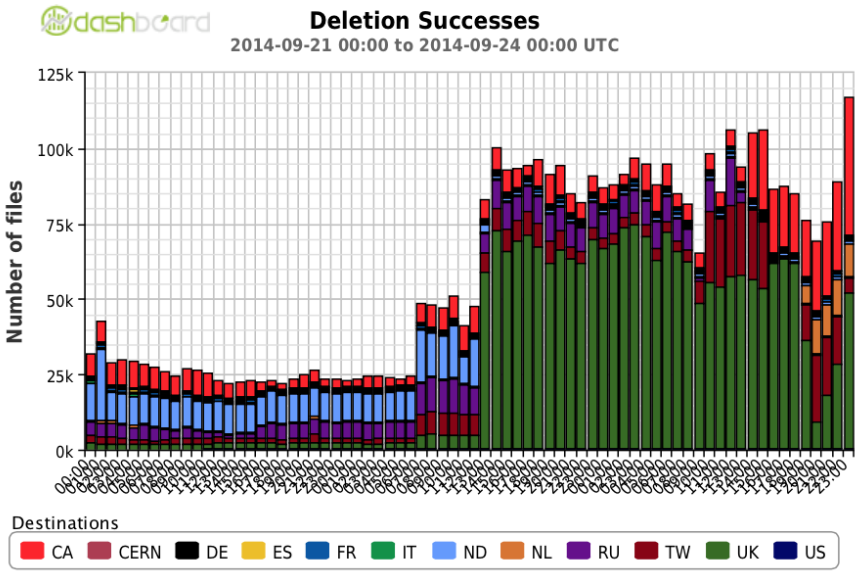
Open and standard technologies:

- WSGI server
- Caching
- Token-based authentication
- New middleware capabilities

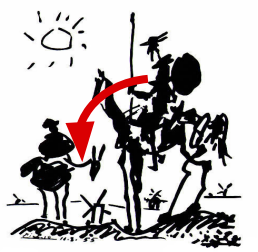
Rucio Scaling Test

- Goal was to transfer 1M/day (run-1)
  - Injecting files/datasets at Tier0
  - Replicating the datasets to Tier1 and Tier2
  - Deleting the datasets after some time
- It ran for 3-4 months with increasing load





Migration Constraints

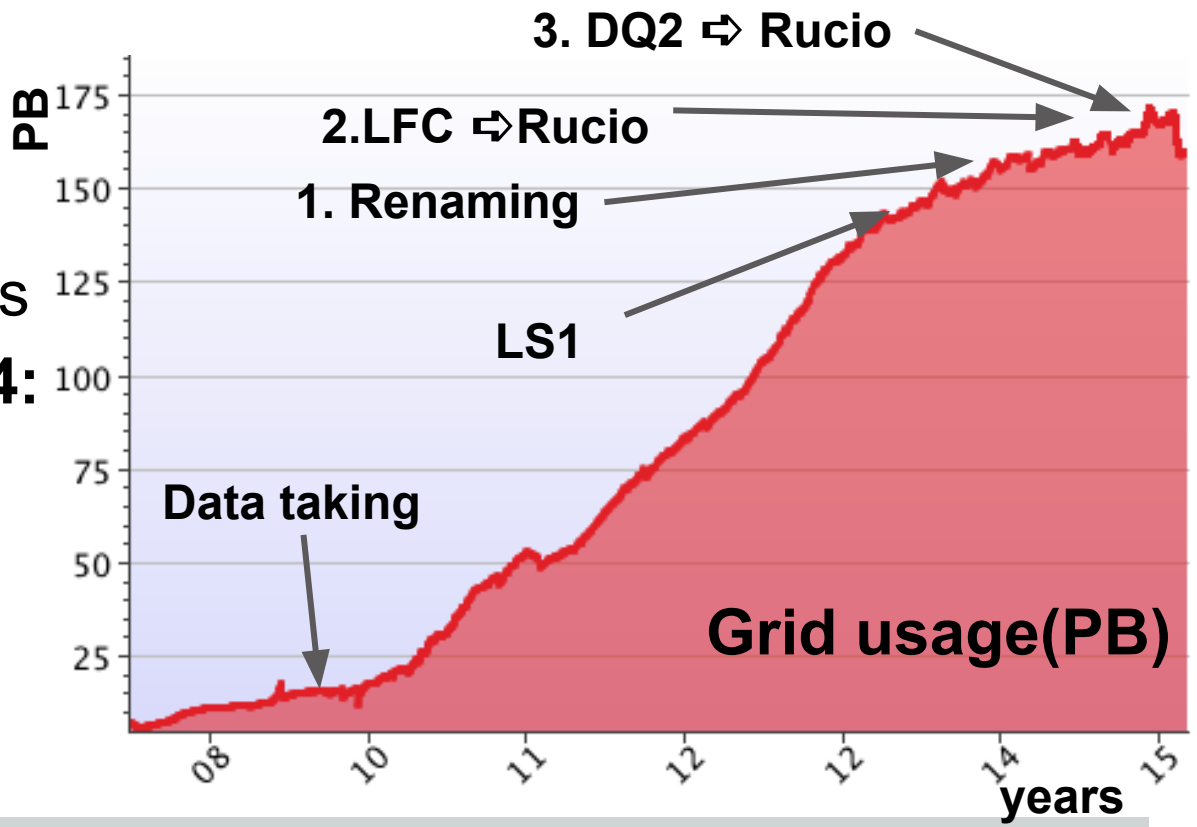


- It was not possible to stop DDM for more than a few hours, even during the shutdown
  - No way to locate file, transfer, delete, etc.
  - No distributed analysis, no production
  - Problem for the users and the sites, e.g,. idle CPUs
- The migration from DQ2 to Rucio was required to be as transparent as possible in order not to disrupt production and analysis
  - Maximum downtime allowed is a few hours !

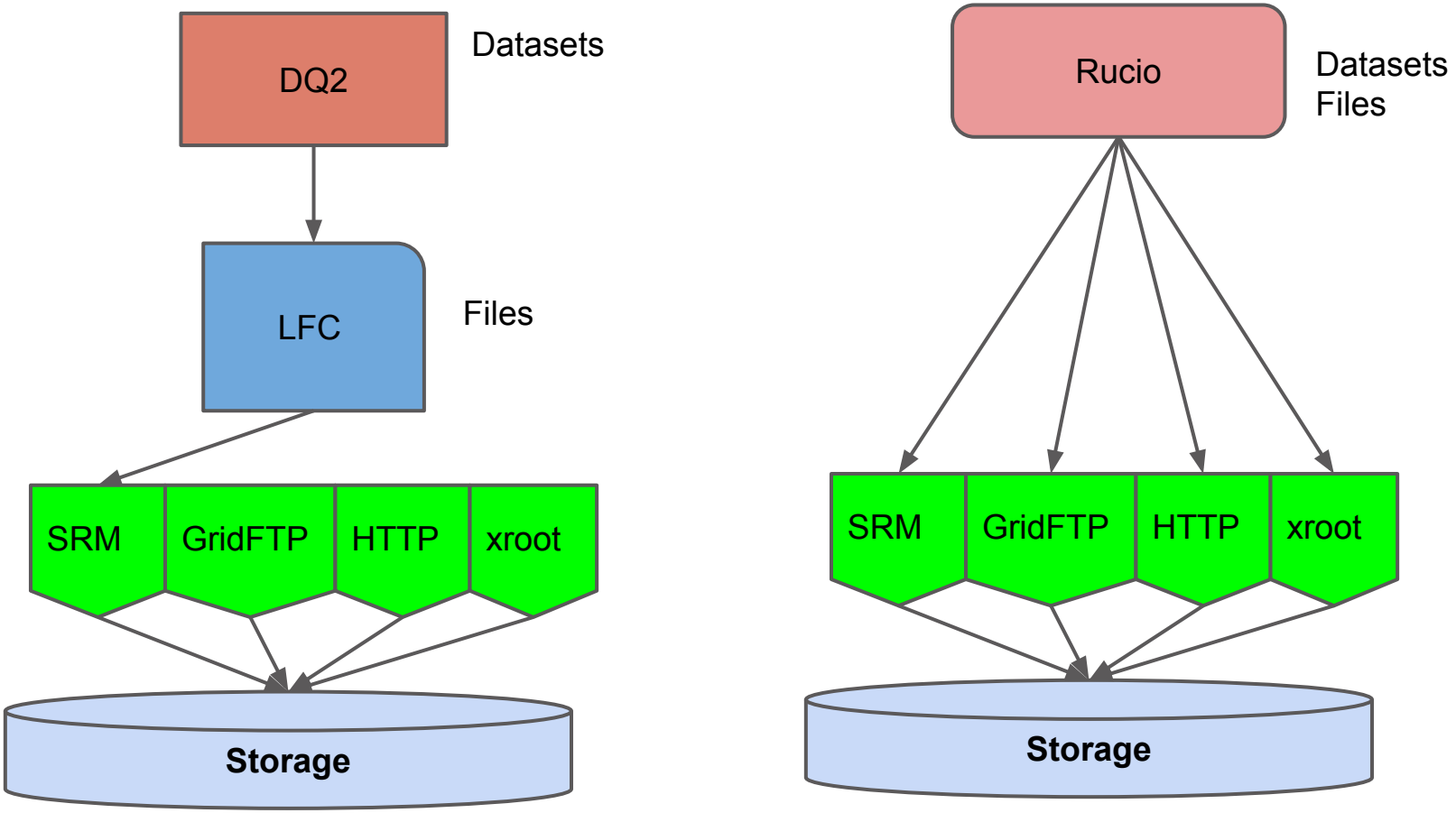
DQ2 ⇨ Rucio Migration

- Rucio is a unified dataset and file catalogue
- The migration has been done in three steps:

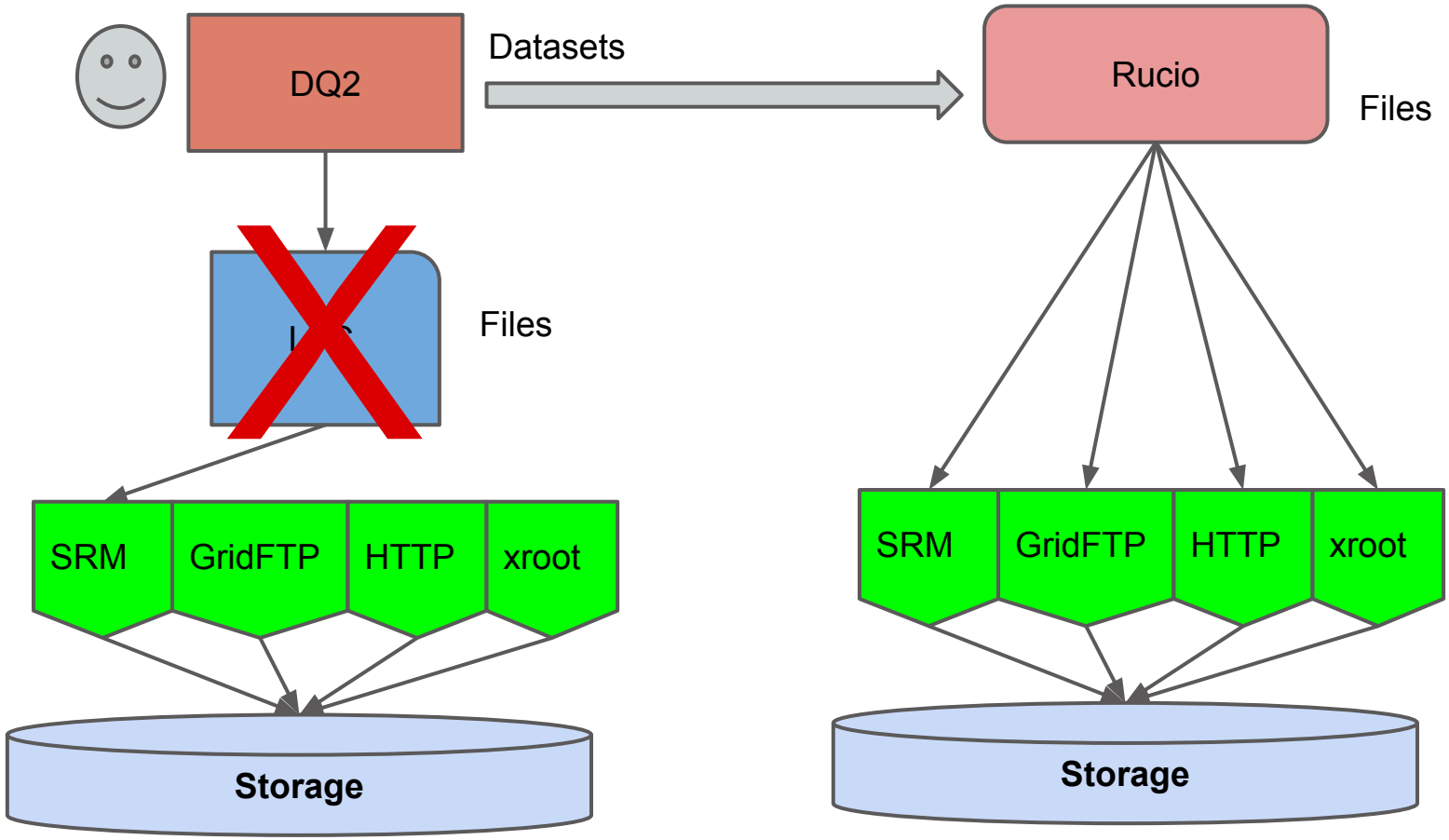
1. **2013 - 2014:**
   - Renaming to deterministic physical file names
2. **March - June 2014:**
   - LCG File Catalog (LFC) ⇨ Rucio
3. **Oct. - Dec. 2014:**
   - DQ2 ⇨ Rucio



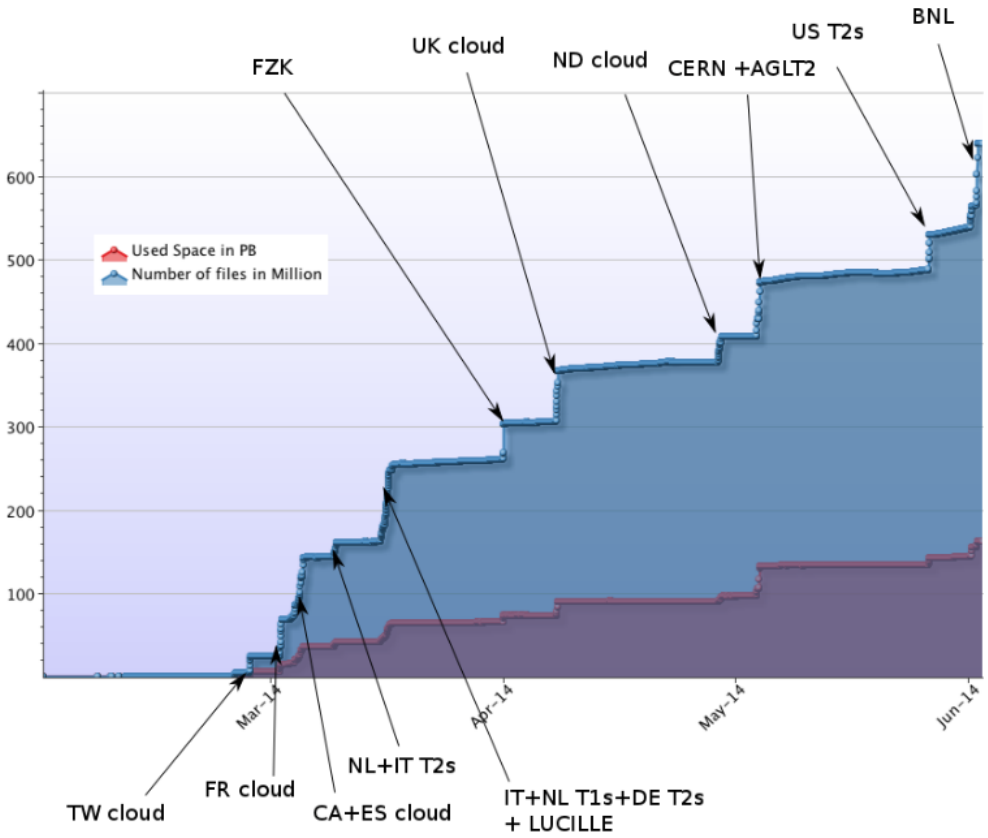
DQ2 ⇨ Rucio Migration Strategy



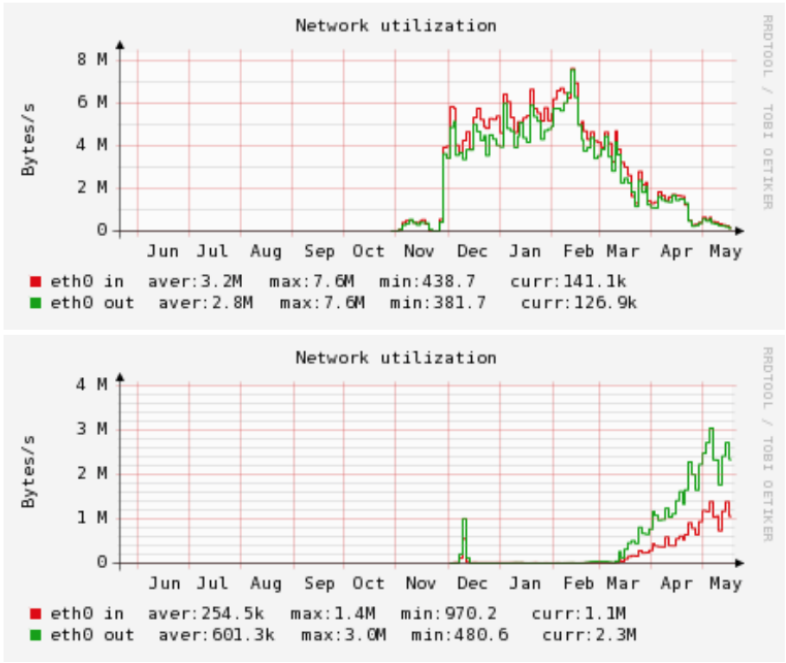
LFC ⇨ Rucio


DQ2
Datasets
Rucio
Files
Files
SRM
GridFTP
HTTP
xroot
Storage
SRM
GridFTP
HTTP
xroot
Storage

LFC ⇨ Rucio

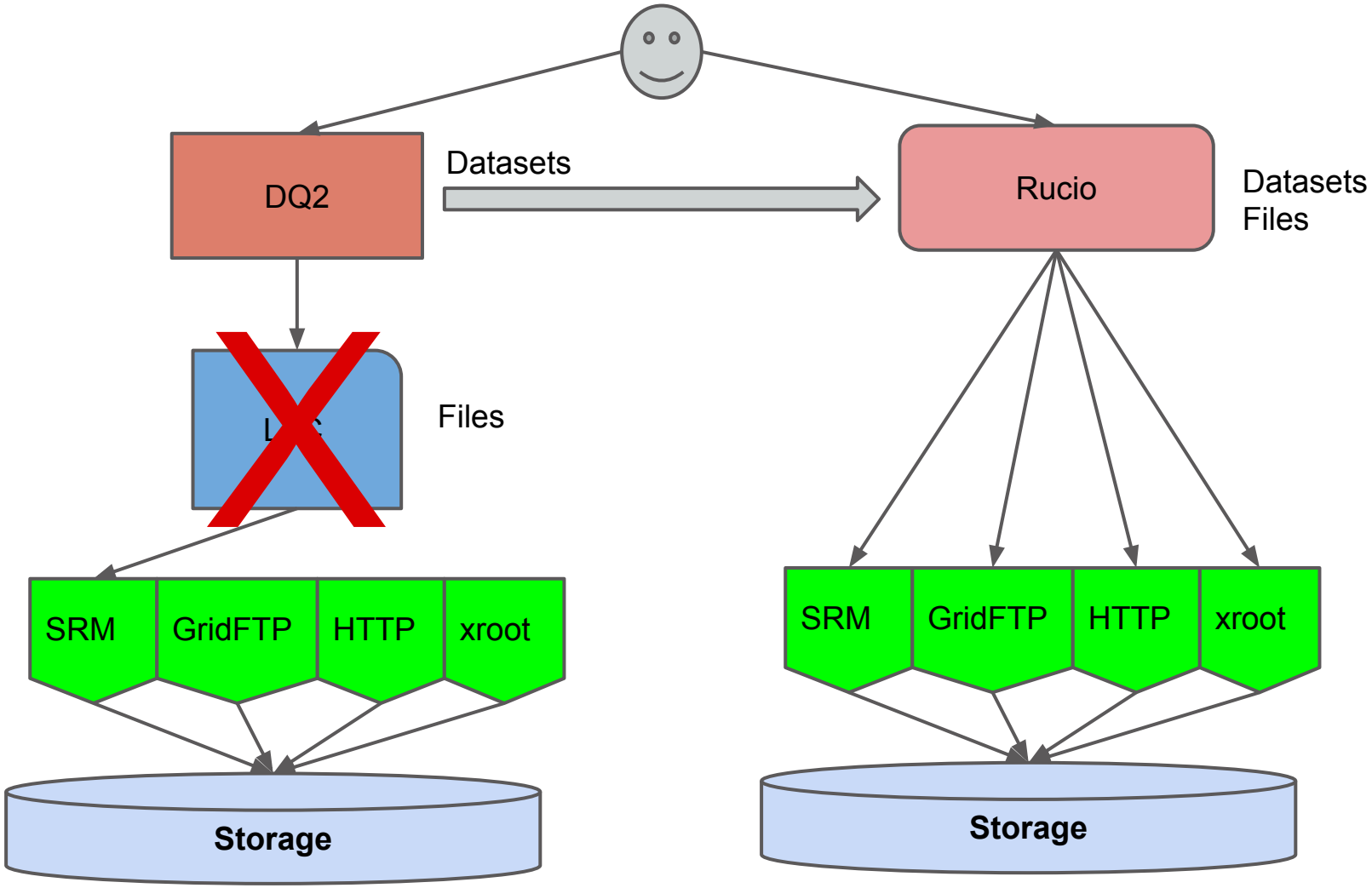


Evolution of the number of files (M) registered in Rucio



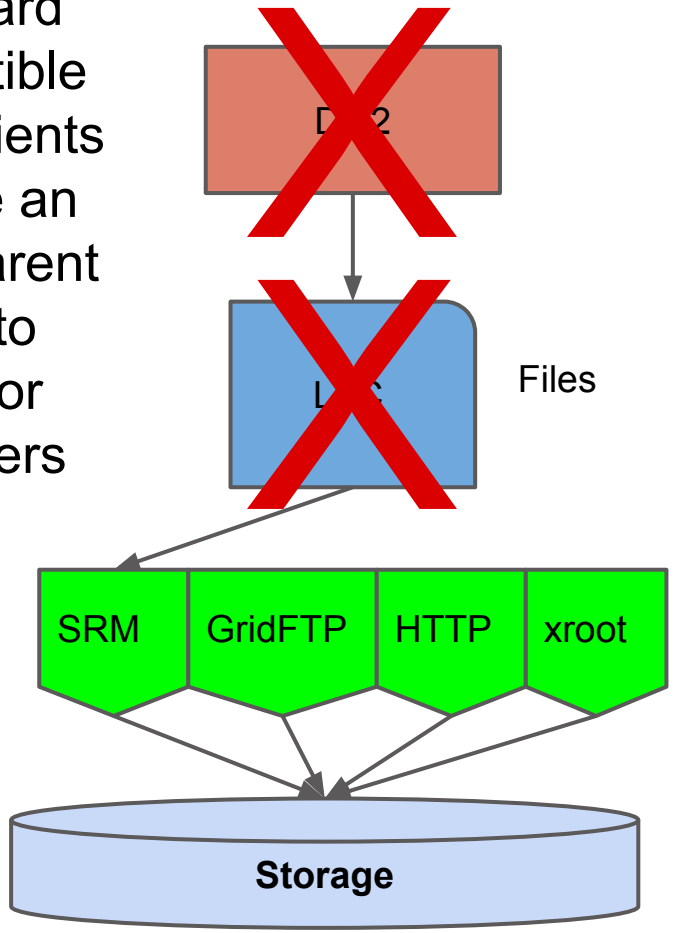
Evolution of the I/O on the LFC servers (up) and Rucio servers (bottom)

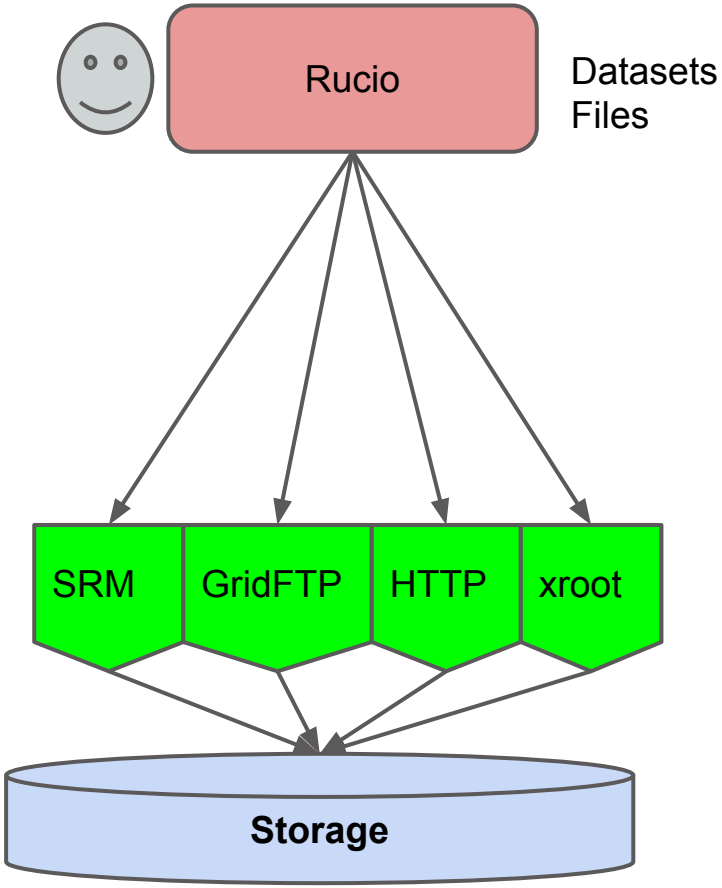
DQ2 ⇨ Rucio


DQ2
Datasets
Rucio
Datasets
Files
Files
SRM
GridFTP
HTTP
xroot
Storage
SRM
GridFTP
HTTP
xroot
Storage

DQ2 Decommissioning

Backward compatible DQ2 clients to have an transparent switch to Rucio for end-users



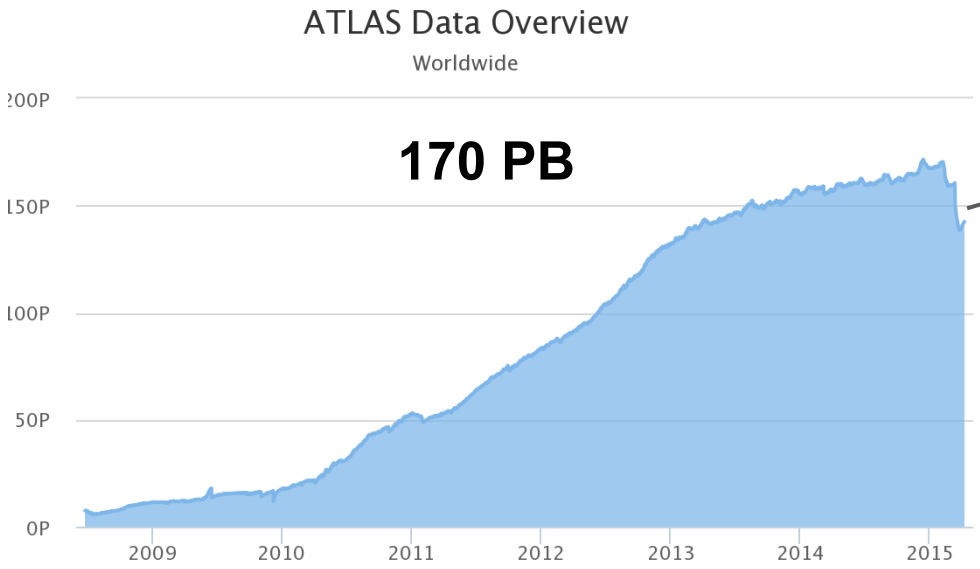


DQ2 ⇨ Rucio: Post-mortem

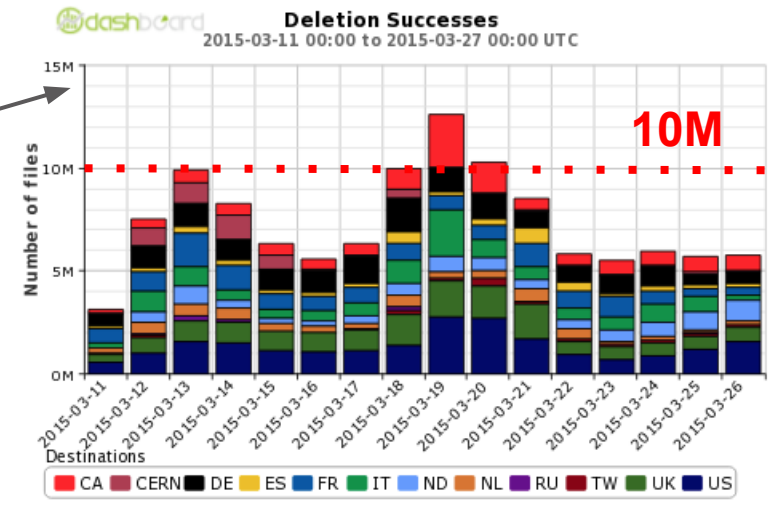
- ‘Almost’ everything worked as planned during the transition
  - Few hiccups and side effects (as expected)
  - Downtime of few hours
- We constantly improved and automated the deployment model
  - ➥ See [224] M.Lassnig
- More visible changes for major external applications, e.g., event based monitoring

➥ DQ2 decommissioned at the end of 2014

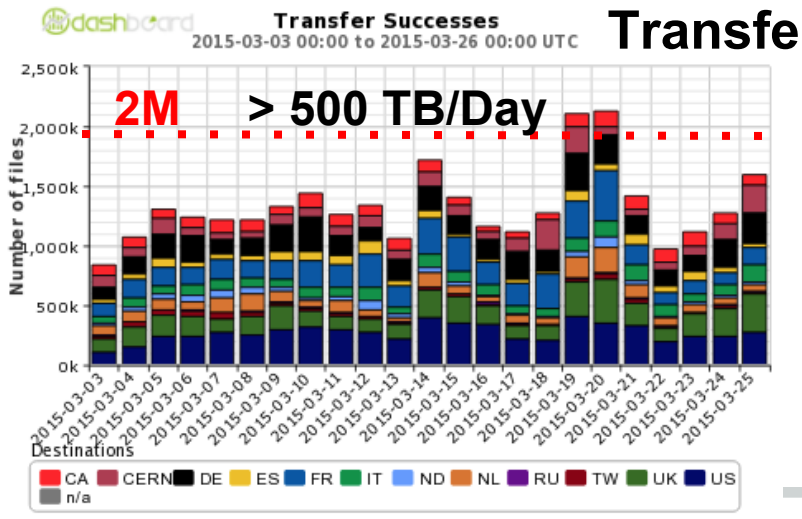
Rucio: Performance



Deletion Campaign



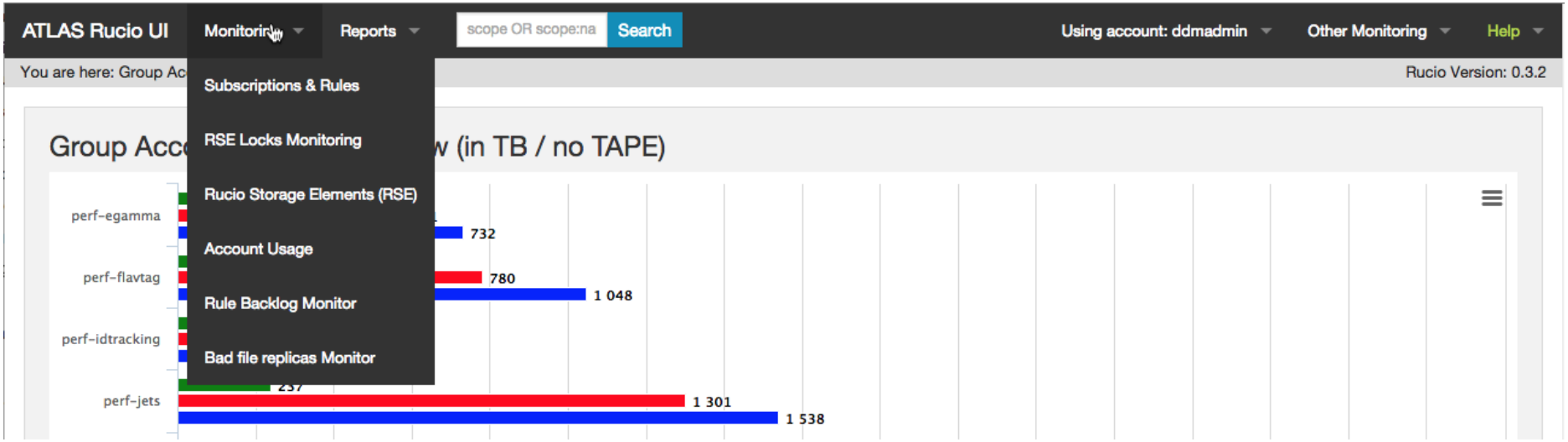
Transfers



- Deletion peaks at 150Hz vs. 60 Hz in DQ2
- Transfer rate > 2M/Day vs. 1M/Day during the Rucio scaling test

➥ Rucio is at a larger scale than DQ2

New features : WebUI



HTTPs based interface allowing to:

- Request, transfer and delete datasets
- Monitor account usage
- Manage site usage, quota and consistency
- etc.

➥See [206] T. Beermann

New features : HTTP Support

- HTTP Rucio redirector can provide HTTP redirection or serve metalink files
- The Rucio server queries the replica table and redirects the query to a selected replica
- The strategy to select the final replica is configurable: random, geoip, selection of the closest replica (IPv4/6 compliant), selected site, etc.

➥See HTTP/WebDAV [157] J. Elmsheuser

Summary

- The Rucio project has a long history, which started well before the end of Run-1
- The migration from DQ2 to Rucio has been as successful as stressful
- Rucio is now fully in production for ATLAS since 1st December 2014
- The performance meets the expectations
- Rucio is at a larger scale than DQ2
- The new features will help to run smoothly during Run-2

CHEP Contribution

- [207] Resource control in ATLAS distributed data management: Rucio Accounting and Quotas
- [157] New data access with HTTP/WebDAV in the ATLAS experiment
- [224] Scalable and fail-safe deployment of the ATLAS Distributed Data Management system Rucio
- [206] Monitoring and controlling ATLAS data management: The Rucio web user interface