Rucio ecosystem

Martin Barisits
on behalf of the Rucio team

RUCIO
SCIENTIFIC DATA MANAGEMENT

ATLAS
EXPERIMENT

Community


eneas
Advanced European Network of E-infrastructures
for Astronomy with the SKA


LCLS


Belle II

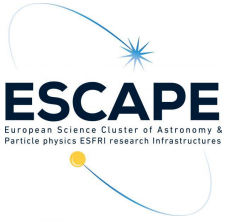

Science & Technology
Facilities Council


COMPASS


CERN

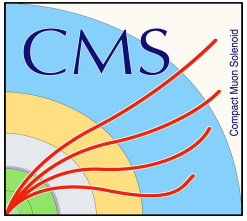

ARC


ATLAS
EXPERIMENT


ESCAPE
European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures


Alpha Magnetic Spectrometer
AMS-02
Europe • Asia
North America


Xe
XENON
Dark Matter Project


CMS
Compact Muon Solenoid

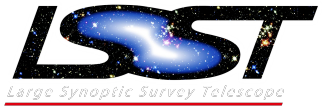

DUNE
DEEP UNDERGROUND
NEUTRINO EXPERIMENT


egi


3D
EISCAT


VIRGO


cta
cherenkov telescope array


LSST
Large Synoptic Survey Telescope


IceCube
SOUTH POLE NEUTRINO OBSERVATORY


FTS


eXtreme DataCloud


Fermilab


BROOKHAVEN
NATIONAL LABORATORY


PAWSEY
supercomputing centre


SKA
SQUARE KILOMETRE ARRAY


LIGO


Open Science Grid


EUXDAT

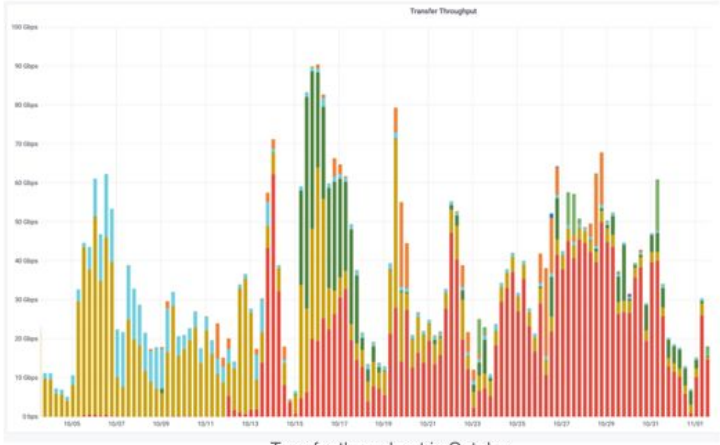
News from the community

- Some recent highlights from the Rucio community
- Not an exhaustive list! More information:
  - 3rd Rucio Community Workshop
  - Weekly Rucio meetings
  - WLCG Grid Deployment Boards
  - …

CMS Migration to Rucio

CMS Rucio Migration

- Transitioned away from PhEDEx and Dynamo, custom CMS tools to community supported Rucio
- SL6 support ended last month, LHC restart in 2021major drivers of the schedule
- Logistical challenge to accomplish this transition with zero downtime in production and analysis systems
- Large effort by many people on CMS
- Fermilab people involved: E. Vaandering, N. Smith, Y. Guo, I. Mandrichenko, C. Huang, N. Ratnikova, F. Garzon, J. Amado
- Significant work to come as we optimize CMS use of Rucio, plan for LHC Run4 with over 10x the data volumes of previous runs
- All using recent technology: Python3, Helm, Kubernetes, Docker. Prometheus, Elastic Search, Grafana for monitoring



Transfer throughput in October
Left: cyan and mustard are PhEDEx; Right other colors are Rucio

E. Vaandering (CMS/FNAL)

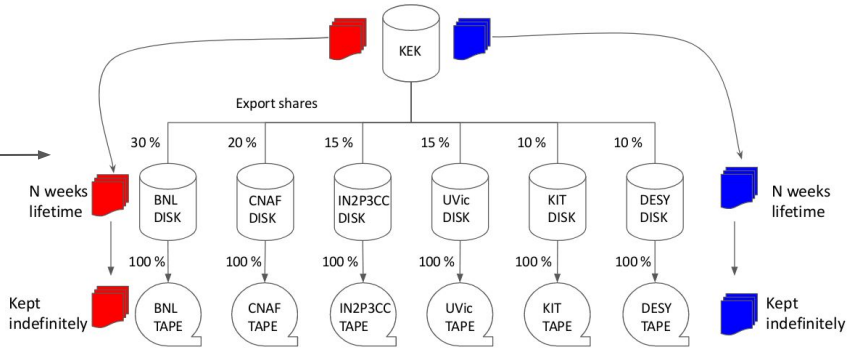
Belle II migration summary

- Migration happened from 14th to 19th January 2021:
  - Jan 14th: Draining of the Grid
  - Jan 15th: Stop services + LFC dump
  - Jan 16th-17th: LFC import + basic checks
  - Jan 18th: Restart of Dirac services + user and production test jobs
  - Jan 19th: Full restart
- Around 100M files imported from LFC to Rucio
- It involved people from many different timezones (JST (UTC+9), CET (UTC+1), EST (UTC-5), CST (UTC-6)). Quite challenging !
- Everything went smoothly. Just a few issues found, but quickly fixed
- The whole migration is the result of many months of preparation and testing

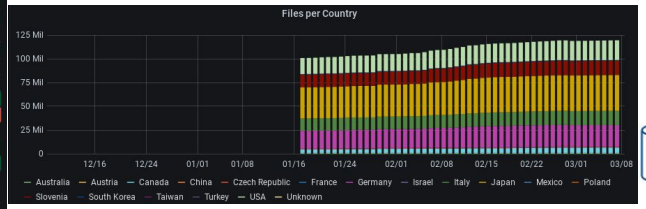
C. Serfon (Belle II/BNL)

Features implemented for Belle II

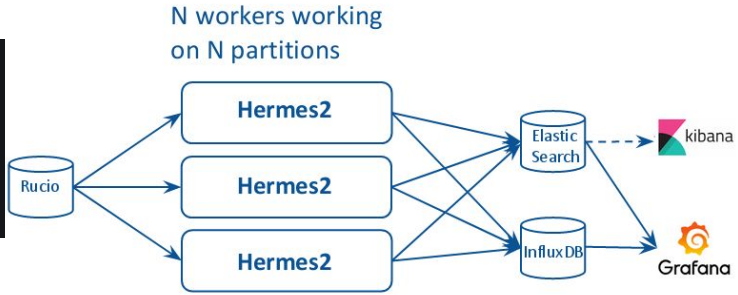
- Rucio File Catalog plugin in BelleDIRAC
- Chained subscriptions
- Simplified monitoring stack using new Rucio daemon
- Grafana dashboards featured to Belle II



Transfer efficiency



Storage accounting



Monitoring stack

C. Serfon (Belle II/BNL)

ESCAPE Full-Dress-Rehearsal

R. Di Maria (ESCAPE/CERN)

ESCAPE DataLake

RSE	Quota	WM
ALPAMED-DPM	100 TB	10 TB
CNAF-STORM	10 TB	1 TB
DESY-DCACHE	40 TB	4 TB
EULAKE-1	300 TB	30 TB
GSI-ROOT	1 TB	10 GB
IN2P3-CC-DCACHE	60 TB	1 TB
INFN-NA-DPM	68 TB	5 TB
INFN-NA-DPM-FED	46 TB	5 TB
INFN-ROMA1	2 TB	200 GB
LAPP-DCACHE	10 TB	1 TB
LAPP-WEBDAV	100 GB	90 GB
PIC-DCACHE	28 TB	27.99 TB
PIC-INJECT	28 TB	27.99 TB
SARA-DCACHE	98 TB	140 GB

ESCAPE DataLake

- Total Quota:

  891 TB

- Watermark:

  113.44 TB

- 10+ RSEs
- 9 sciences
- 50+ accounts

8 November 26th, 2020 Riccardo.Di.Maria@cern.ch

Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Full-Dress-Rehearsal

R. Di Maria (ESCAPE/CERN)

DataLake **24-hour Full Dress Rehearsal** Objectives

Goal: exercise covering **experiment data workflow** needs on a single day (data injection, replication, and access).

Three fold goal: perspective from **scientists**, perspective from **sites**, and the assessment of the **ESCAPE DataLake tools and services** under **pseudo-production conditions**: RUCIO, FTS, CRIC, IAM, perfSONAR, monitoring, QoS, clients, etc.

DataLake Objectives

- Stable infrastructure:
  10 sites, 5 storages technologies, 3 protocols ✓
- Monitoring: automated tests ✓
- 1M files - to demonstrate stable and sizeable data movement ✓
- 3 QoS: CRIC as reference point ✓

Compute Integration Objectives

- Interactive access to files in DataLake
  e.g. using JupyterLab Notebook ✓
- Batch access to files in DataLake ✓
- Data access through caching layer ⊖
- Process data locally downloaded ✓
- Use cases covering small-to-large files range; simple from compute perspective (focus on mimicking data flow) ✓

QoS Objectives

- Demonstrate compute-driven QoS staging ⊖
- Demonstrate cost-performance QoS trade-off ✓
- Demonstrate VO-specific workflows (data lifecycles) ✓
- Demonstrate data injection with targeted QoS ✓
- Demonstrate computational match-making ⊖

Network and Monitoring Objectives

- Test suite development ✓
- Rucio events into a dashboard ✓
- Dashboard adjustments ✓
- Develop and run rucio-level tests ✓
- Develop example data lifecycle definitions and scripts ✓

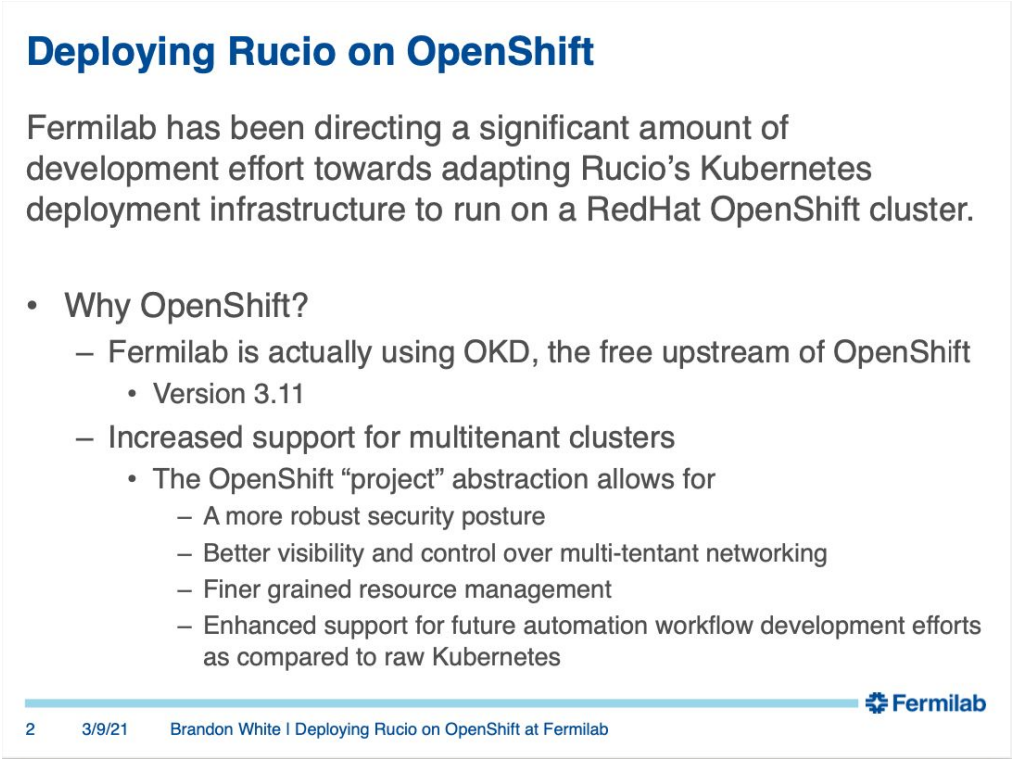
AAI Testbed Objectives

- User enrollment flow in place ✓
- X.509/VOMS AuthN/Z in place ✓
- Token-based AuthN/AuthZ in place ⊖
- Continuous monitoring tools to assess that AuthN/Z work as expected at sites ✓

9 November 26th, 2020 Riccardo.Di.Maria@cern.ch

Funded by the European Union's Horizon 2020 - Grant N° 824064

DUNE & Openshift @ FNAL



Deploying Rucio on OpenShift

Fermilab has been directing a significant amount of development effort towards adapting Rucio's Kubernetes deployment infrastructure to run on a RedHat OpenShift cluster.

- Why OpenShift?
  - Fermilab is actually using OKD, the free upstream of OpenShift
    - Version 3.11
  - Increased support for multitenant clusters
    - The OpenShift "project" abstraction allows for
      - A more robust security posture
      - Better visibility and control over multi-tentant networking
      - Finer grained resource management
      - Enhanced support for future automation workflow development efforts as compared to raw Kubernetes

Fermilab

2 3/9/21 Brandon White I Deploying Rucio on OpenShift at Fermilab

B. White (FNAL)

DUNE & Openshift @ FNAL

Fermilab Usage of Rucio on OKD

- The DUNE experiment is leading the way in this endeavor
  - Most of the groundwork for adapting Rucio's Kubernetes deployment system to OpenShift has been completed
    - OpenShift templates differ in slight, but critical, ways from Kubernetes templates
    - Utilizing Helm version 2 solely for template-based generation of deployment configurations, but not for deployment management
    - OpenShift 4 will allow for much better integration with Helm 3
  - Production testing of the service is ongoing
    - Successfully transferred approximately 150 terabytes to RAL in the first week of March 2021
    - Testing utilization of an FTS3 deployment also running on OpenShift as the transfer tool
- The ICARUS experiment and AccelAI project at Fermilab are also looking to utilize Rucio on OpenShift going forward

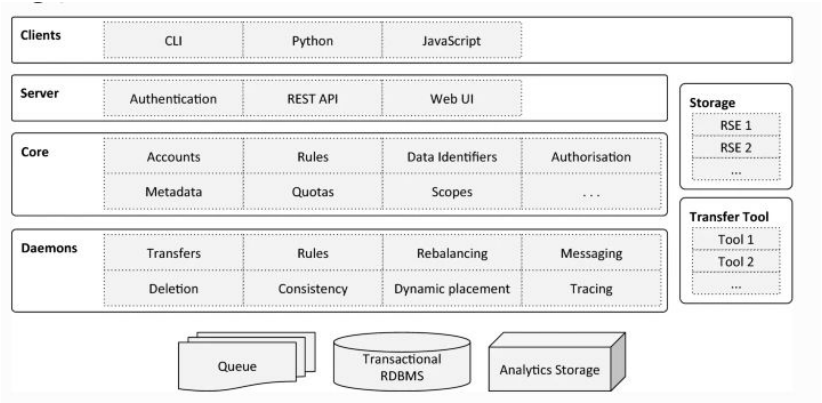
Fermilab

3 3/9/21 Brandon White I Deploying Rucio on OpenShift at Fermilab

B. White (FNAL)

Project organisation

- Rucio consists of 29 components (Vertical/horizontal slices through the architecture)
  - 1-3 persons responsible for the maintenance and feature development of a component
  - Point of contact in case of issues, coordinating development effort and contributions to the component
  - Historically all components were lead by ATLAS people



Project organisation

- Rucio is a community project; objective is to foster long-term community involvement not only one-off contributions → formally include other communities in the development and evolution of the project
  - Personpower, expertise, community “buy-in”
- New component (co-)leaders from Oct 2020
  - OIDC Authentication: Rizart Dona (ESCAPE/CERN)
  - Core & Internals: Gabriele Fronze (VIRGO/LIGO)
  - Multi VO: Ian Johnson (RAL)
  - Policies: James Perry (DUNE/U. Edinburgh)
  - Probes & Alarms: Eric Vaandering (CMS/Fermilab)
  - REST & API: Benedikt Ziemons (CERN)
  - Transfers: Eric Vaandering (CMS/Fermilab)
  - DIRAC: Cedric Serfon (Belle II/BNL)

Project communication

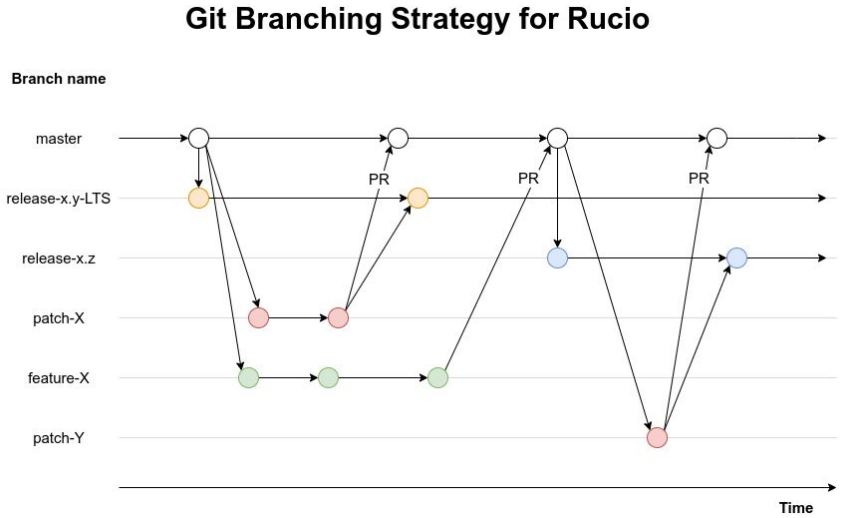
- Very active Slack community (Developers, Admins, Users)
  - 300 members, 10k messages per month
  - → Community support model
- Weekly Developers meeting
  - **Community DevOps roundtable**
    - Operational issues, deployment topics, integration with external applications, etc.
  - **Developers roundtable**
    - Priority follow up on milestones for the next feature release
    - Other development topics
- Feature release planning meeting to set milestones for each feature release
- Yearly workshops
- Yearly coding camps

Release plan

- Feature releases in 2020
  - **1.22** “Green Donkey” **February 2020**
  - **1.23 LTS** “The Incredible Donkey” **June 2020**
  - **1.24** “Aquadonkey” **November 2020**
- Feature releases in 2021
  - **1.25** “Rat-Donkey” **February 2021**
  - **1.26 LTS** “Donkey League of La Mancha” **June 2021**
  - **1.27** “Batdonkey v Superdonkey” **November 2021**
- Long Term Support releases maintained with security/critical fixes for 2 years
  - LTS model was introduced after Rucio Workshop 2019
  - Very well received by the community, several communities deploy LTS releases
  - Effort to maintain these LTS releases minimal

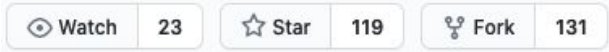
Code management model

- After 8 years switched to a new code management model in June 2020
- Advantages:
  - Similar to most other open source projects
  - Easy to understand/contribute for newcomers
- **Very positive** feedback from developers
- Complete overhaul of CI system



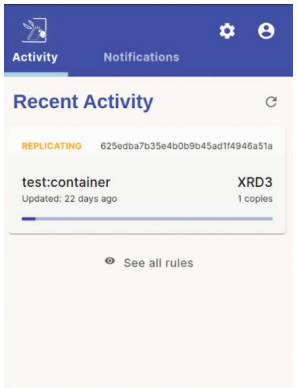
Contributions in 2020

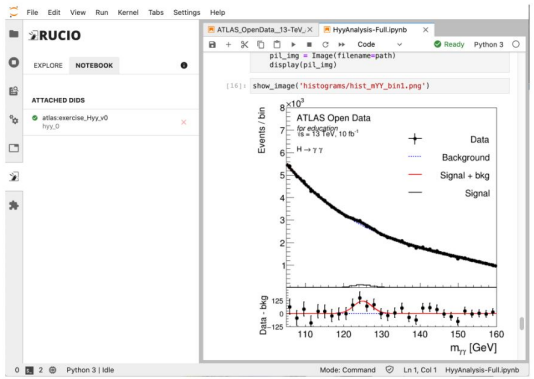
- 30 different contributors
  - Mixture between very active and one-off contributors
- Lines of Code added
  - ATLAS ~35k LOC
  - Non-ATLAS ~19k LOC
  - Disclaimer: LOC not the best indicator in terms of feature-impact and effort, but still gives an insight of community involvement
- 131 github forks
- 39 people part of the github team (Repository permissions)



Google Summer of Code & Season of Docs 2020

- Three Rucio Summer of Code projects finished this year
  - Rucio SWAN/Jupyter integration together with SWAN team
    - Access Rucio data directly from SWAN/Jupyter notebooks
  - Rucio desktop app (+ fuse plugin)
    - Lightweight desktop application for search, download and notifications
  - Rucio NLP Bot
    - Support chat bot based on NLP
- Google Season of Docs
  - New docusaurus based documentation
  - https://rucio.cern.ch/documentation/ (DRAFT)
  - Simplified editing (Markdown), auto-publish via github actions
    - Encourage community to evolve documentation
  - Will replace current documentation in 2021





Kubernetes & Containers

- Big benefits not only for deployment and scaling but also for isolation of the software due to full control of the environment with containers
- Common development effort with CMS to establish common deployment charts
  - ATLAS and CMS run the same deployment charts
  - Recommendation to new communities, due to ease of deployment
- Majority of ATLAS Rucio daemons now run on K8 cluster instead of puppet
  - All Rucio K8s containers running on Py3
  - Memory intensive daemons still problematic (need larger memory profile) → Improvements coming
  - Will fully switch ATLAS to K8s in 2021

Capability based auth & AAI

- Capability based authentication (Available since Rucio 1.22.0)
  - Funded by european XDC project
- Ramping up continuous tests on Rucio DOMA instance
  - Successfully demonstrated full chain transfers (Rucio ⟶ FTS ⟶ dCache)
  - Successfully demonstrated rucio uploads to storage (Rucio ⟶ dCache)
- Feature works well, still some open decisions which will require Rucio development
  - Large-scale tests very important due to involvement of the entire infrastructure
- Long road still ahead, but Rucio is in a good state

Quality of Service

- Integration of storage quality of service in Rucio
  - User expresses QoS needs with replication rules
  - Rucio communicates QoS needs to storage
  - Rucio **transitions** (instead of TPC) data in between QoS zones
- Alignment with DOMA QoS group, where technically necessary
- Lots of interest from different communities
- Development ongoing
  - Strong dependency on features developed in FTS and storage
  - Will require large-scale testing
    - Prototypes will be tested within ESCAPE

Globus (Online)

- Globus Online transfer functionality added in 2019/2020 by Matt Snyder (BNL)
- Important functionality for Data ingress/egress to US-based HPCs
- Several successful (small scale) test transfers from fts-managed RSEs to globus-managed RSEs within ATLAS
- Late 2020 started a second development phase to bring the functionality in a production-ready status
  - Part of Rucio 1.25 release
- Lots of interest from other communities as well

Conclusion & Summary

- Lots of activity in the Rucio community
  - Major experiments just migrated
  - Much interest from communities within HEP & beyond
- Streamlining support, development, and documentation
  - Making it as easy as possible for newcomers to contribute
  - Support from the community for the community
  - Common operations: Benefiting from each others operational experience
- Project organisation
  - Rucio is a common software project!
  - Increase formal responsibility of non-ATLAS developers in Rucio
- Important ongoing developments
  - Kubernetes, AAI, QoS, GO, and many more!
- Rucio Community Workshop 2021 → To be announced soon!

More information

Website	http://rucio.cern.ch
Documentation	https://rucio.readthedocs.io
Repository	https://github.com/rucio/
Images	https://hub.docker.com/r/rucio/
Online support	https://rucio.slack.com/messages/#support/
Developer contact	rucio-dev@cern.ch
Publications	https://rucio.cern.ch/publications.html
Twitter	https://twitter.com/RucioData