

# Rucio and ATLAS

*Perspectives from ATLAS Distributed Computing*

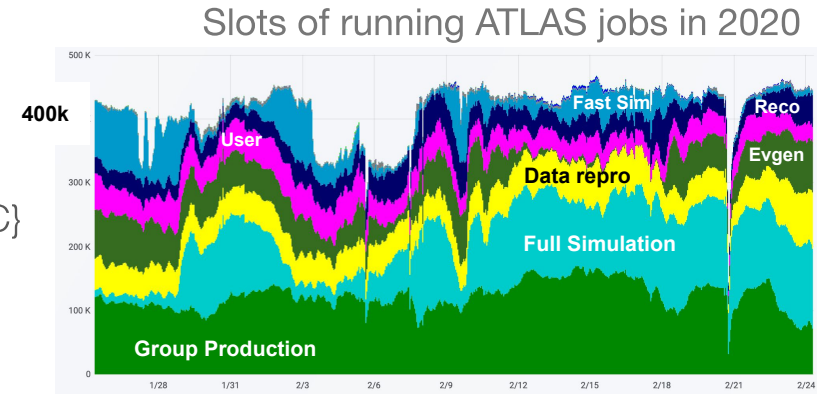
3rd Rucio Community Workshop, 10th March 2020



David South (DESY)



- ADC comprises the hardware, software and operations needed to support distributed processing, simulation and analysis of ATLAS data
  - ADC are therefore responsible for the *delivery of distributed computing* to ATLAS, and for the *ongoing developments* to support evolving ATLAS needs
  - **“To enable as much physics as possible without getting in the way”** *D. Cameron*
- ADC is running 24 hours, 7 days a week, 365 days / year
- Activities can be broken down into key areas
  - Workflow Management System (**WFMS**)
    - PanDA, Pilot(2), Harvester, ProdSys
  - Distributed Data Management (**DDM**)
    - **Rucio**, activities in DOMA {access, QoS, TPC}
- In addition
  - Monitoring, Analytics, Daily Operations
  - Ongoing development, adaptation, innovation..





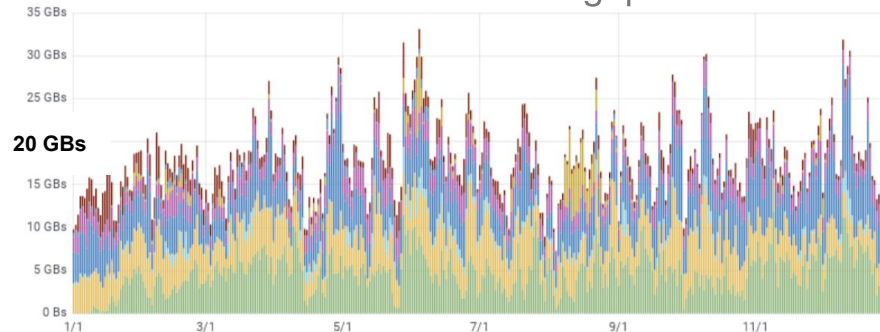
# A large amount of data - and we move it around

## ATLAS data volume managed by Rucio

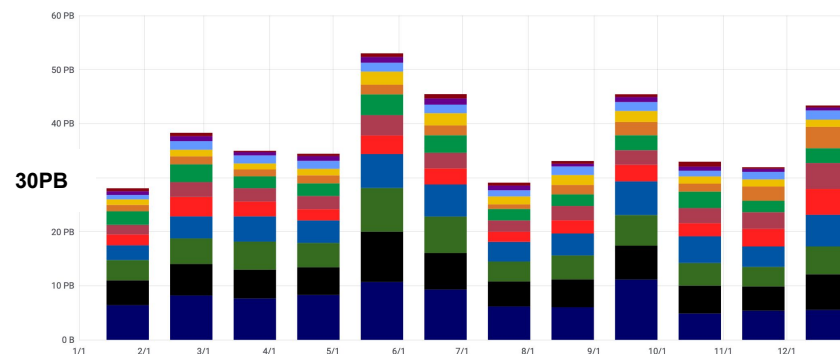


- Average transfer throughput of 18 GB/s at a rate of 22 Hz (with peaks up to 33 GB/s at 56 Hz)
  - Consistently transferring 30PB+ / month
- Why do we transfer so much?
  - Pilot based job scheduling system, late binding
  - Chained production workflows with wall times of O(10-24h), storage of intermediate outputs
  - Movement of production input files to available CPU slots, aggregation of outputs at larger sites

## 2019 Transfer Throughput



## 2019 Transfer Volume



- Rucio arose from its predecessor, DQ2
  - Need for a more scalable, robust and efficient model with additional functionality
  - Development of initial architecture 2011-2013
  - Stress tests, commissioning and migration 2013-2014
- Integration with other day-to-day DDM
  - PanDA and ProdSys, gaining user familiarity
  - Monitoring of storage and transfers
  - Replication policies and accounting
- And just when you think you're ready.. things happen..
  - Experience gained invaluable since full adoption
- Project moved from development to operations in 2014
  - This was not the end but the beginning!
  - Still very active with continuous development, adaptation
  - ATLAS continues to have weekly DDM DevOps meetings, devoted to the discussion of new features.. and fixes..





... ooopss

- November 2014:
  - Just few days before the final migration

**File Deletion: Incident Report**

**Summary**  
Approximately 500k files were physically deleted from storage between the 24 November and the 27 November because of a configuration problem of the Rucio integration infrastructure. Out of these 500k files approximately 40k were single replica files. Those 40k files were secondary only.

**Explanation of incident**  
The Rucio integration infrastructure serves as a testbed, *stresstest*, as well as functional test, including third party systems, necessary for ATLAS DDM. It is therefore necessary to operate it selectively next to the production system. Safety measures are in place, such that the integration infrastructure does not interfere with the production infrastructure. Due to the caching of storage access protocols, however, this safety was effectively circumvented, leading to the construction of file access paths of the production system within the integration system. To effectively disable parts of the production system, a privileged grid certificate is needed. The integration system had this privileged certificate installed, due to an ongoing test with an external system (TSS) that required it. The deletion service pointed to this privileged certificate due to a configuration error. This led to the attempted deletion of 1.1 million files, of which 600k were non-existent files created by the Rucio emulation using the integration infrastructure, 400k were files with other replicas and hence recoverable, and a minority 60k files which were unique replicas without a potential backup. Investigation into recovery and possible reconstruction of files is ongoing.



A. di Girolamo, 1st Rucio Workshop



<https://dev.to/javinpaul/10-things-every-software-developer-should-know-399e>

- Rucio Development Plans
  - Full Kubernetes production deployment, including autoscaling
  - Storage QoS support in Rucio
  - Metadata unification (Multiple metadata backends addressable in single interface)
  - Fully scalable rule components
  - RPG Integration into Rucio subscriptions
  - Further work on Token-Based authentication
  - Stronger integration of archive workflows (TPC of constituents? Recovery?)
  - Documentation overhaul
- Proposed Ops Features
  - Scalable Reaper and Judge for deletion of large datasets/small files
  - Multihop required for CTA
  - Global quota for group space consolidation
  - Plugin transfer tool (ARC, GlobusOnline)
  - Upload service to replace SCRATCHDISK
  - Subscriptions aware of other replicas with subscription consistency checker
  - RPG integration into Rucio with delayed rules
  - Automatic releases on CVMFS of rucio clients

- Rucio Development Plans

- Full Kubernetes production deployment, including autoscaling
- Storage QoS support in Rucio
- Metadata unification (Multiple metadata backends addressable in single interface)
- Fully scalable rule components
- RPG Integration into Rucio subscriptions
- Further work on Token-Based authentication
- Stronger integration of archive workflows (TPC of ...)
- Documentation overhaul

- Proposed Ops Features

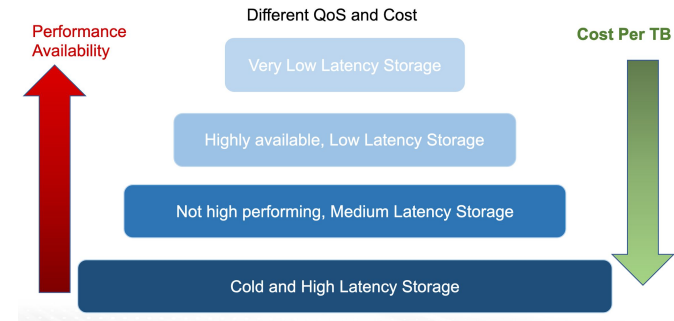
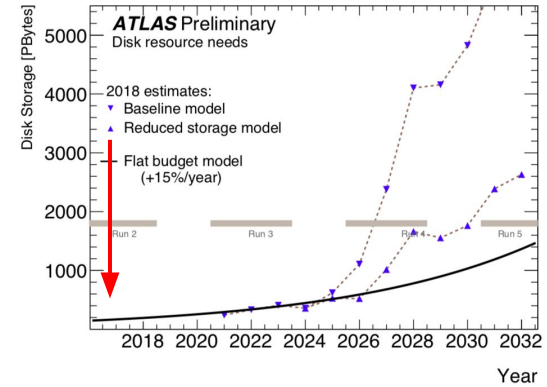
- Scalable ...
- Metadata ...

This is not the talk to describe all the new rucio features and plans listed here

Take away: Lots going on. Pick out a few recent examples of particular relevance to ATLAS

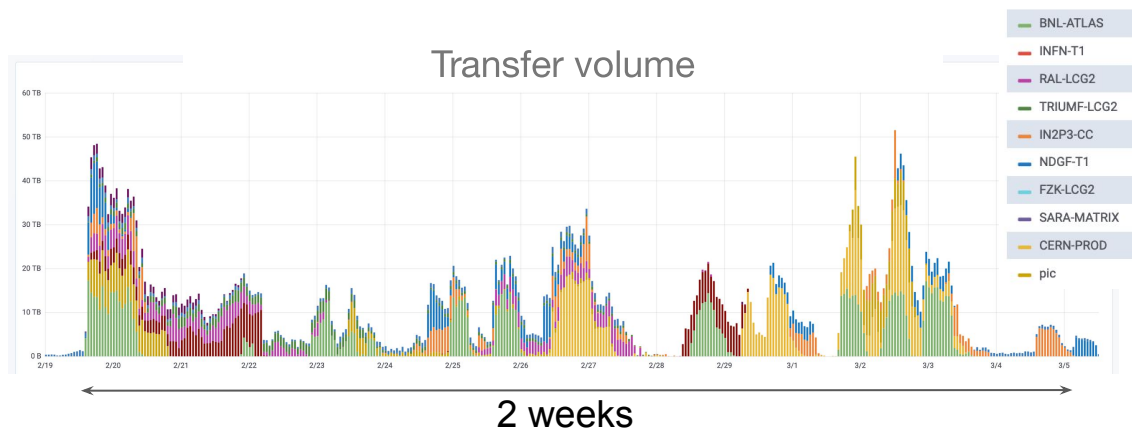
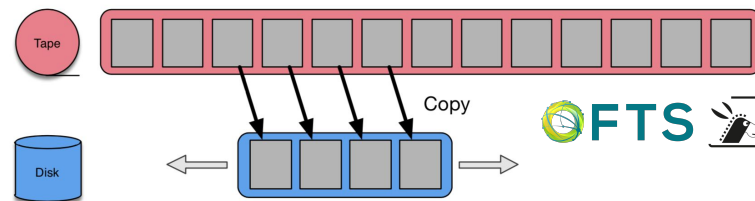
- ... consolidation
- ... (e.g., GlobusOnline)
- ... to replace SCRATCHDISK
- ... subscriptions aware of other replicas with subscription consistency checker
- RPG integration into Rucio with delayed rules
- Automatic releases on CVMFS of rucio clients

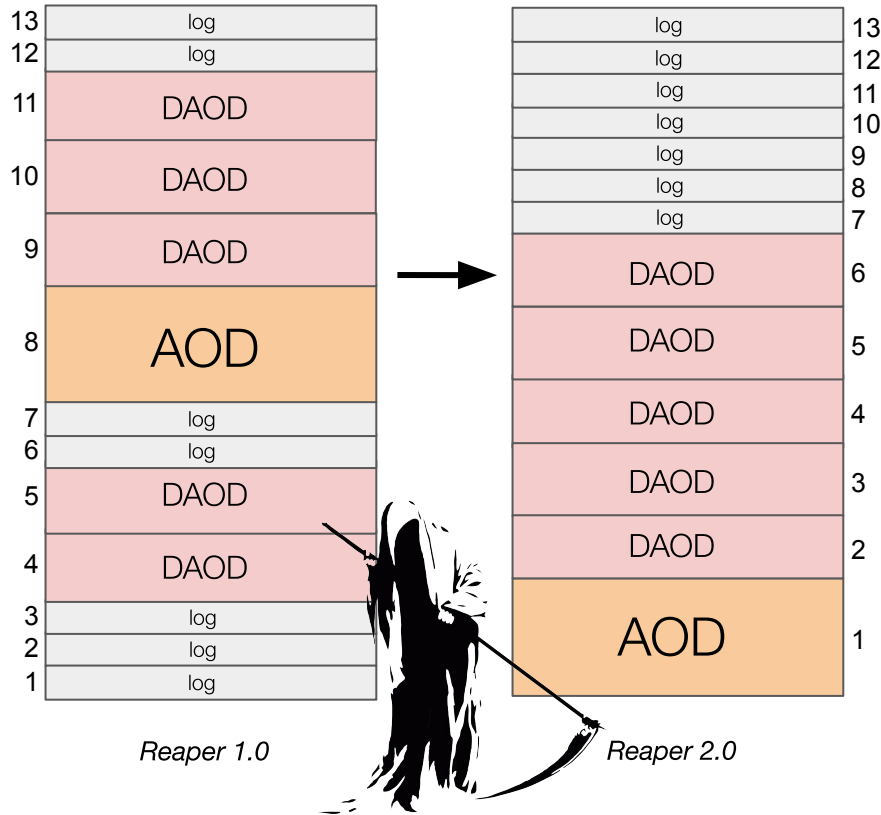
- The current funding model for disk space and the projected disk storage requirements diverge for run 4 and beyond
  - Work is being done on reducing the disk footprint by reducing event size in analysis data formats. But what else can be done?
- One possible solution to alleviate this imbalance: introduce layers with different availability and cost, “Quality of Service”
  - Different types of data: different types of storage
- Another area is the use of caches such as xCache and Data Lakes
  - Popular data more readily available, shared between sites
  - Introduction of more diskless sites and cached data access
- What are the implications for Rucio?
  - QoS becomes explicit part of the replication rule
  - Specify QoS requirement together with RSE selection
  - Multiple QoS classes within one storage system
  - Reduction of load Rucio Ops with use of caches/diskless sites





- Another handle on this is to better utilise the existing disk resources we have: the Data Carousel
- On demand reading from tape without pre-staging
  - Rolling disk buffer with adjustable size, tuned to suit available resources and production requirements
  - E.g. DAOD production, where input is full of AOD: reduced disk footprint
- Development: close collaboration between T1s, FTS and Rucio
- Now used by ATLAS in production: Run 2 DRAW\_RPVLL reprocessing
  - 2017 total RAW inputs: 5.7PB
  - < 1PB on disk at one time
  - Several 100TB processed and removed from disk after ~ days

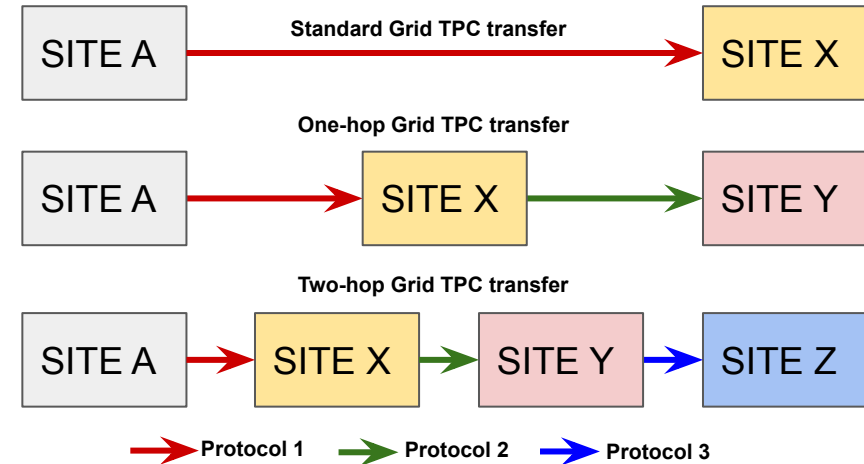




- A third example of optimisation of disk usage via the Rucio deletion daemon, Reaper
- Large scale deletions can be slowed by the presence of many log files: little immediate space gain if they're ahead in the queue
- Optimisation of deletion may provide more space, in a more timely manner
  - Build in some intelligence: re-ordering and/or group files before deletion performed
  - Can we go further and bundle/zip/archive files together before deletion?

- ATLAS grid sites are not homogeneous, with a large variety storage technologies **and** transfer protocols coming in via DOMA TPC effort to replace GridFTP
- Rucio Mover: crucial interface between WFMS and DDM, to handle the large number of data transfers
- Multi-hop transfers: Grid TPC transfers via intermediate site(s) with compatible transfer protocols
  - Developed within Rucio with FTS support
  - Possible for Rucio to request Multi-hop transfers for RSEs without connectivity
  - Blocking of existing connections also possible, to optimise data flow

Storage (sites)	SRM/ GridFTP	XRootD	WebDAV	WLCG SRR
Echo (2)	GridFTP		WIP	ATLAS json
EOS 4.6.9+ (1)	GridFTP	only testbed	only tokens	auto
DPM 1.13.2+ (33)				auto (https)
dCache 5.2+ (30)				cron (doc)
Dynafed (1)				ATLAS json
StoRM 1.15+ (17)	SRM	native(?)	only tokens	cron (doc, *)
XRootD 4.11+ (~6)	GridFTP		WIP	cron (slacxrd)

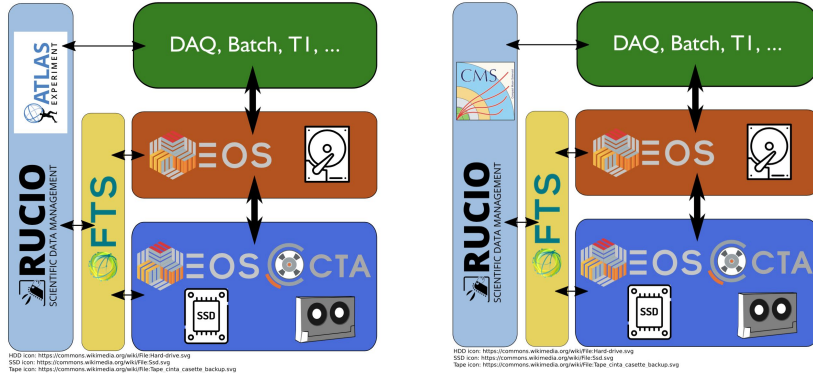


- Resources and Compute
  - Opportunities provided by Kubernetes, actively scaling to current needs and availability
  - Integrating ever more heterogeneous resources such Google Cloud - cost/benefit analysis?
  - Authentication issues on “non-standard” (non-X.509) resources?
- Storage and Networks
  - Full integration of QoS in a “transparent” way, on multiple storage and transfer technologies
  - Caching/xCache/Data Lakes and disk-less sites, non-gridFTP TPC with http/xrootd
  - Use of Virtual Placement ideology and Lakes as storage constraints become tighter
  - Impact on network of finer granularity (event by event) access models such as iDDS
  - FTS is vital to all TPC transfers - and drastic improvements are possible, with time/personpower
- Monitoring and Analytics
  - Changes to metadata full exploitation of dataflow traces
  - Further drive for more Operational Intelligence to reduce manual operations

- Pleasing for ATLAS to see the adoption of Rucio by other experiments and communities
- Makes sense as many things are already shared or similar (infrastructures, data models..)
  - Common experiences; successes, pitfalls..
- Importantly not only users joining but also developers



Taken from M. Barisits' talk in Oslo - almost certainly out of date



- Another recent example, presented as part of the plans for the CTA migration, shows identical ATLAS and CMS workflows with Rucio at the core
- Sign of things to come?

- Resources are often scarce, personpower limited
  - Yet we are often doing similar with things in similar environments
  - This is particularly true at multi-VO sites
  - We share a lot of the same problems - and the same solutions
- Shared infrastructure
  - Dynamic scaling between experiments
  - Common monitoring accounting and procedures
- Operational Intelligence
  - Automating computing operations, e.g. automatic ticketing of sites, ML algorithms to suggest shifters most probable cause of the problem, autodiscovery of network paths degraded

- Distributed Data Management
  - It's complicated and it's difficult
- Rucio makes it easier
  - It works. It's stable. But there's always more to do.
- From Development to Operations and back to Development (in parallel with Operations!)
  - We don't stand still
- As the community builds, we can work more together
  - Strengthen cross collaboration, joint projects, closer cooperation