Methods of Data Popularity Evaluation in the ATLAS Experiment at the LHC

M. Grigorieva, T. Beermann, O. Chuchuk, A. Di Girolamo,
A. Klimentov, M. Lassnig, M. Schulz, A. Sciabà and E. Tretyakov

Data Popularity: objectives and evaluation methods

- Some data management procedures *(deleting, replicating, moving between tapes, disks and caches)* are still carried out in a semi-manual mode.
- Reasonable replication and placement policies for the popular data is crucial for the efficiency of data processing.
- *Automated and dynamic replication* may help to improve tasks execution process.
- **Data popularity - how to measure?**
  - How often datasets were requested on the grid?
  - How many tasks were executed with the datasets as an input?

Sources of Data Popularity Meta-information

- Distributed Data Management System (DDM) Rucio
  - *Rucio Traces*
  - *Rucio API*
- EOS Report Logs
- PanDA Database (Production and Distributed Analysis)

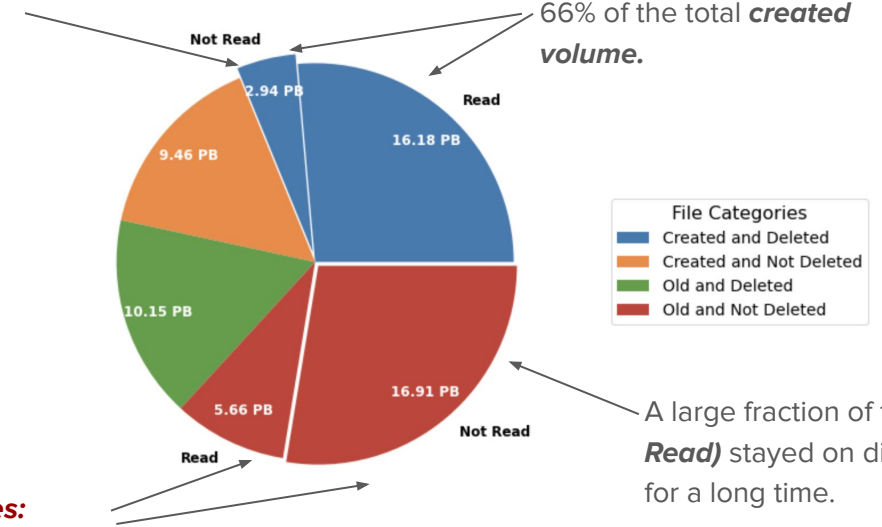
Analysis of ATLAS EOS instance at CERN using EOS Report Logs

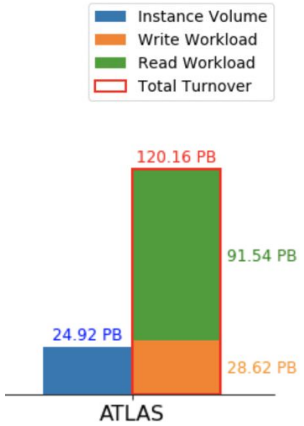
3 months period

- The total turnover ***(the sum of all bytes read and written)*** is over 480% of the instance volume.
- ATLAS read ~2-3 times more data than they wrote.

Almost ***3 PB*** of files *were not read* between their creation and deletion.

Short-lived files:
66% of the total ***created volume.***





A large fraction of files ***(Old, Not Deleted, Not Read)*** stayed on disk without being accessed for a long time.

Persistent files:
37% of the total volume.

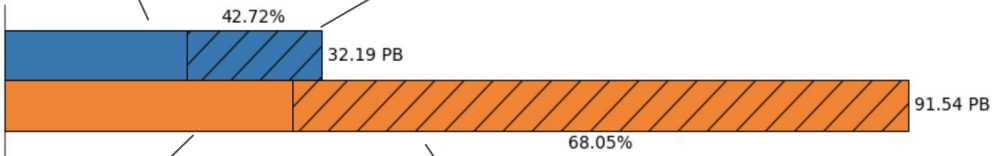
Analysis of ATLAS EOS instance at CERN using EOS Report Logs

3 months period

- *The larger volume (85%) is occupied by files with a size of >1 GB, while the average file size ~400 MB*
- *The average fraction of the file that is read per access at the ATLAS instance is 95%*
- *A **big fraction** of the read workload re-reads the same files*

Read Volume - ∑ sizes of all accessed files

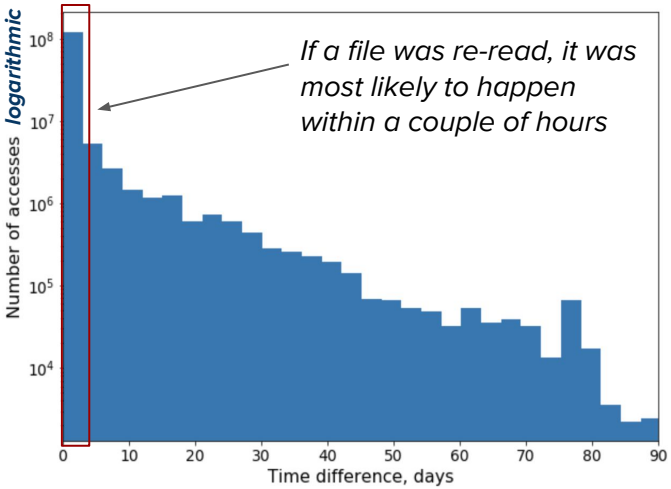
Repeated Read Volume - ∑ sizes of files accessed ***more than once***



Read Workload - ∑ all bytes read ***(including multiple reads)***

Repeated Read Workload - ∑ bytes read of files read ***more than once***

The distribution of the time difference between consecutive accesses



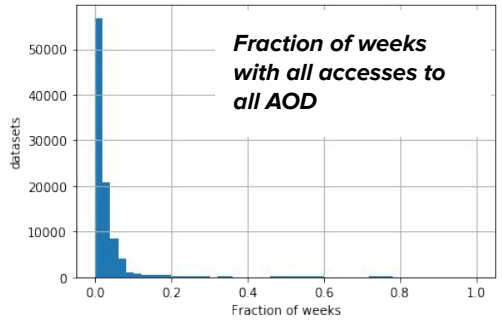
Rucio Access Metrics for 2020 year

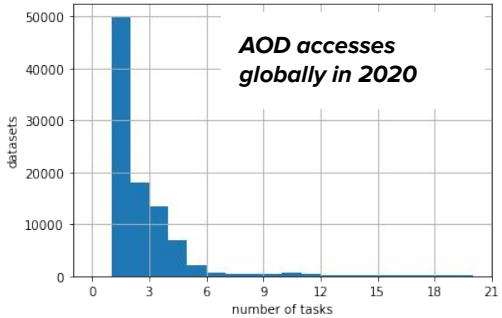
- Analyzed accesses to AOD, DAOD and HITS datasets in 2020 showed that ***site to site differences are relatively small***
- Analysis of how data accesses are spread over time or concentrated can have a big impact on data management
  - Very few ***AOD (Analysis Object Data)*** dataset accesses in a very small fraction of the year
  - The vast majority of ***AOD (Analysis Object Data)*** datasets are accessed only a few times during a year
- ***Analysis of disk storage efficiency***
  - Files spend a large fraction ***(~60% at CERN-PROD)*** of their time on disk without being accessed ➔ The time spent on disk by dataset replicas might be decreased

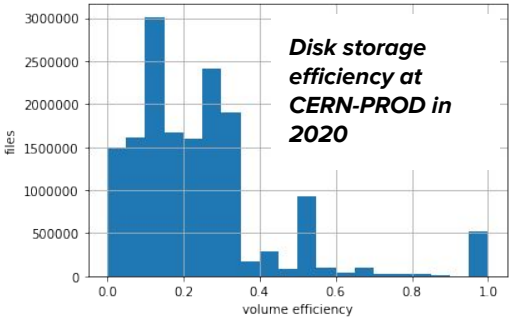
Volume efficiency

$V_{eff} = W_{acc} / W_{disk}$

W_{disk} = *weeks the file spends on disk*
W_{acc} = *weeks with accesses*

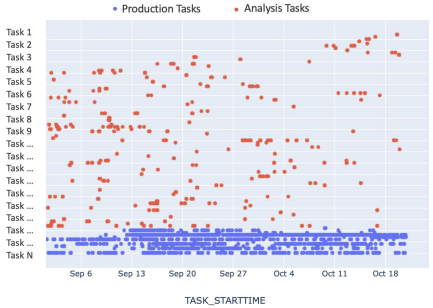






User Analysis Data Popularity Overview

Start time of user Analysis tasks fluctuates a lot, while Production tasks are executed centrally.



Popularity Metrics ***Data Processing Parameters***

Project	Data Format	Number of tasks	Number of jobs	Number of users	Number of Home Institutes	Number of Countries	Avg Task Execution Time (min)	Avg Task Waiting Time (min)	Number of Sites	Max Input Volume (TB)
mc16_13TeV	BPHY	1162	10999	8	7	7	2100.39	90.87	60	10.42
	EGAM	293	197516	19	13	10	4711.89	186.78	61	10.56
	EXOT	18138	603013	77	48	17	2782.47	263.95	70	9.74
	SUSY	125571	3097544	52	34	13	1545.41	258.69	70	2.61
	TAUP	411	7628	3	3	3	1864.00	56.01	52	2.18
	TCAL	45	4205	1	1	1	711.73	27.41	44	0.05
	TOPQ	20285	2826810	108	60	22	3290.79	591.58	71	26.59

*The most popular group of datasets **by the number of tasks***

*The most popular group of datasets **by the number of jobs and distribution of users interest***

The longest waiting time (~10 hours)

Max input volume

Conclusion and Future Plans

- It is now possible to understand the complete life cycle of data:
  - from the global distribution down to the fraction of files.
- The main outcome of the studies covering AOD and DAOD datasets:
  - DAOD datasets are the most popular for end-user analysis and access pattern is changing in time,
  - About 50% of the AOD data occupies disk space for extended periods without seeing active use,
  - AOD datasets are the most popular for derivation production.
    - It is conducted centrally and that's why ATLAS AOD handling policy was changed in 2020
    - There are no persistent AOD dataset replicas on disks, but there 1-2 copies on tapes
  - Most accesses take place within short intervals (couple of hours).