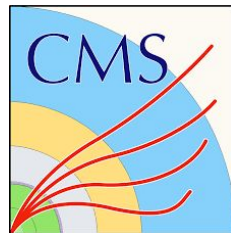
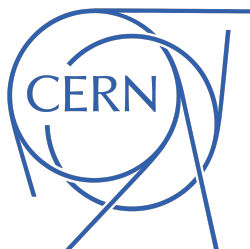


Recent Experience with the CMS Data Management System

Hasan Öztürk, Panos Paparrigopoulos,
Rahul Chauhan, Hugo Gonzalez Labrador, Dmytro Kovalskyi, Andrew Wightman,
Katy Ellis, Christos Emmanouil, Andres M. Ardila, Eric W. Vaandering



CHEP 2024

23.10.2024



- Recent CMS data management contributions
- CMS data management overview
- Challenges
- Technology upgrades & migrations
- Monitoring
- Summary

Recent CMS Data Management Contributions



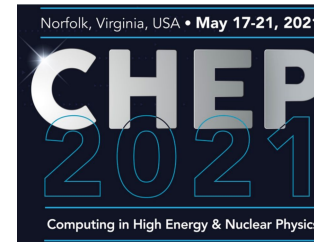
- **2019: Transitioning CMS to Rucio Data Management**

- Eric Vaandering, CHEP 2019



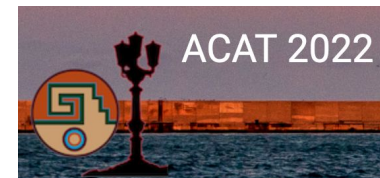
- **2021: Experience with Rucio in the wider HEP community**

- Martin Barisits et al., CHEP 2021



- **2022: Data Management interfaces for CMS experiment: building an improved user experience**

- Rahul Chauhan, ACAT 2022





CMS Data Management Overview



Rucio is the high level data orchestration service



FTS is used to TPC data between sites
(Mostly via Rucio)

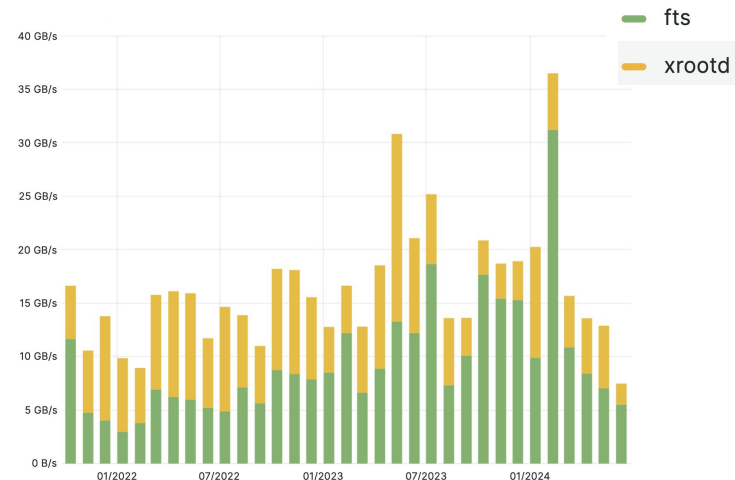


XRootD is mainly used to remotely
stream data to grid jobs



8 tape and 50+ disk storage elements

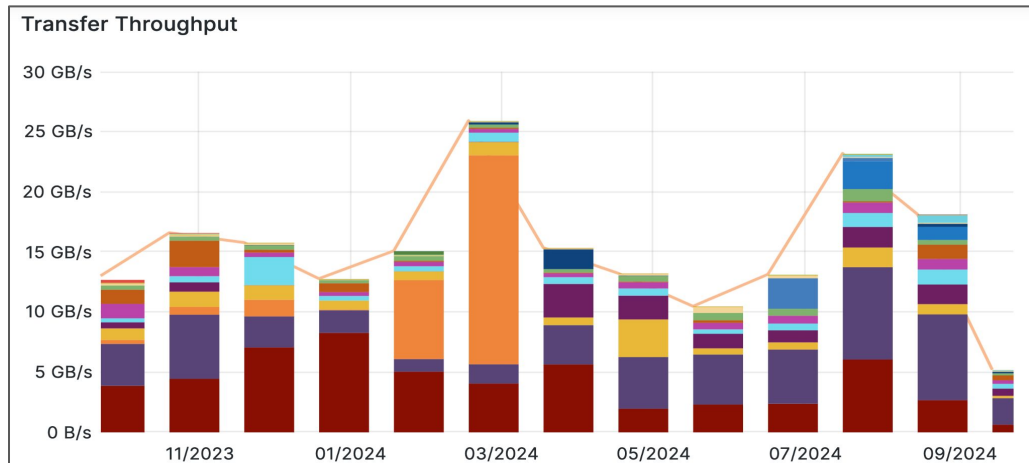
Transfer Throughput



- **Occupancy**

- **~250 PB** on disk (96% of total quota)
- **~550 PB** on tape (80% of total quota) – increase by 22PB last month

- **FTS Transfers: ~ 10-15 PB of FTS transfers weekly**



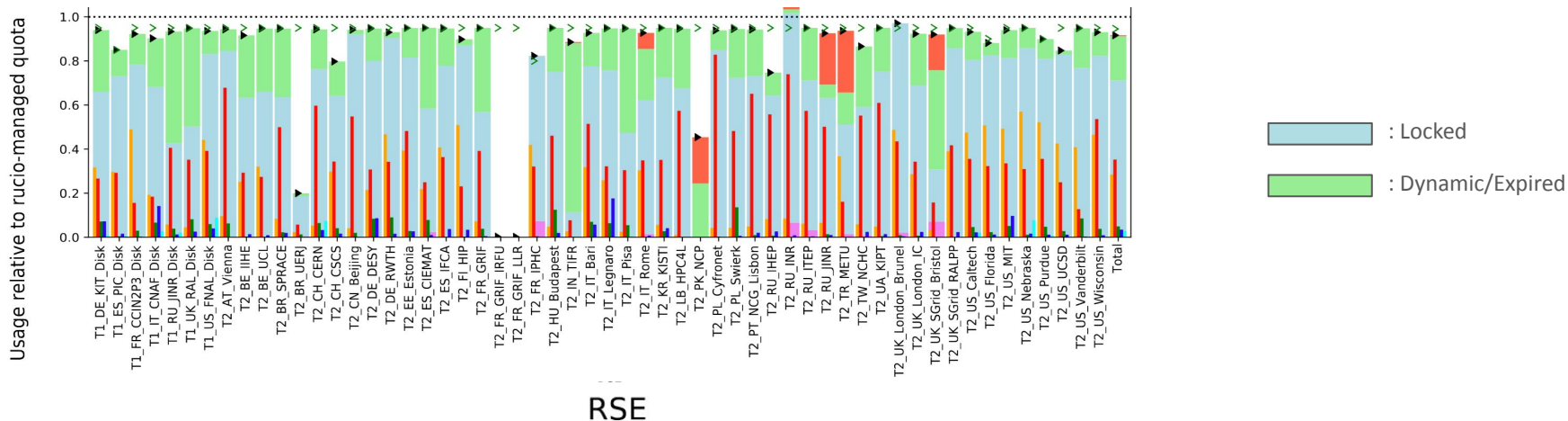
	avg
Total	15.2 GB/s
Production Output	4.19 GB/s
T0 Tape	3.79 GB/s
Data Challenge	2.03 GB/s
Analysis Input	1.04 GB/s
T0 Export	944 MB/s
User Subscriptions	740 MB/s
Data rebalancing	570 MB/s
Production Input	504 MB/s
ASO	420 MB/s
Data Rebalancing	262 MB/s
Manual Transfers	218 MB/s
Recovery	177 MB/s



Challenges

What to keep on disk

- **Disk is precious**
 - Constitutes **32%** of total storage (2023)
 - Certain disks such as CERN and FNAL are more popular due to their location and extra capabilities
- **Current disk policy: Keep multiple copies of analysis format data on disk permanently**
 - **2 copies of MiniAOD (26.6% of all disk usage), 6 copies of NanoAOD (2.7% of all disk usage)**
 - Remaining data is kept with lifetime, expired data (green parts in the figure below) is deleted based on LRU
- **Users contact the Ops team to stage data from tape**
 - Extra burden on the team as many of the requests are trivial and tedious to execute



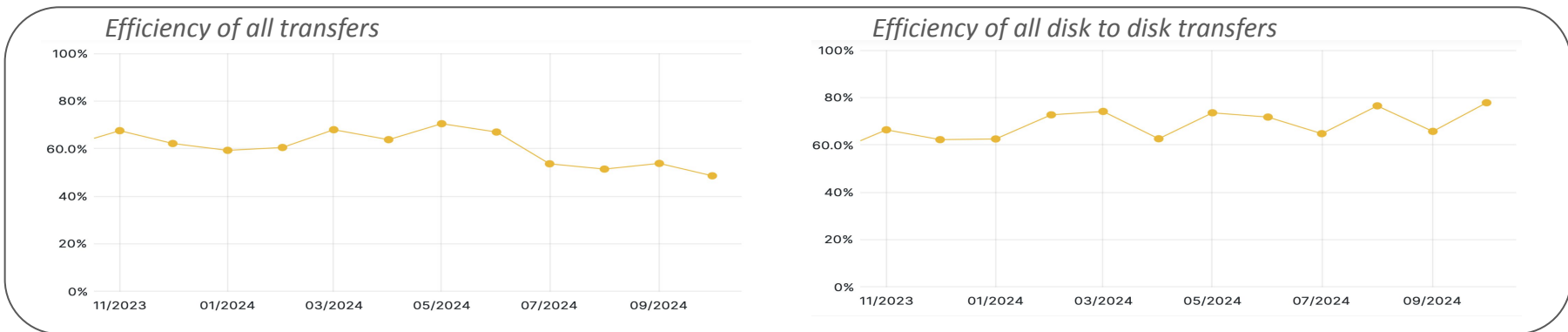
- **New feature: Users can lock any data on disk without any approval**
- **Changes made in Rucio so that each community can define their policy**
 - To be done via rucio policy package: [Link to the PR](#)
- **Protections:**
 - At most **10 PB** can be locked by this activity
 - Single user can lock at most **500TB** at all sites
 - Single user can lock at most **50 TB** at a single site/RSE
 - Data can be locked at most **1 year**
- **~1.5 PB of data is locked by this activity on average in the last 6 months**
 - Reduced user requests and the manual effort required by the team significantly



A happy physicist according to ChatGPT :)

- **Although the transfer efficiency isn't very high, often times retries work**

- Efficiency of all transfers: **58.8%** – Efficiency of all disk to disk transfers: **69.7%**
- ~ **150k** successful rucio rules in the last 30 days while only ~**300** failed rules (SUSPENDED)



- **Failures in tape transfers are critical as 100% success is required**

- **Top failure: Destination file exists and overwrite is not enabled (71% of all tape transfer failures)**
- Occurs in the retries of a failed transfer with which file is made it to the buffer or the tape
- As a policy, CMS doesn't allow the overwrite flag (e.g. `xrdcp -f`) in transfers to tape

- **Improvement: overwrite-when-only-on-disk**

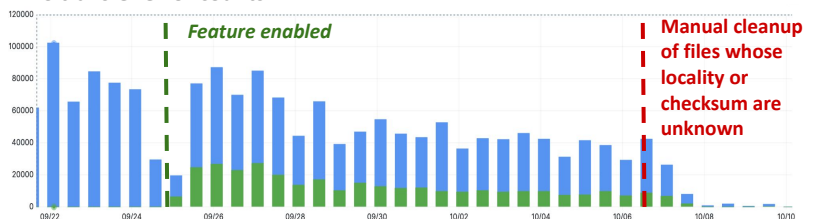
- Contribution provided by the CERN IT department for CMS (Code changes in FTS & Rucio) 🤝
- CMS provides this flag to FTS via Rucio to make FTS overwrite the files on tape buffer (disk)
- Files that are on tape are still handled manually

- **First results after enabling the feature for 2 tape storages (CERN CTA & CNAF STORM):**

- Error count didn't decrease significantly, but we have better monitoring thanks to enriched error messages
- **Remaining challenges:**
 - *Tape storage doesn't report the locality: BAD_ADDRESS (Green bars below)*
 - *Files on tape but either they are corrupted or SE doesn't report checksum: FILE_EXISTS (Blue bars below)*

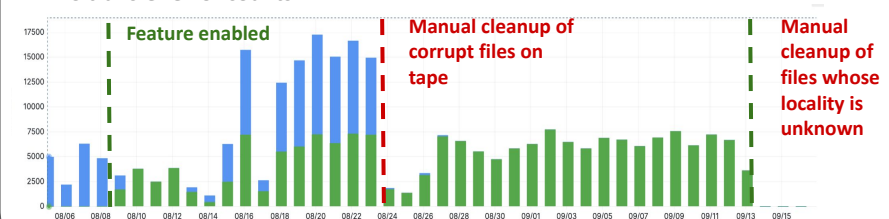
Transfers to CNAF Tape

FTS transfer error counts



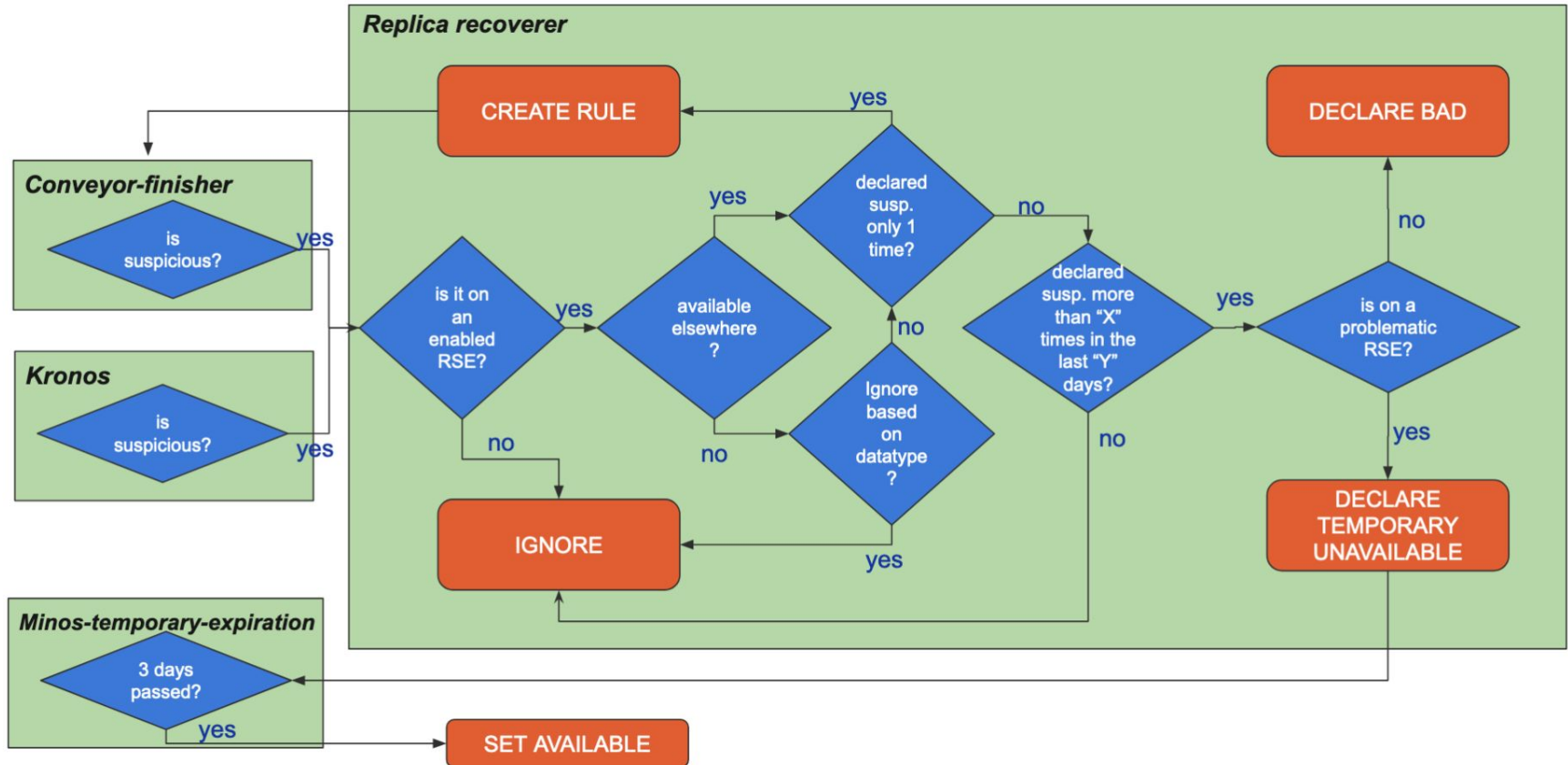
Transfers to CERN Tape

FTS transfer error counts



- **Corrupt replicas are a common problem**
 - Several incidents reported monthly
 - Operations team check the replica and replace it with a healthy replica (if any)
- **How to detect such replicas automatically?**
 - Rucio/FTS transfer error messages,
 - E.g. “CHECKSUM MISMATCH Source and destination checksums do not match”
 - Production & Analysis job error messages
- **Improvement: Adapted Rucio replica recoverer daemon**
 - Daemon is made VO-agnostic such that each community can use it now
 - Fixed issues: [1](#), [2](#), [3](#)

Suspicious replica recovery via Rucio



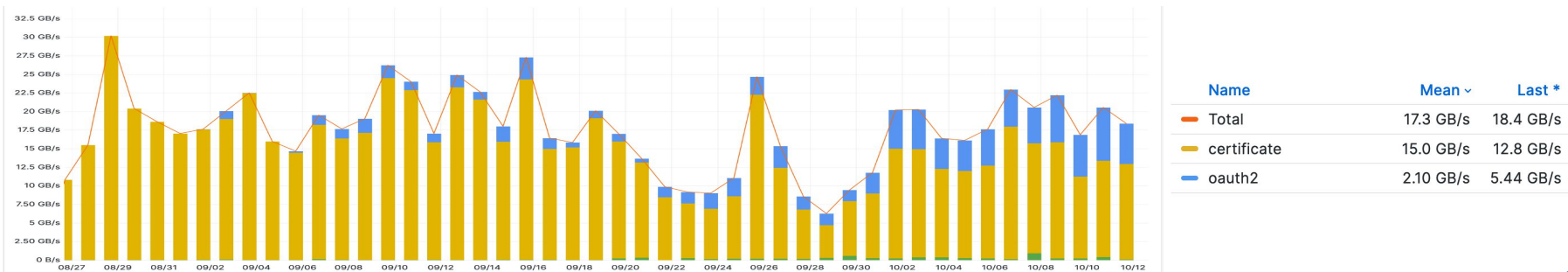


Technology upgrades & migrations

● Token transition

- CMS used the first version of token implementation in DC24 for around half the traffic
- Rolled back to certificates after DC24 to wait for the token improvements
- Gradually moving to tokens: Currently **~30%** of all CMS traffic is using tokens (**~55%** of all disk2disk traffic)
- Awaiting FTS development to use tokens for tape
- For more details, see [Rahul Chauhan's talk from the FTS/XrootD workshop](#)

Transfer throughput by auth method in the last 3 months



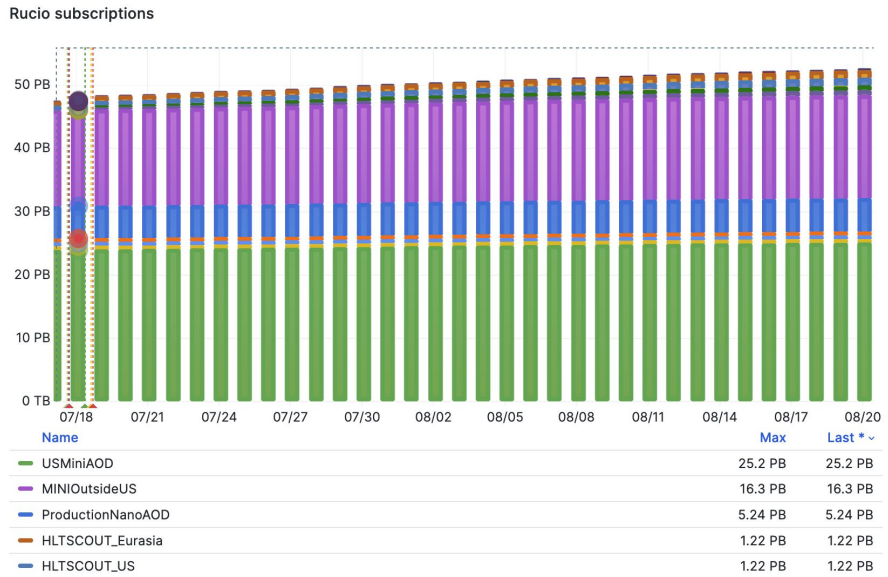
● Tape REST API transition

- All tape endpoints except FNAL has been migrated to HTTP (previously using SRM)
- This will allow us to communicate transfer metadata to tape systems which will optimise staging performance, prioritization etc.

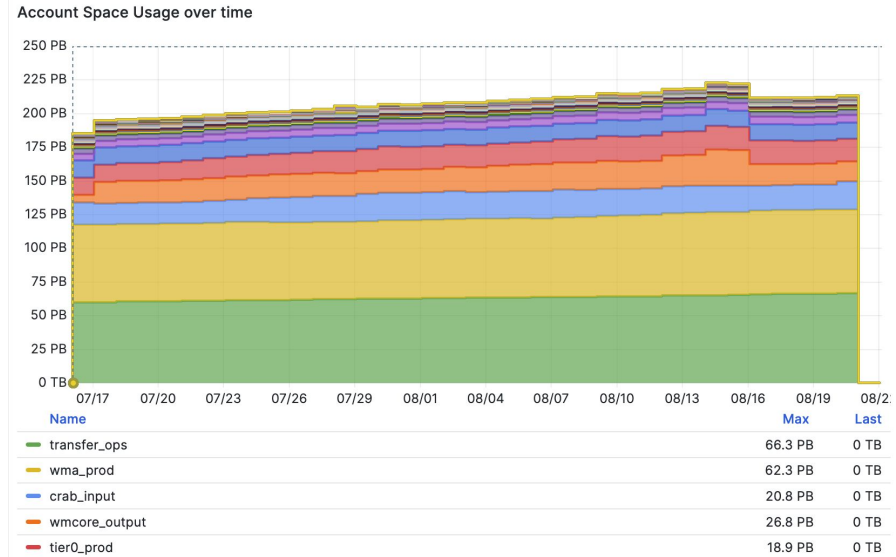


Monitoring

Monitoring improvements



Rucio subscriptions: How much data each rucio subscription locks over time

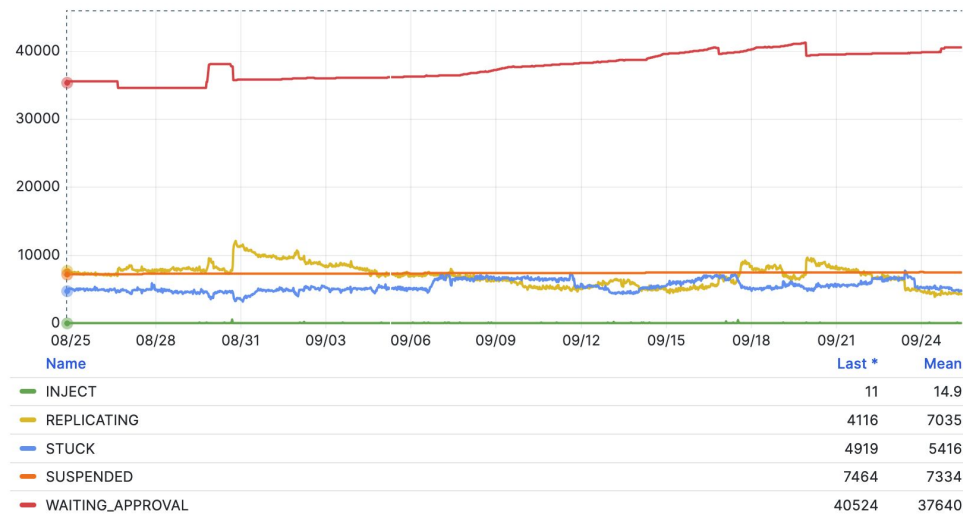


Account monitoring: How much data each rucio account locks over time

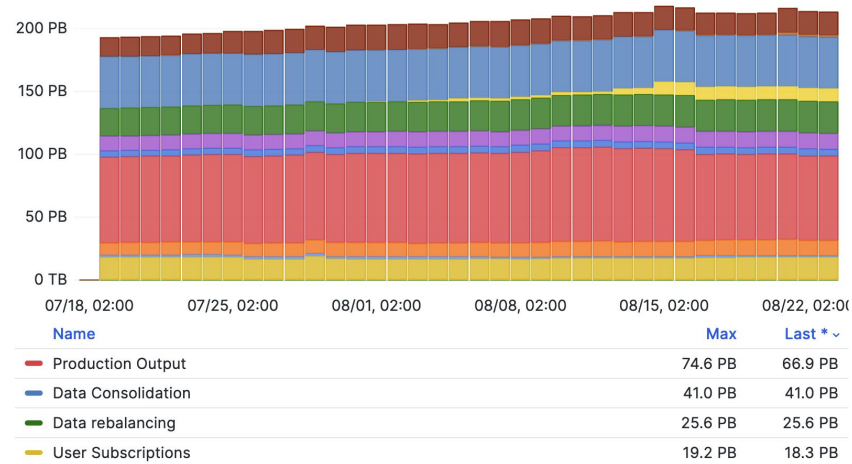
Monitoring improvements



Rule count overview (excluding OK rules)



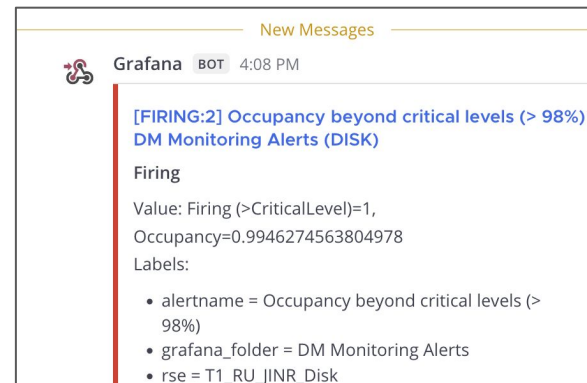
Activity locks over time



Rucio rules: Number of rules per state

Activity monitoring: How much data each rucio activity locks over time

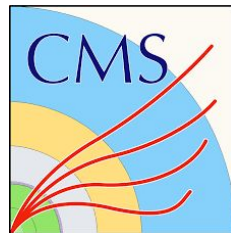
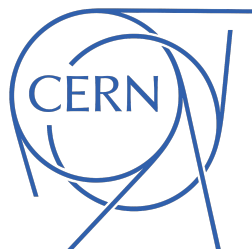
- Grafana alerts channelled to Mattermost
- Occupancy
 - Fire if occupancy goes beyond **98%** in a site
- Transfer failures
 - Fire if **5k+/hour FTS failures for 6 hours** for a given src or dst
- FTS submissions
 - Fire if FTS submissions failures are beyond 10% per RSE





- Letting users use disks more freely to increase data availability (User AutoApprove)
- Improvements to minimize tape transfers failures & corrupt replicas
- On track in token and tape REST API transitions
- Better account, activity, subscription & rules monitoring

THANKS!





- **New error modes**

- Tape storage doesn't report the locality (BAD_ADDRESS)
 - DESTINATION [14] Could not check destination file locality (overwrite-when-only-on-disk requested)
- Files on tape but either they are corrupted or SE doesn't report locality (FILE_EXISTS)
 - DESTINATION [17] Destination file exists and is on tape (overwrite-when-only-on-disk requested)



- **Synchronization of Rucio with DBS (Meta-data handling system)**
 - CMS uses a separate service to handle its metadata called DBS
 - Data injected into Rucio & DBS synchronously, but modifications in one data source cause inconsistency e.g. valid files (DBS) that have no replica (Rucio)
 - **Improvement: An internal sync tool is developed**
- **Data rebalancing: Popular sites getting over-occupied**
 - Currently manually rebalancing data when needed
 - Looking into different solutions such as **Rucio BB8 rebalancing daemon**
- **Overloading storages & throttling**
 - Large transfer requests targeting a single site might result in storage failures
 - **FTS Optimizer** observed to be slow to react
 - Currently using **Rucio throttler daemon**, but this requires manually setting limits per RSE