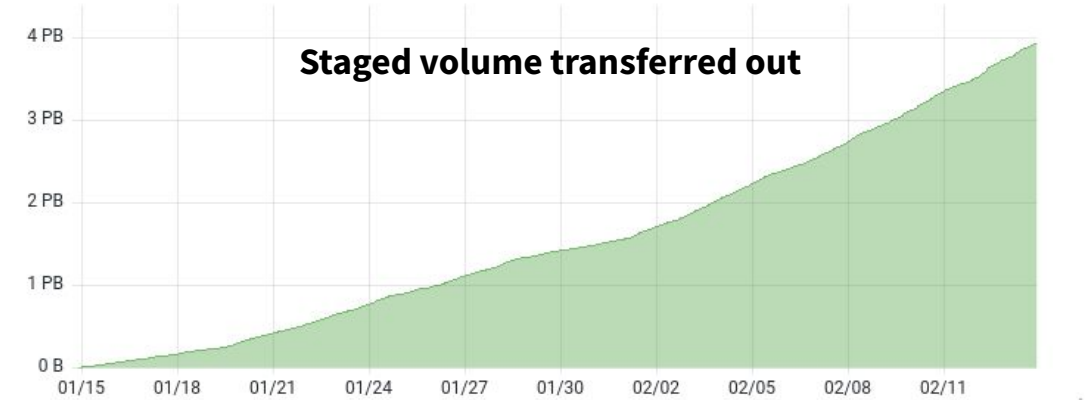
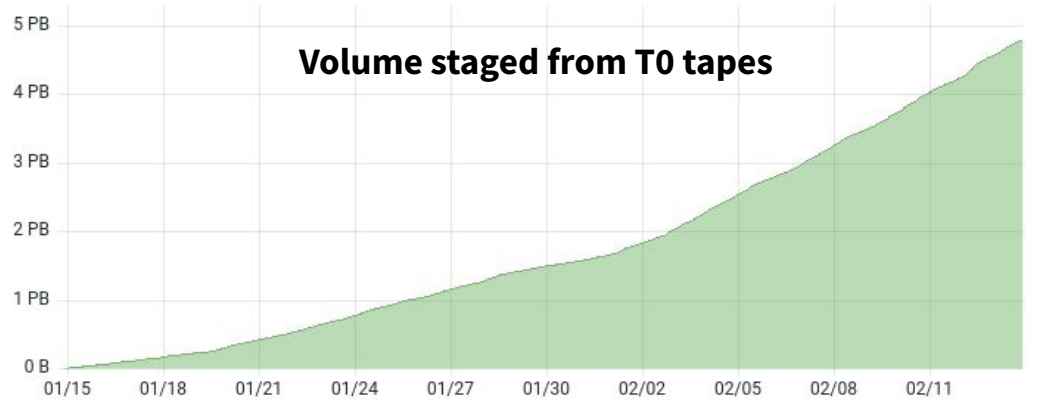


CMS DM - T0 Storage meeting: CTA

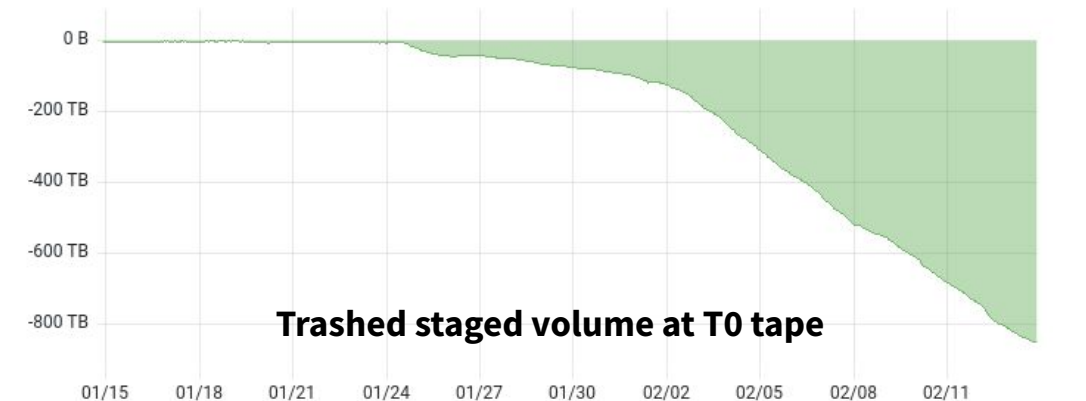
CTA team

14/02/23

Parking staging from T0 tape



- Parking recall trashing now represents 24% of staged data
 - trashing is increasing over time
 - out of 30 reading tape drives 7 are working for the bin
- *Bring online timeout* set to 7 days needs to be increased
 - defaults to 30 days?
- Other staging activity?
 - parking staging consume all tape drives leaving no bandwidth for other *activities*
 - **Currently staging priority can be set by activity**
 - parking staging priority could be lowered
 - CTA team needs activity string



Beginning of Run3: *legacy* CMS placement tweaks

- CTA maps tape family with directory
 - *CASTOR legacy*
- Improving written data placement with the experiments
 - Improve per directory tape collocation on tape
 - CMS split of MC, 2022 data, 2023 data...
 - parking data is interleaved with 2022 data
- Several limitations as
 - CTA directory structures is dictated by experiment namespace
 - remapping in CTA tape buffer

Tape site need hints: *archive metadata* discussions are starting with experiments

Current CMS interest:

- tag parking data for specific tape family routing
- *archive_priority*: lower priority of MC data over DAQ

Files on broken tapes at T0

- **Files on broken tape imported from CASTOR in CTA 2 years ago**
 - Communication hell and too many people involved: <https://github.com/rucio/rucio/issues/4447>
- **Rucio workflows and T0 assumed files would be deleted from tape by DM team**
 - or experiment should not reuse site/filepath
- **Technical proposal to end all this: fix wrong assumption**
 - T0 tape will explicitly flag files on broken tape to prevent these from being revalidated
 - **file size will be set to 0**
 - **other metadata unchanged**
 - **Files on broken tapes will be published (along with all files) in weekly experiment namespace dumps**
 - **example: file on broken tape since 1/1/2018**

```
{"adler32":"8ac7845f","error_text":"","on_tape":true,"path":"/eos/ctacms/archive/some_directory/some_file.txt",  
"size":0,"tape_status":"broken","tape_status_since":1514761200}
```

- **CMS developments on RUCIO Namespace / CTA Namespace comparison?**

Clean up datachallenge data on T0

- **> 2 PB of datachallenge test data in eosctacms.cern.ch/eos/ctacms/archivetest**
 - Pure test RSE that must be separated from production in RUCIO CMS
 - Site should be allowed to regularly delete from this location (data older than X months?) if they need to recover space/tapes
 - How to synchronize deleted files? directory trees?
 - Could site be allowed to issue Rucio REST API queries against CMS Rucio instance?