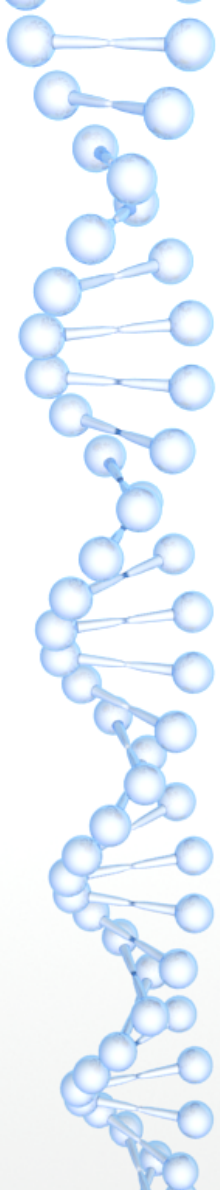


Data management in jAliEN - reclaiming storage space

Volodymyr Yurchenko



Storage operations



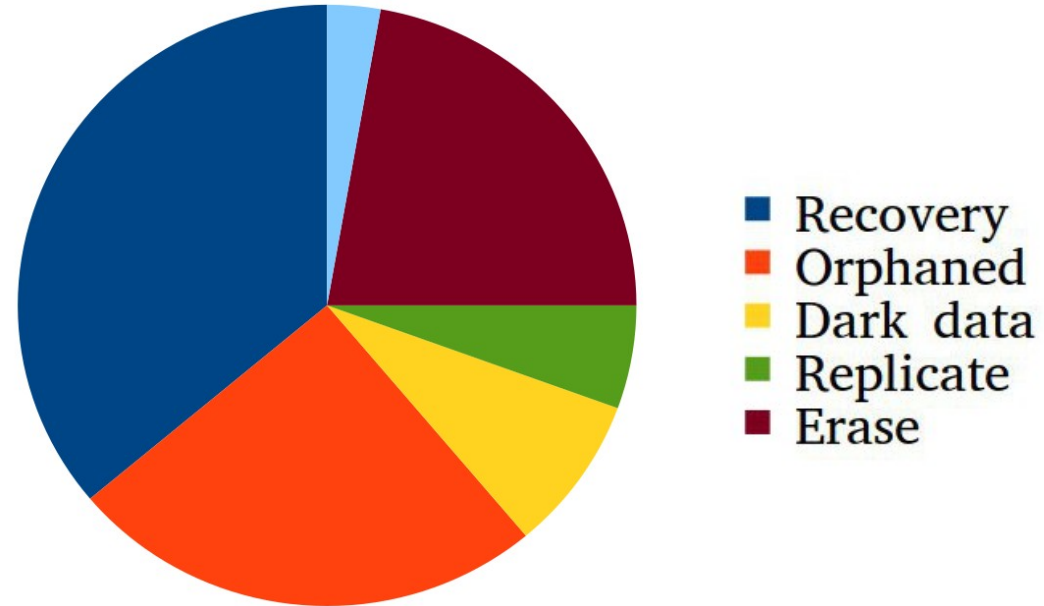
Typical operations

- Recover lost files on a storage
- Clean up orphaned entries in the Catalogue
- Find dark data and sync SE
- Replicate/move SE content
- Storage decommission

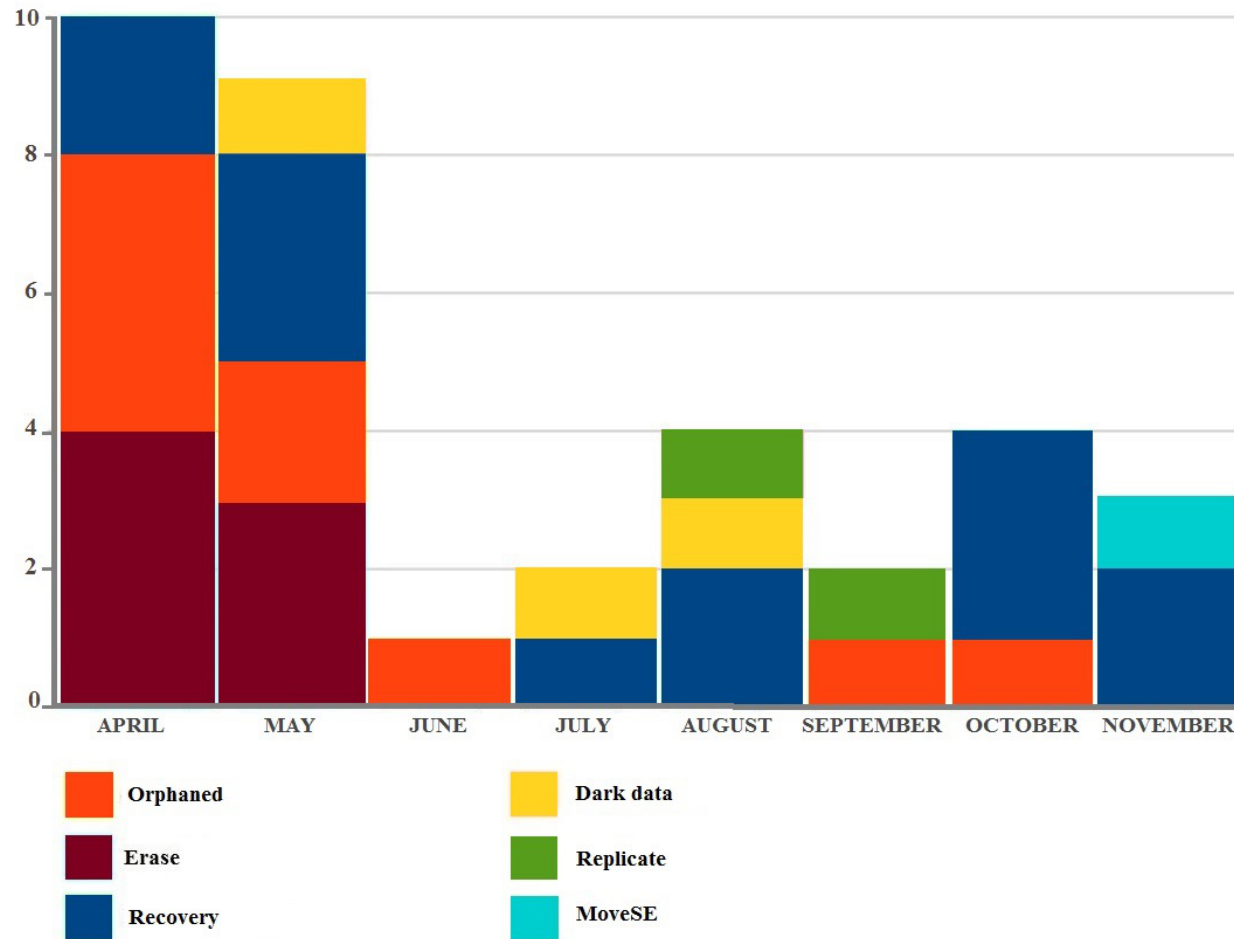
Site admins report the incident and prepare a list of lost files or files present on the storage.

Operations Apr2018-Nov2018

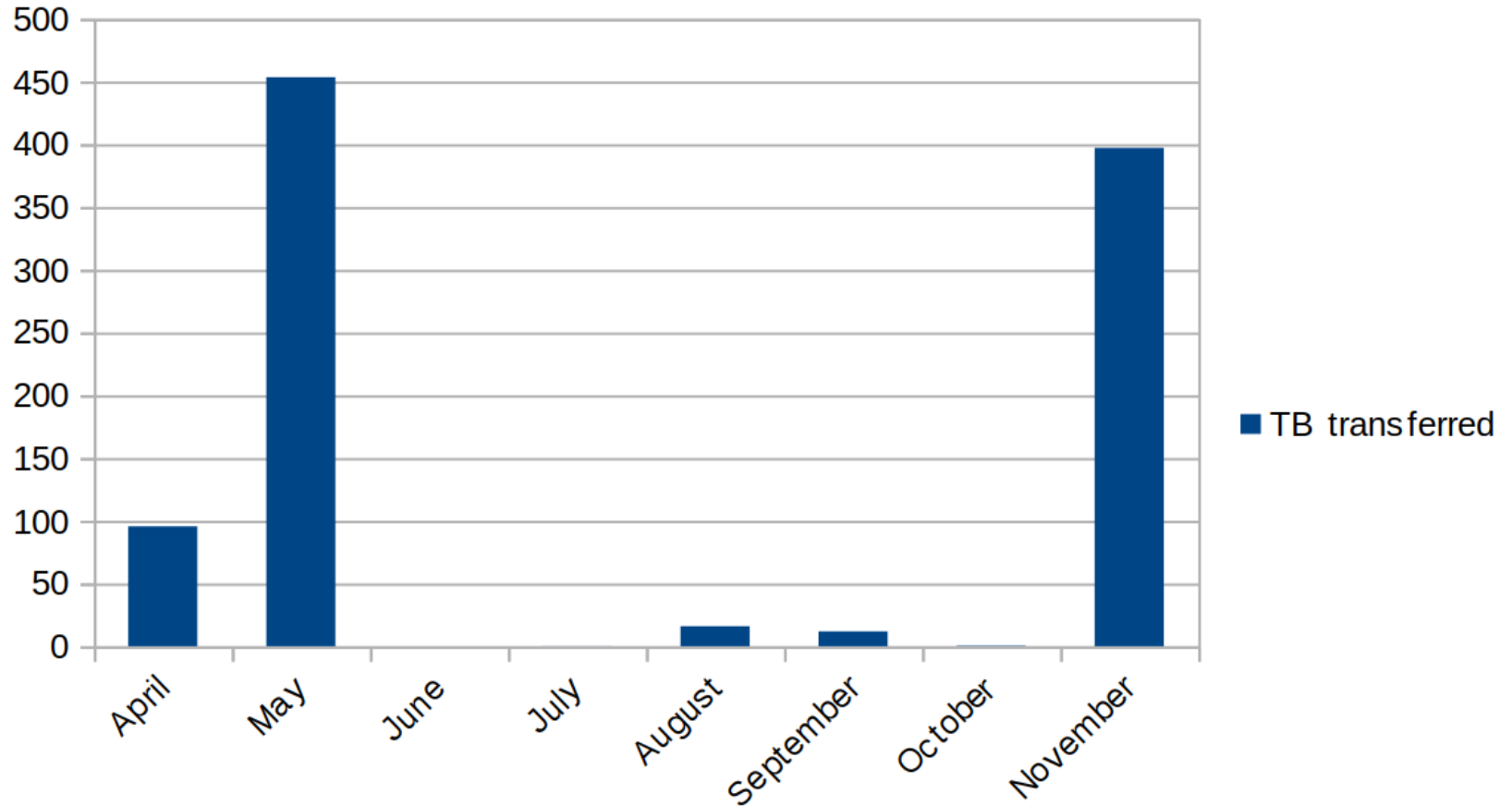
- Total: **36**
 - Recovery of lost files: **13**
 - Clean up of zero-replicas files: **9**
 - Dark data & sync: **3**
 - Replicate/move SE content: **3**
 - Erase storage: **8**



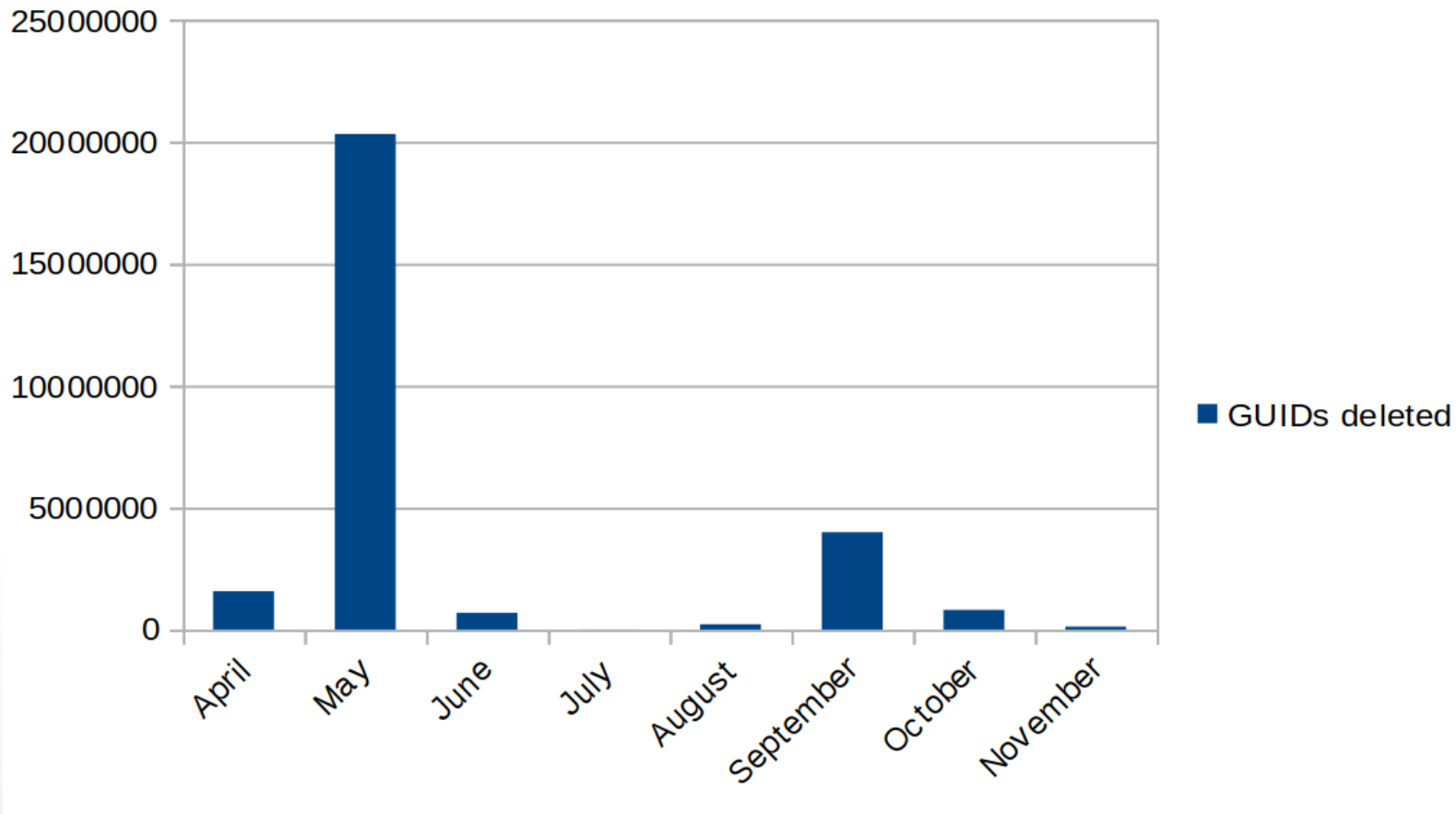
Operations rate by months



Transfers (SE Recovery & Copy) by month



Orphaned GUIDs clean up by month





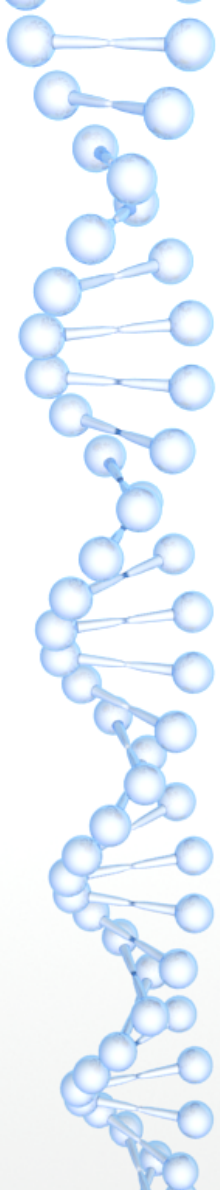
Time Estimations

- 100 000 files are scheduled for transfer in ~**30 min**
- 100 000 orphaned entries cleaned up in ~**14-15 hours**
- 100 000 files (4.7 TB) transferred in **12 days** – *heavily depends on the target storage network throughput, xrd3cp enabled, internal configuration, etc*



Decommission (Erase) SE

- ALICE::CyberSar_Cagliari::SE
- ALICE::ISMA::SE
- ALICE::Trujillo::SE
- ALICE::Tsukuba::SE
- ALICE::UPB::SE
- ALICE::Strasbourg_IRES::SE
- ALICE::SARFTI::SE
- ALICE::Birmingham::SE



TrackRefs and AliESDfriends Removal



TrackRefs: Theory (Initial plan)

- Download archive
- Extract
- Create new archive
- Upload
- Delete old archive members and register new
- Delete old archive and rename the new one



TrackRefs: Practice (Slightly more complicated)

- Clean up previous iterations
 - Decide whether previously uploaded files are valid
 - If yes, move them to the target directory
- Check if PFNs are accessible
- If the file is the only member of archive or is not a member of any archive – just delete it
- Download archive
- Extract
- Create new archive
- Upload it to the temporary directory
 - If failed – delete temp dir
- Register new members in the temporary directory
 - Create subdirectory structure (e.g., BKG/ subdir for some productions)
 - If failed – delete temp dir
- Delete old archive members
- Delete old archive
- Create a file marker to leave traces
- Move the uploaded archive
 - Check if there is no old archive left (errors during deletion)
- Move the archive members
 - Check if there is no old member left (errors during deletion)
- Delete the temporary directory



TrackRefs: jAliEn in CVMFS

- **8** tags in total (to cover more and more use cases)
- Now available as `VO_ALICE@JAliEn::1.0.7-1`



TrackRefs: Removal Statistics

- Total number of found files: **83 446 942**
- Total number of processed: **4 017 163 (4.8%)**

- Total size of found files: **5048 TB**
- So far reclaimed space: **167 TB (3.3%)**
- Failed to reclaim: **76 TB (1.5%)**
- Fail ratio: 31%

Lessons learned, removal will continue!



TrackRefs: What is success?

- A successful job has processed all 100 files
- Failure usually means at least one failed file, and it fails on retries later on
- Therefore the file processing success ratio is always higher than job success number


TrackRefs: Newer productions

| Output directory | % | Total | Done | Active | Wait | Err. | Oth. | Running | Saving | Size |
|--|------|-------|------|--------|------|------|------|-----------|---------|----------|
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18a7c/0 | 97% | 383 | 372 | | | 11 | | 8d 17:41 | 6:29 | 86.88 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18e1c/0 | 84% | 147 | 124 | | | 23 | | 6d 8:54 | 2:37 | 49.53 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18e1b/0 | 83% | 278 | 232 | | | 46 | | 8d 14:44 | 4:09 | 69.74 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18e1a/0 | 62% | 133 | 83 | | | 50 | | 7d 1:03 | 1:34 | 15.89 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18e1/1 | 95% | 346 | 331 | | | 15 | | 2d 14:07 | 6:56 | 6.04 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18e1/0 | 57% | 518 | 299 | | | 219 | | 18d 4:04 | 7:28 | 49.32 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18d7_extra/0 | 95% | 145 | 138 | | | 7 | | 5d 23:52 | 2:22 | 720 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18d9b/0 | 100% | 10 | 10 | | | | | 44m 20s | 10m 54s | 941.9 KB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18d9/0 | 84% | 13 | 11 | | | 2 | | 20m 32s | 10m 50s | 655 KB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18d8/0 | 95% | 41 | 39 | | | 2 | | 17:17 | 40m 33s | 12.33 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18d7/0 | 96% | 150 | 145 | | | 5 | | 1d 2:25 | 2:24 | 39.47 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18d6a/0 | 83% | 509 | 424 | | | 85 | | 50d 11:21 | 9:04 | 153.2 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC18b10b_calofast/0 | 100% | 1 | 1 | | | | | 30s | 53s | 9.272 KB |



TrackRefs: Older Productions

| | | | | | | | | | |
|--|------|-----|-----|--|--|-----|-----------|---------|-----------------|
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC09a9/0 | 99% | 348 | 347 | | | 1 | 14d 10:15 | 9:20 | 29.04 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC08d3/0 | 0% | 14 | | | | 14 | | | |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC08-LHC08c11/0 | 0% | 10 | | | | 10 | | | |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC08-LHC08c5/1 | 0% | 11 | | | | 11 | | | |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC08-LHC08c5/0 | 11% | 9 | 1 | | | 8 | 26s | 1m 44s | 92.11 KB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC08-LHC08c4/1 | 8% | 12 | 1 | | | 11 | 1m 5s | 1m 24s | 233.2 KB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC08-LHC08c4/0 | 0% | 11 | | | | 11 | | | |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC08-LHC08u/19 | 66% | 45 | 30 | | | 14 | 1 | 18:25 | 37m 40s 2.56 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC08-LHC08u/3 | 95% | 42 | 40 | | | 2 | 5:14 | 41m 14s | 2.746 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC07-LHC07g/123 | 0% | 40 | | | | 40 | | | |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC07-LHC07g/105 | 0% | 44 | | | | 44 | | | |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC07-LHC07g/48 | 0% | 9 | | | | 9 | | | |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC07-LHC07g/25 | 0% | 9 | | | | 9 | | | |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC07-LHC07g/68 | 0% | 9 | | | | 9 | | | |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC07-LHC07f/45 | 0% | 1 | | | | 1 | | | |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC07-LHC07e/24 | 100% | 9 | 9 | | | | 6:40 | 9m 14s | 1.437 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/PDC07-LHC07c/0 | 42% | 501 | 211 | | | 290 | 1d 3:06 | 3:29 | 112.6 MB |
| /alice/cern.ch/user/a/aliproduct/TrackRefs/output/LHC17j3b2/0 | 76% | 21 | 16 | | | 5 | 1:29 | 16m 56s | 4.525 MB |



TrackRefs: Job Example

| % | Total | Done | Active | Wait | Err. | Oth. | Running | Saving | Size |
|-----|-------|------|--------|------|------|------|----------|----------|----------|
| 73% | 2584 | 1901 | | | 671 | 12 | 8d 15:14 | 1d 21:49 | 109.4 MB |

- Production: **LHC11a10b** (chunk #4 out of 17)
- Total number of files: **45000**
- Successfully processed: **41419 (92%)**
 - Not data availability, but an execution success rate!
- Took **14:43 hours** of the real time to run



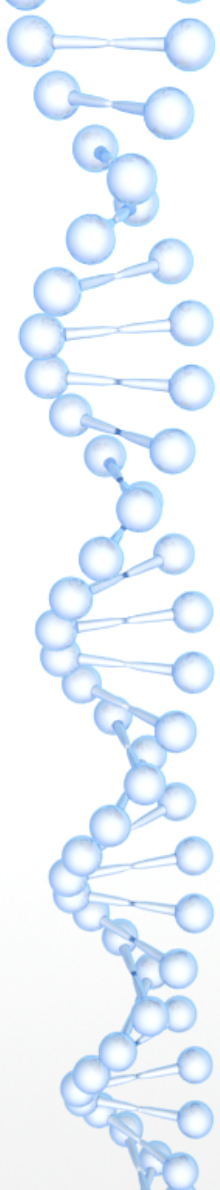
TrackRefs: Failure Reasons

- Failed to download the file
- Connection to JCentral rejected (services scalability issues)
- Job is killed by a WN for a random reason
- `ERROR_E, EXPIRED`



TrackRefs: Implementation and Run Problems

- Job can be killed at any time → complex error handling with all kinds of failsafe procedures
 - Sometimes revealing accidental DB defects (e.g. wrong value type for particular table)
 - Sometimes showing jAliEn bugs and unaccounted cases
- Compression bug
 - „Zero compression“ in jAliEn was not equal to „no compression“ - had to change the flags
 - And to re-do **153 TB** of processed files
- Database overload with a lot of queries
 - Many lookups for archive members, renaming, moving, deleting etc, required by the main procedure and failsafe mechanisms
- We had to limit the number of such jobs running concurrently because of the high load they put on the DB



MC productions deletion and archival



Reclamation Statistics

Production tags deleted

- Total number: **28**
- Space reclaimed: **957.038 TB**

Production tags archived

- Total number: **32**
- Space reclaimed: **188.012 TB**



Conclusions

- During 8 months (April 2018 – November 2018) there were:
 - Reclaimed: **1312,05 TB** of disk storage
 - Recovered: **626.55 TB** of lost files
 - **27 697 391** orphaned GUIDs (and LFNs) deleted
- The TrackRefs deletion was started and will be continued next year