



# Repack, where we are

CTA DEPLOYMENT MEETING – THU 25/06/2019

# Summary

1. Functionality that have been implemented
  1. Repack "Just move" mode
2. Tests
  1. Tests that needed to be done
  2. Tests that have been done
  3. Bugs discovered
3. **Current development**
  1. Investigating the explosion of the repack buffer
4. What remains to be implemented
  1. Bug correction
  2. Functionalities
  3. Timelines

# 1. Functionality that have been implemented

1. Repack "just move" mode
  - ▶ Goal : Move the content out of a tape onto one or several other tapes
  - ▶ Implemented that functionality
  - ▶ Implemented unit tests
  - ▶ Implemented system tests (CI)

# 2. Tests

1. Tests that needed to be done
  - ▶ Unit tests of the complete repack "just move" workflow
  - ▶ Stress test of Repack
  - ▶ Repack a single tape (10TB -> 15TB)
  - ▶ Repack all tapes of 10TB -> 15 TB (Vlado)

# 2. Tests

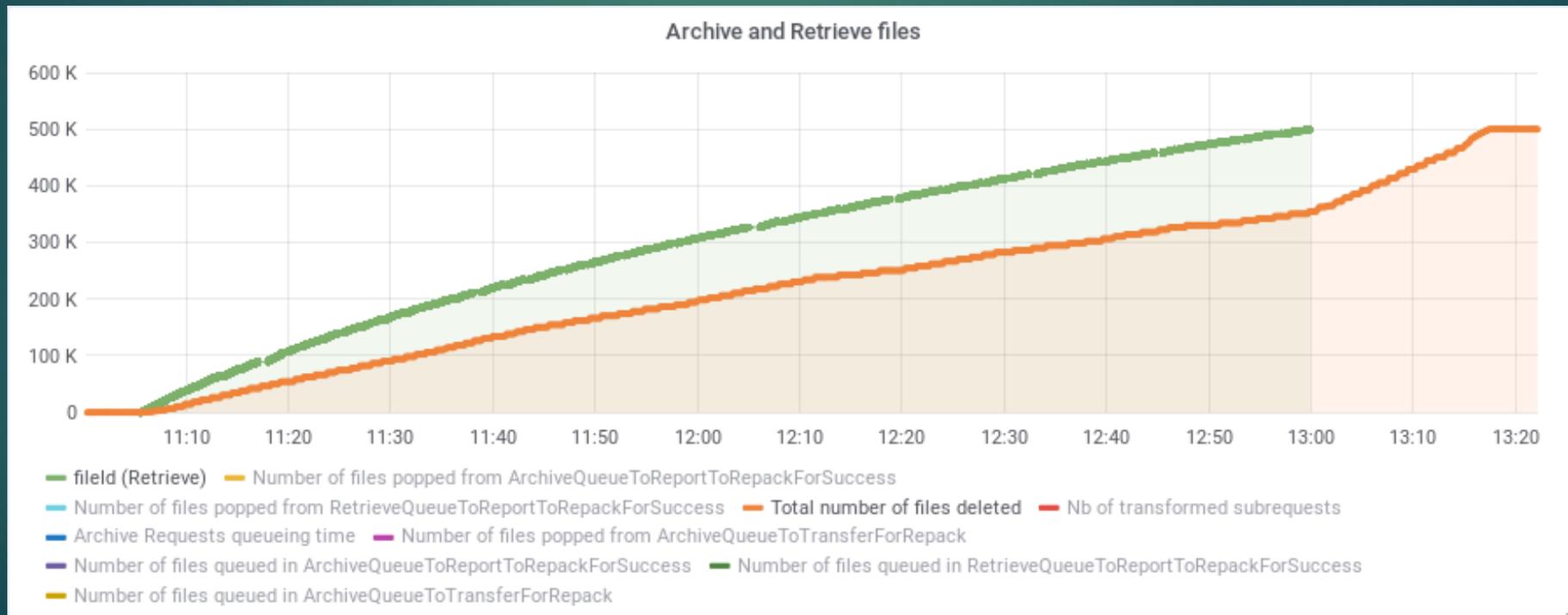
## 1. Tests that have been done

- ▶ Unit tests of the complete repack "just move" workflow ✓
- ▶ System test (CI)
  - ▶ Ping pong Repack --> Round trip repack with one file ✓
  - ▶ Repack of a tape containing 1153 files ✓
  - ▶ CI-Stress test ✗
    - ▶ Repack of a tape containing 500 000 files
    - ▶ Explosion of the repack buffer (Issue link : <https://gitlab.cern.ch/cta/CTA/issues/531>)

# 2. Tests

## 1. Bugs discovered

### 1. Explosion of the Repack Buffer



# 2. Tests

## 1. Bugs discovered

1. Explosion of the Repack Buffer
2. Repack is blocked if a Retrieve Request fails (no tape available found)
  - ▶ Exception "Unexpected job status" : To be corrected
3. cta-admin sq
  - ▶ In QueueAndMountSummary::getOrCreateEntry(): Unexpected mount type.

# 3. Current development

1. Investigating the explosion of the Repack's buffer
  - ▶ Added some monitoring during the repack workflow (size of queues, time to do the reporting, ...)
    - ▶ Link <https://meter-cta.web.cern.ch/d/JDpWes7Wz/repack?orgId=1&from=1561456823000&to=1561464955000>
2. Corrected some possible points of slowness
  - ▶ Long time to enqueue Archive Request : corrected
  - ▶ Metadata of the Repack Request processing lag
    - ▶ Workaround : Multithread the reporting of Repack subrequests : Improved the explosion : max 80K files in the buffer for a 500K files repack
    - ▶ Shard the Repack Request : Not implemented

# 3. What remains to be implemented

## 1. Bug correction

- ▶ Explosion of the Repack buffer
  - ▶ Backstop is coming, not a problem anymore ?
- ▶ Repack is blocked if Retrieve Request fails
- ▶ cta-admin sq command

# 3. What remains to be implemented

## 2. Functionalities

- ▶ Repacking of a disabled tape
  - ▶ Linked to the dedication, overriding of the "disabled" flag not supported
- ▶ Put the repack buffer URL parameter in a configuration file
- ▶ Repack tape-repair workflow
  - ▶ If files are in the buffer, we need to directly create the Archive Requests and not try to retrieve them from the tape
- ▶ Repack "just add copies" workflow
  - ▶ In case of a storage class modification, create new copies of a file

# 3. What remains to be implemented

## 3. Timelines

Functionality	Estimated time	Estimated date
1. Buffer explosion	???	
2. Repack is blocked if Retrieve fails	1 week	
3. cta-admin sq command bug	2 days	
4. Repacking of a disabled tape	1 week	
5. Repack buffer URL in config file	1 day	
6. Repack tape-repair	2 weeks	
7. Repack "just add copies	2 weeks	
8. Testing Repack in a single tape (10TB --> 15 TB)	0.5 day	
9. Testing Repack multiple tape	0.5 day	

Thanks for your attention