The GridKa Tape Storage: various performance test results and current improvements

Track 4: Data Organisation, Management and Access

Author/co-author(s): H. Musheghyan, A. Petzold, A. Heiss, D. Ressmann, M. Beitzinger

After the first ATLAS recall test within ATLAS Distributed Computing (ADC) tape carousel in 2018, various local tests were performed on GridKa tape system.

The main goal of these tests is

- a detailed study of the current tape setup,
- identifying and eliminating bottlenecks,
- properly adjusting and improving the current setup.

Increase the overall performance of the tape storage system

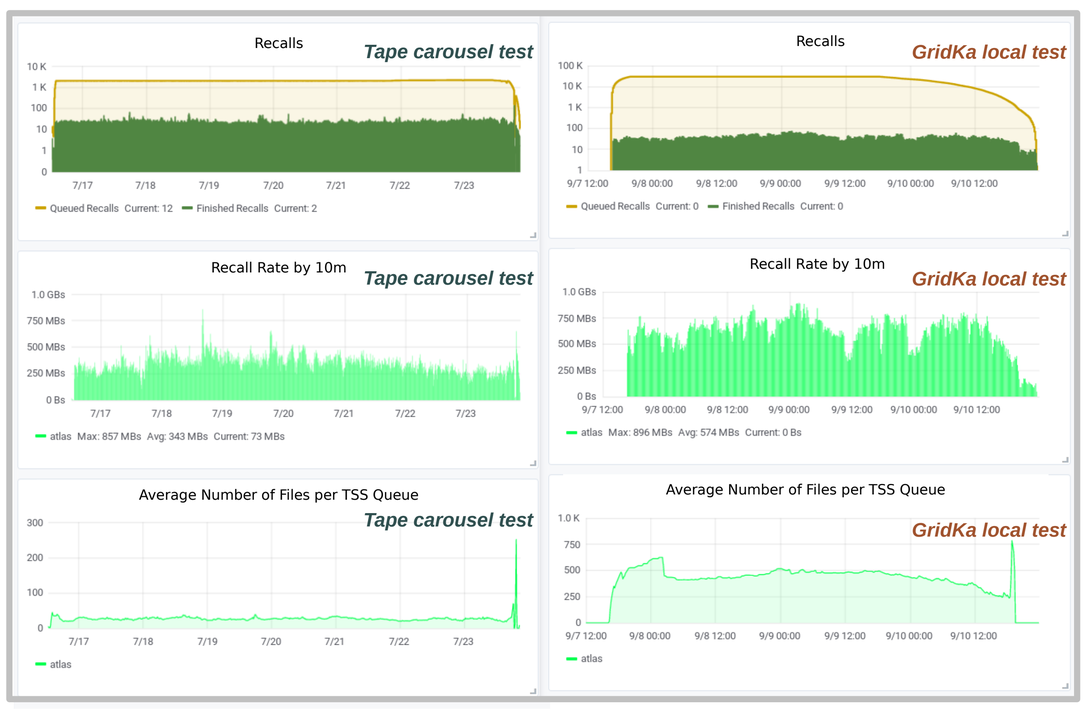
Read files directly and indirectly from tape

Tape Carousel Test Results

- Nr. of files: ~92,000
- Queued requests: 2,000
- Avg read rate: ~400 MB/s
- Nr. of drives: 8

GridKa Local Test Results

- Nr. of files: ~92,000
- Queued requests: 30,000
- Avg read rate: ~600 MB/s
- Nr. of drives: 8

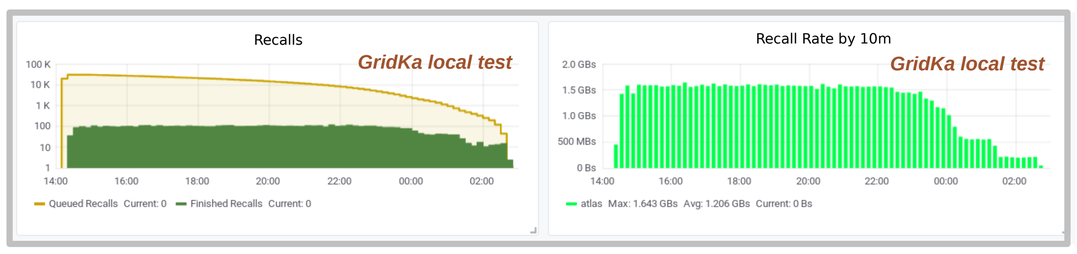


More details in the following article: Ressmann, Doris, et al. "The GridKa Tape System: status and outlook." EPJ Web of Conferences. Vol. 214. EDP Sciences, 2019.

Read the entire tape (12 tapes)

GridKa Local Test Results

- Nr. of files: 31,411
- Queued requests: 30,000
- Avg read rate: ~1600 MB/s
- Nr. of drives: 12



Read/write large files

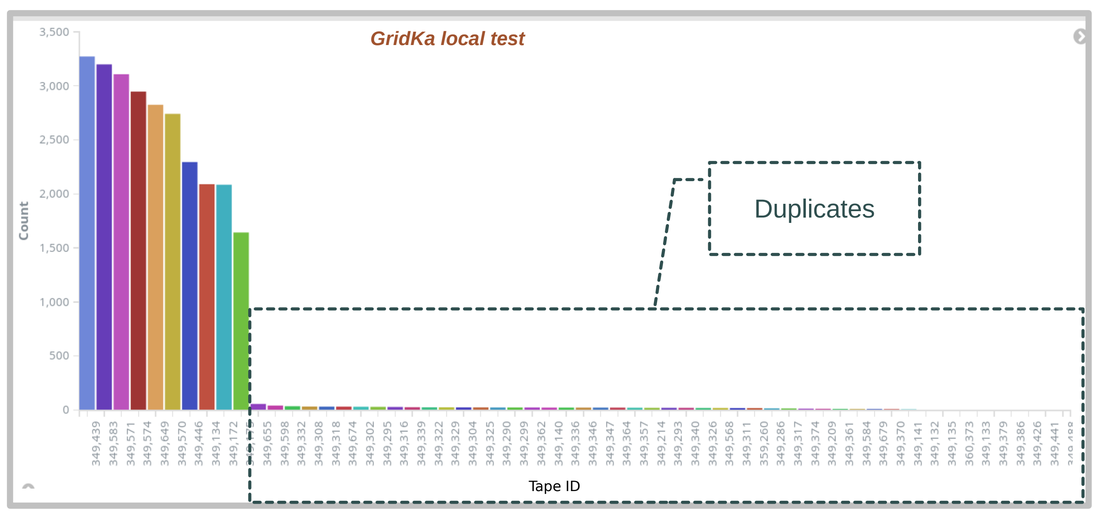
GridKa Local Test Results

- Avg read rate: ~160 MB/s
- Filesize: >10 GB
- Nr. of drives: 1
- Avg write rate: ~110 MB/s
- Filesize: >10 GB
- Nr. of drives: 1

Duplicated files on tape

GridKa Local Test Results

- Nr. of files: ~27,000
- Tape file list: 10
- Nr. of tapes actually used: 58
- TSM (tape backend) allows duplicated files



Summary

We found 3 main bottlenecks that can be eliminated both on GridKa and on the VO side, these are:

- Increase the number of concurrent requests from 2,000 to 30,000
  - ~50% improvement in overall tape recall rate
- Remove duplicated files from tape
  - Significantly reduces the number of tape mounts/sessions
- Write and then read large files (>10 GB)
  - Increases the overall tape recall rate ~3,0 times

In our test setup, we achieve at least 50% better performance than in production setup.