Karlsruher Institut für Technologie

Lessons learned from TSM to HPSS migration

Haykuhi Musheghyan, Doris Ressmann & GridKa Team

STEINBUCH CENTRE FOR COMPUTING - SCC

GridKa Tape

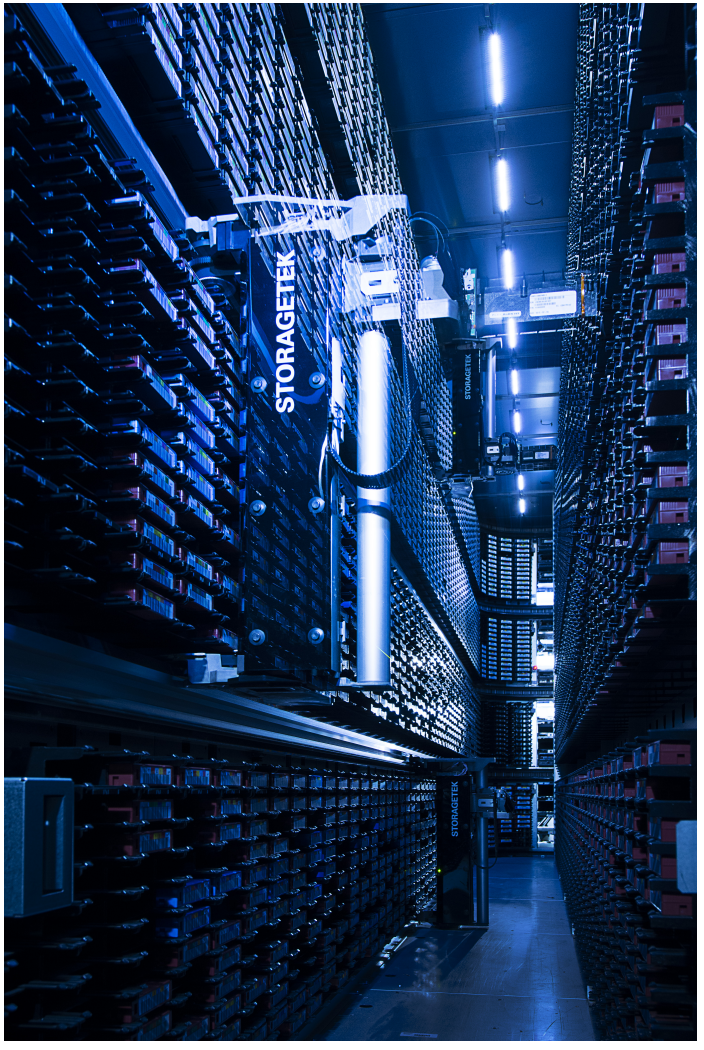
Tape Library (TSM current production)

- 1 Oracle SL8500 Library
- 35 T10k-D drives
- ~10,000 cartridges

Tape Library (HPSS migration underway)

- SpectraLogic TFinity
- 29 TS1160 drives
- 9000 slots (1000 Slots for TSM Backups)

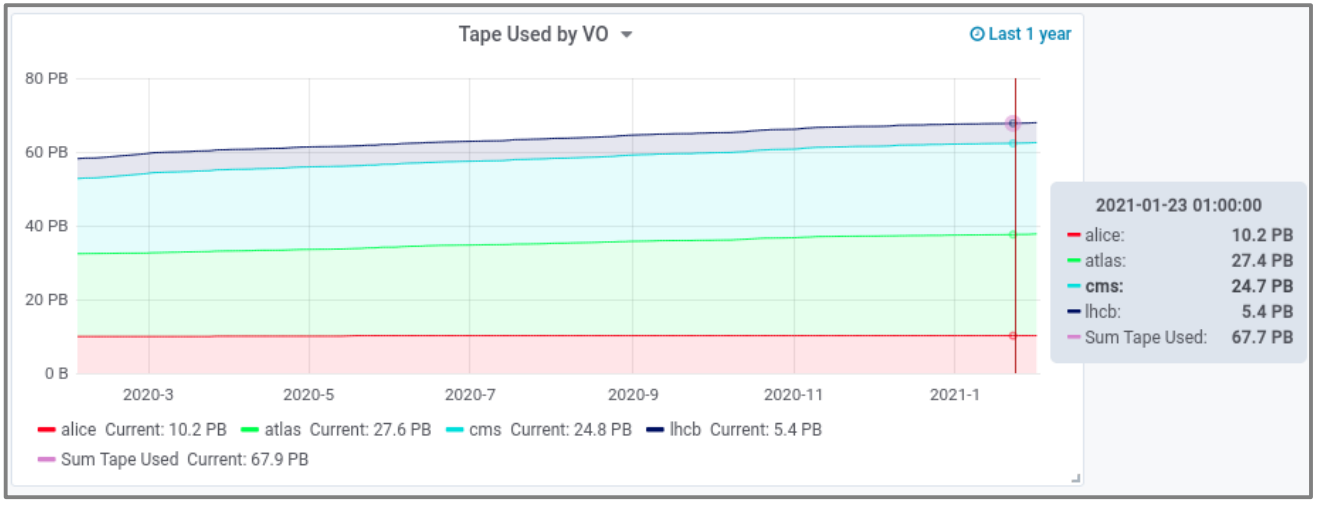
KIT operates more tape libraries outside of Tier-1 context.



TSM: Recent changes in production setup

Current GridKa Tape Setup

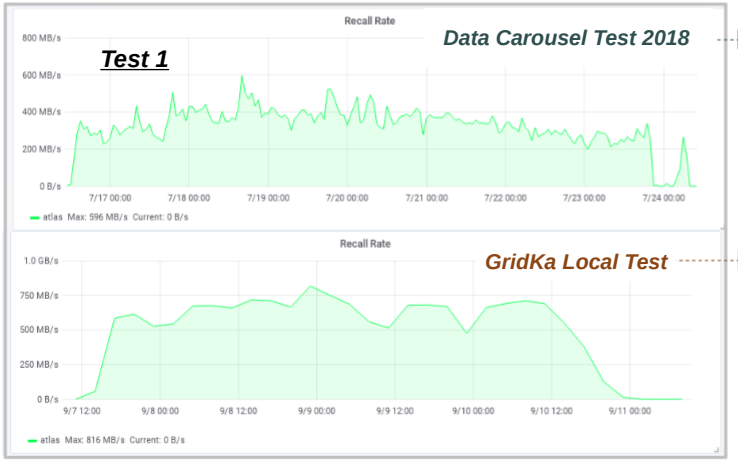
- dCache (ATLAS/BelleII/CMS/LHCb) & xrootd
- GPFS
- TSM
- TSS: queuing and scheduling software developed at GridKa based on TSM API
- Endit-TSS software: intermediate software between dCache ENDIT-Provider plugin and TSS-client
  - In production since Jan 2020 (glue scripts (before Jan 2020))



Back to 2018...

After the 1st Data Carousel Test, various tests were done

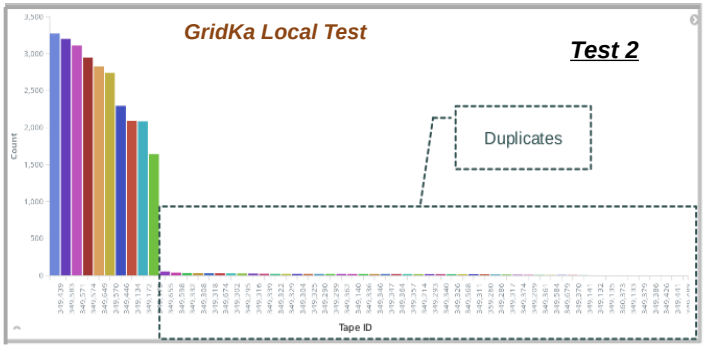
- *Test 1:* How much can the maximum number of concurrent requests be increased on a standalone TSS machine? → ***from 2,000 to 30,000***
- *Test 2:* How files are distributed on tapes? → ***many tapes and duplicated files***
- *Test 3:* Does file size affect overall tape rate and how? → ***the bigger the better***



MaxRequests: 2.000
Avg rate: ~400MB/s
Drives: 8

MaxRequests: 30.000
Avg rate: ~600MB/s
Drives: 8

Max rate: ~210MB/s
Drives: 1



Lessons Learned from TSM/TSS

Bottlenecks that can be eliminated both on GridKa and on the VO side

- The number of concurrent requests can be increased from 2,000 to 30,000.
  - New limit (30,000) due to limitations on the TSS side
  - ~50% improvement in overall tape recall rate
- TSM allows duplicated files.
  - Removing them significantly reduces the number of tape mounts
- Writing and then reading large files (>10GB) directly affects the tape recall/migrate rate.
  - Increases the overall tape recall rate ~3,0 times

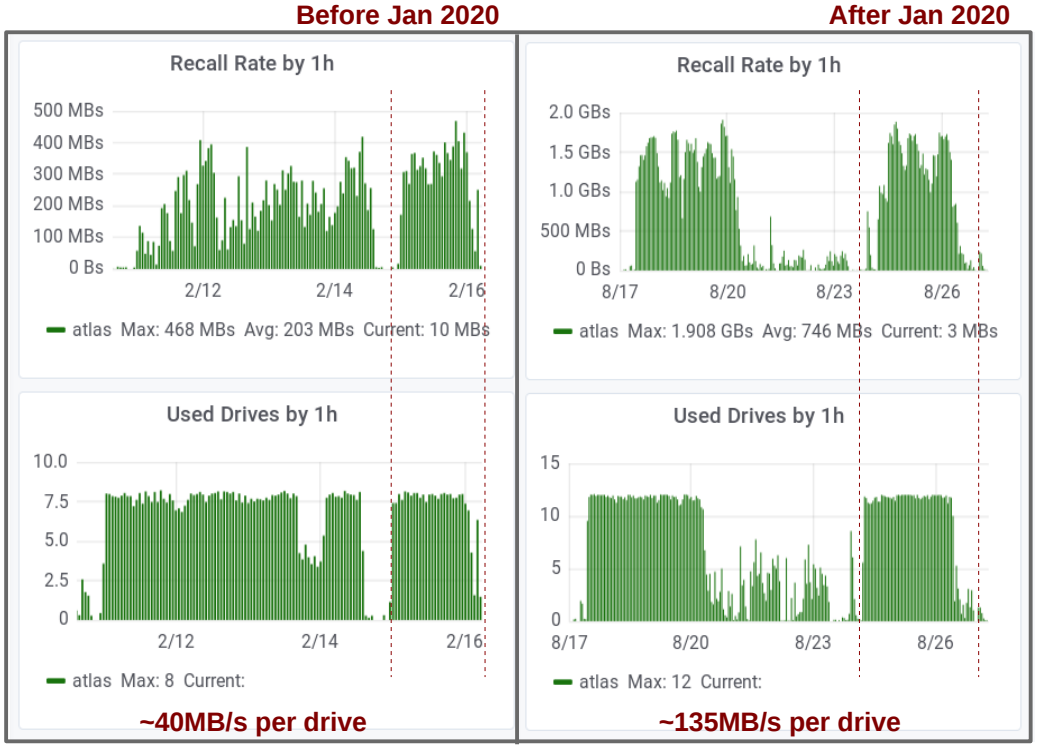
In our test setup, we achieve minimum 50% better performance than in production setup.

More details in the CHEP article.

TSM: Bringing Test Results into Production...

Latest Results from Production Setup

- detection and elimination of bottlenecks,
- use of the dCache Endit-Provider plugin,
- new software Endit-TSS for efficient recalling from tape,
- new hardware and faster network.

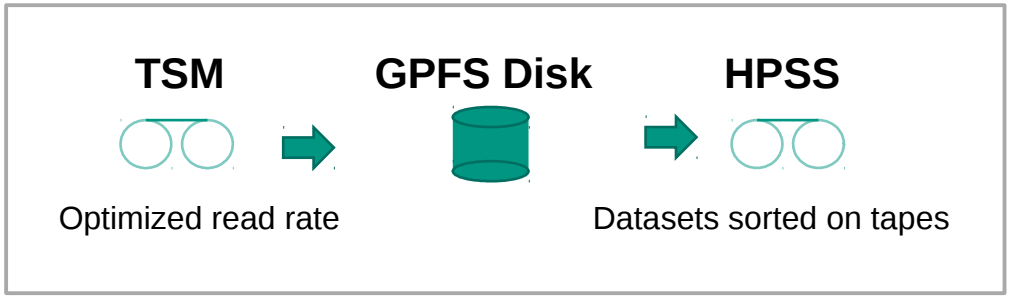


As a result, the tape recall rate is improved by more than factor of 3 per tape drive.

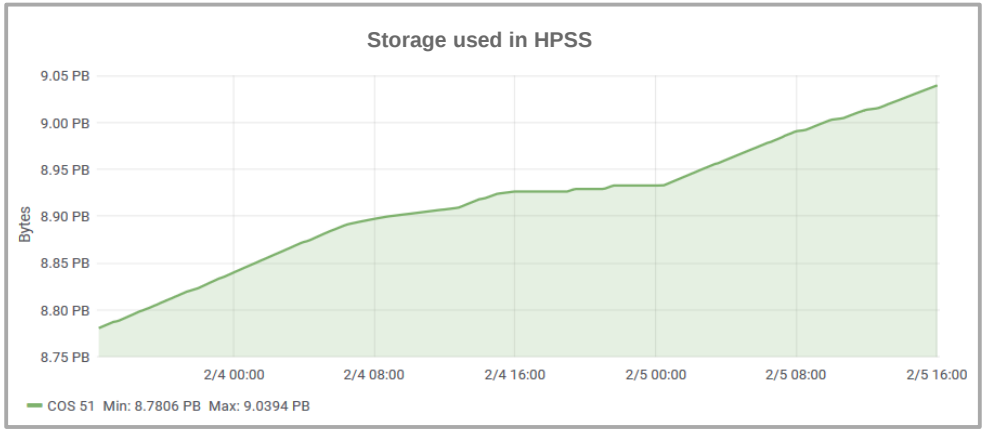
HPSS: Data Migration from TSM to HPSS

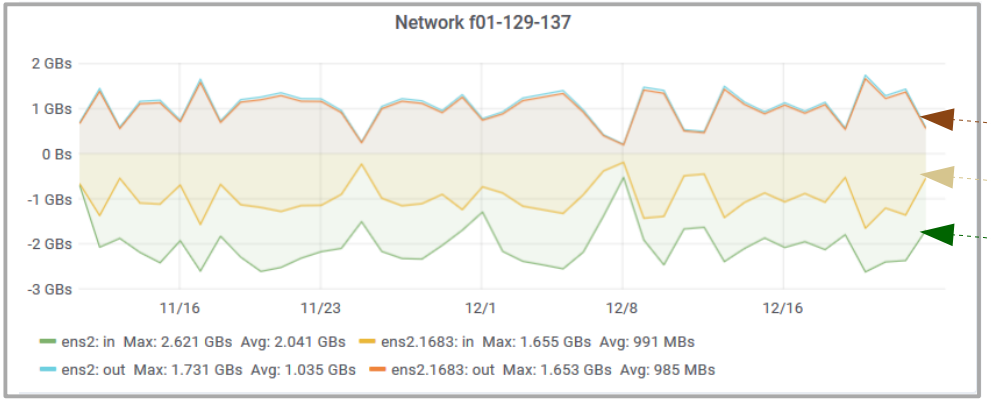
HPSS Migration Process

- Data is transferred outside of dCache
- Query chimera DB for file names
- Recalling complete datasets via TSS to Disk
- Writing dataset via pftp to HPSS
- Comparing the checksum



Transfer Rates





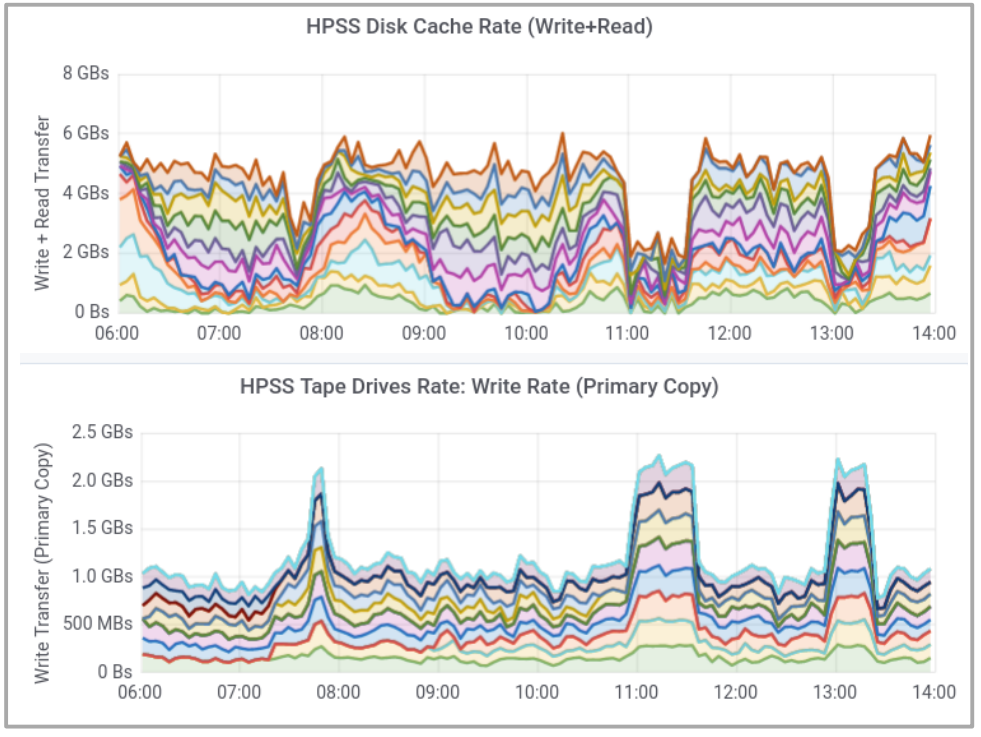
Migrating via pftp to HPSS

Calculating checksum

Staging via TSS to disk

Lessons Learned from HPSS Migration

- Sorting data on HPSS side
- Use of one drive per dataset to write
- Write big files to tape (reduce tape marks)
- Optimize recalls with maximal request rates (bulk recalls)
- Good disk cache performance is essential!



Thank you