Karlsruhe Institute of Technology

Steinbuch Centre for Computing

Efficient interface to the GridKa tape storage system

Track 1: Data and Metadata Organization, Management and Access

Author/co-author(s): **H. Musheghyan**, A. Petzold, D. Ressmann, S. A. Pérez, X. Mol, P. Konstantinov, D.Lobontu, A. Gottmann

The GridKa tape storage system was recently migrated from IBM SP to HighPerformance Storage System (HPSS) for LHC and non-LHC HEP experiments.

HPSS Tape System Overview

Writing files to tape

- Files are transferred from the dCache pool
- Files are written to HPSS disk buffer
- Checksum verification
- In HPSS, tape writes are initiated in file aggregates
- Up to 100 files ≤ 10 GiB in one directory collected in aggregates

Reading files from tape

- File read requests collected for file aggregates
- Using full aggregate recall mechanism (FAR)
- Files are read from the HPSS disk buffer into the dCache pool
- Checksum verification is done by dCache

- HPSS disk buffer/cache is shared between all VOs.

See Frame **1**

Writing Files to HPSS

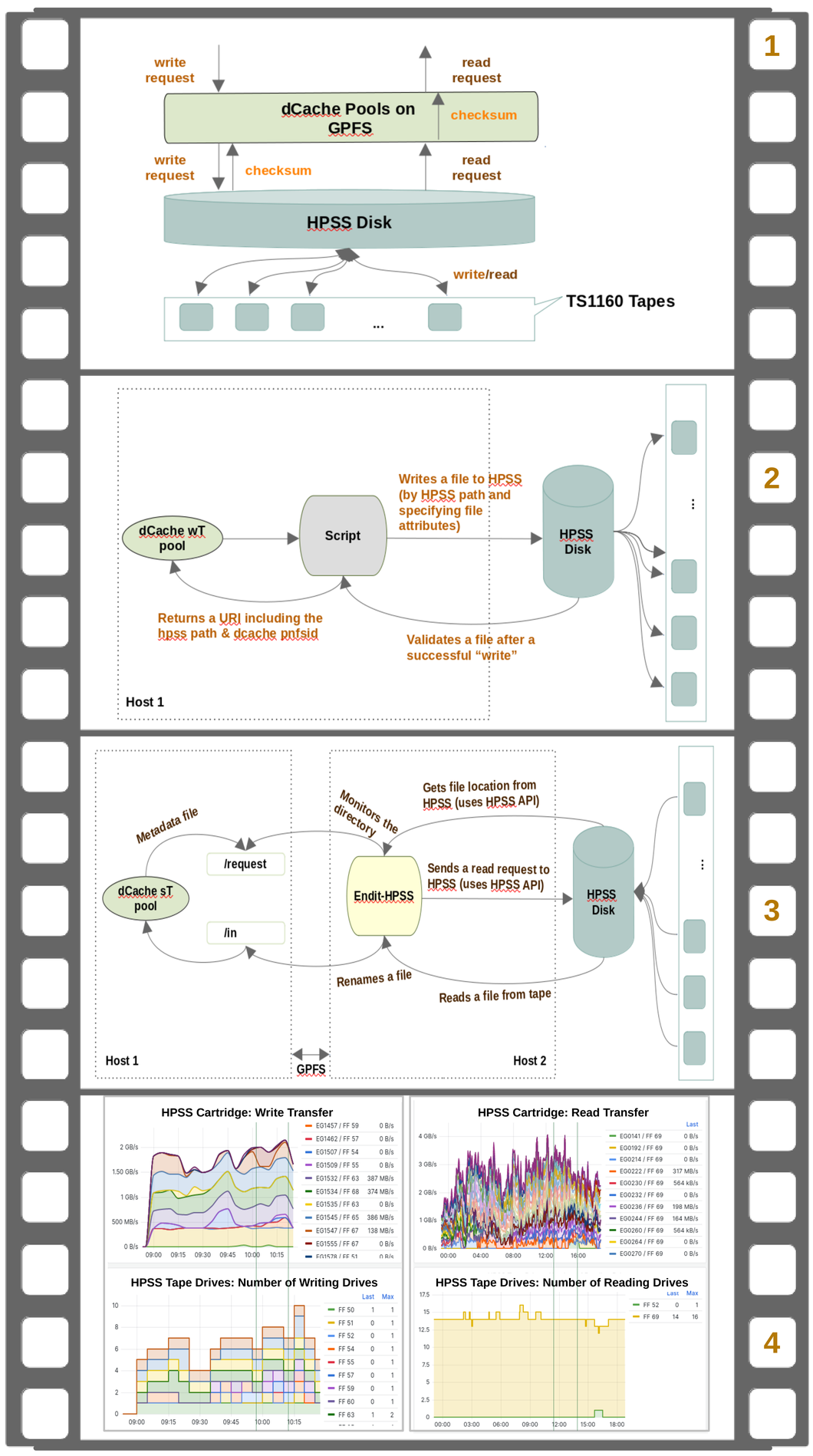
- Get LFN from dCache and converting it to HPSS path
- Write a file to HPSS disk via pftp and by setting specific HPSS attributes
- File families (FF) are integers and are reused within the same VO
- Calculate and set file family number on the fly per file
- 1 tape drive per FF for writing to HPSS is used

See Frame **2**

Reading Files from HPSS

- Extracts a file HPSS path from the dCache URI
- Queries file location attributes from HPSS
- Groups files per tape and per aggregate and creates a list(s) of them
- Iterates through the created lists and reads files sequentially from tape
- Allows to control the number of used drives per VO

See Frame **3**



Results

- Max write rate (disk→tape): ~2.0 GB/s (~390 MB/s per tape drive)
- Nr. of used tape drives: 8
- Max read rate (disk←tape): ~4.0 GB/s (~380 MB/s per tape drive)
- Nr. of used tape drives: 14

See Frame **4**

Summary

- Grouping files into aggregates before writing them to tape and reading those aggregates works very efficiently.
- Having well performing HPSS disk cache is essential

The new setup works very well and the overall tape rate is improved by more than factor of 2 per tape drive.

Contact: haykuhi.musheghyan@kit.edu