

Tape facility at INFN CNAF

INFN CNAF provides storage resources for 4 LHC experiments (Alice, Atlas, CMS, LHCb) and ~30 non-LHC collaborations

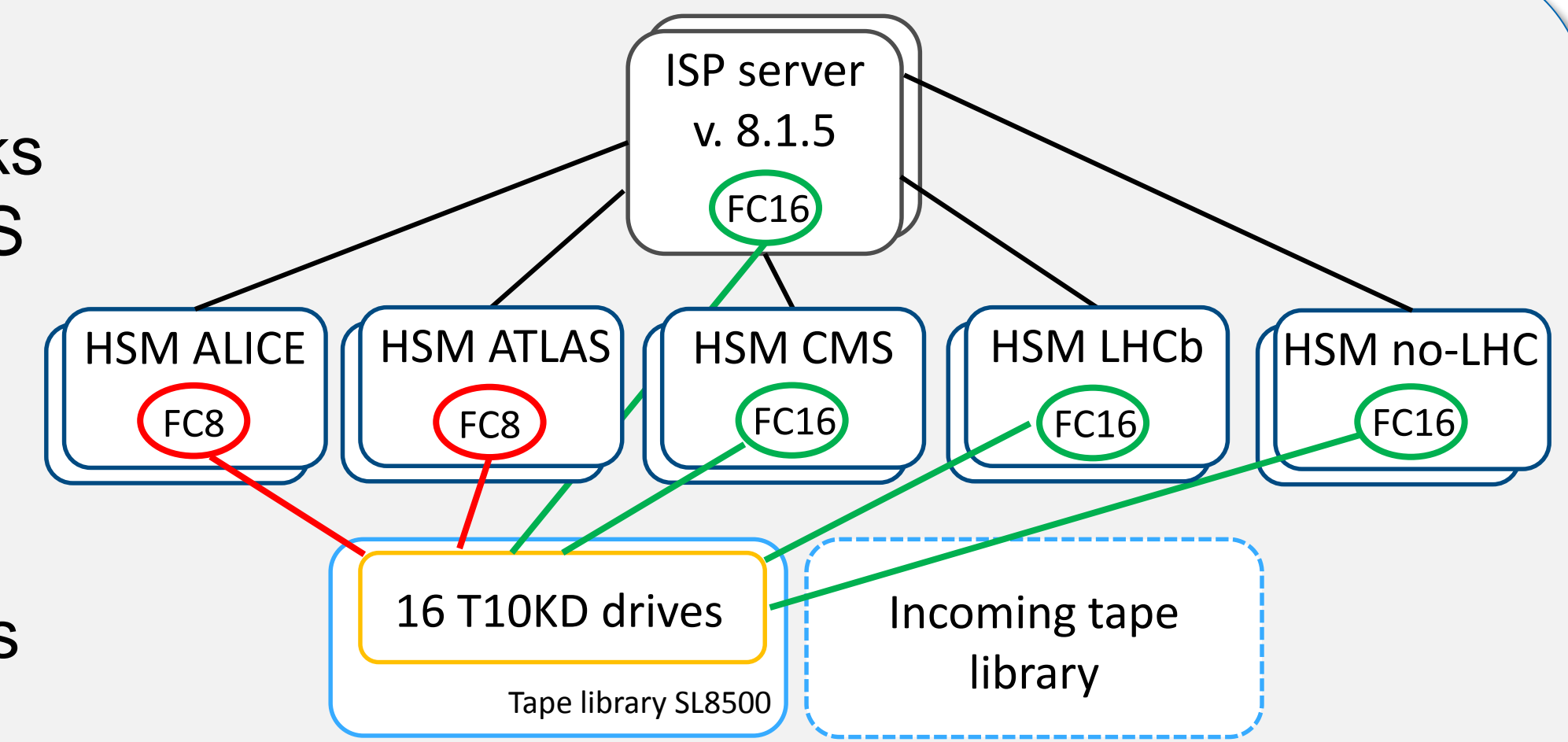
- ✓ ~ 30 PB on disk
- ✓ ~ 47 PB on tape (85 PB of total space)

Tape infrastructure components

- ✓ **1 tape library Oracle-StorageTek SL8500** (10000 slots)
- ✓ **16 tape drive T10KD** for scientific data
- ✓ **9 tape drive T10KC** for backup/archive
- ✓ **GEMSS** (Grid Enabled Mass Storage System) software developed by INFN that provides a full HSM (Hierarchical Storage Management) integration of:
 - ✓ **StoRM** (Storage Resource Manager): software released by INFN based on SRM (Storage Resource Management) interface to access storage resources
 - ✓ **IBM Spectrum Scale (GPFS)**: the disk storage software infrastructure
 - ✓ **ISP (IBM Spectrum Protect - TSM)** software: the tape system manager

HSM infrastructure

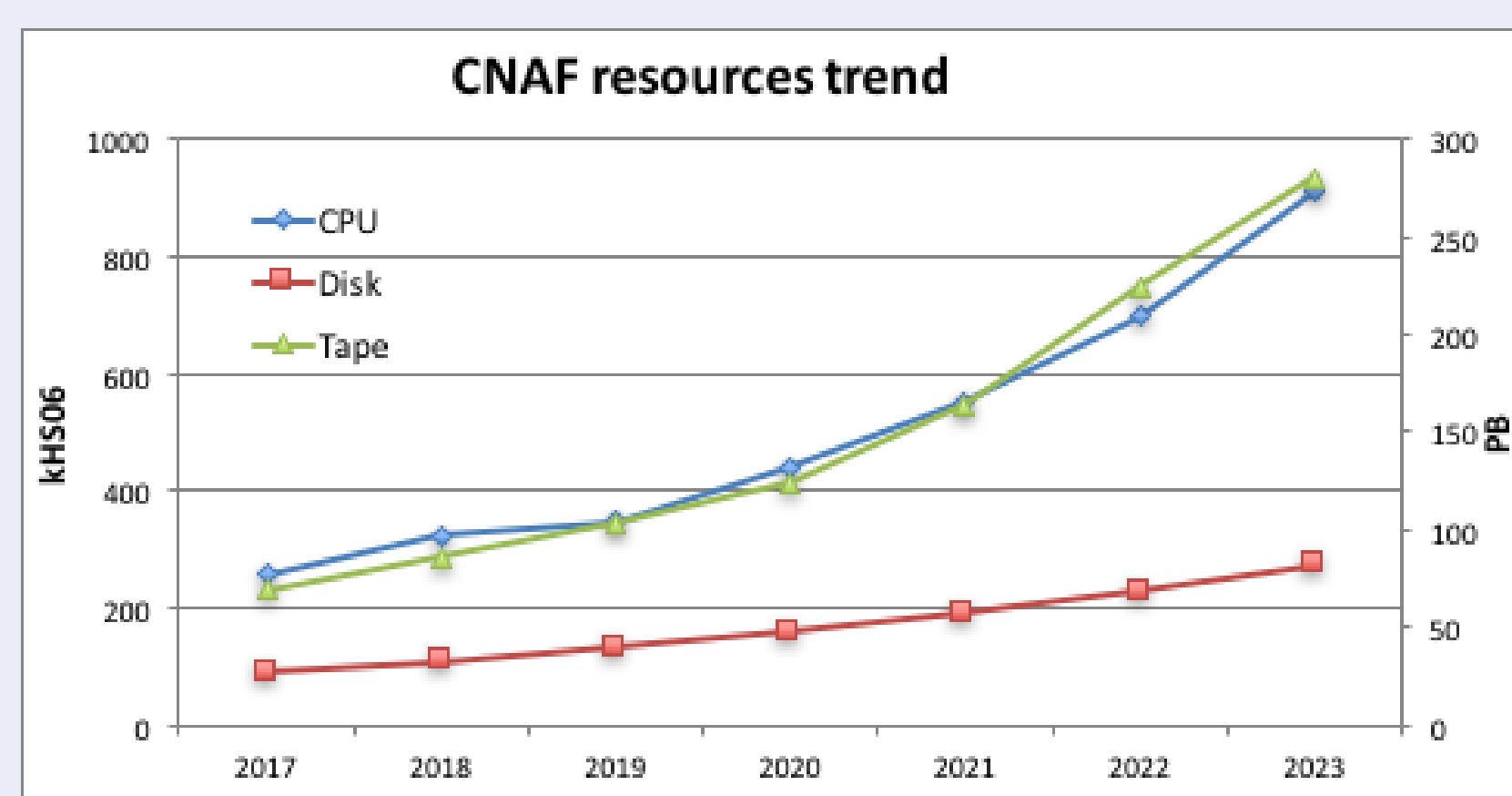
- ✓ Each HSM server works with one or more GPFS filesystems
- ✓ A disk buffer is in front of tape resources to allow write/read operations



- ✓ Migrations/recalls managed by GEMSS
 - ✓ Optimization in migration/recall management
 - ✓ Migrations are managed through GPFS policies
 - ✓ Recalls can be triggered by periodic scan of StoRM bring-online file list or by direct user requests
 - ✓ Periodic regeneration of tape ordered file lists to include new requests in already existing lists

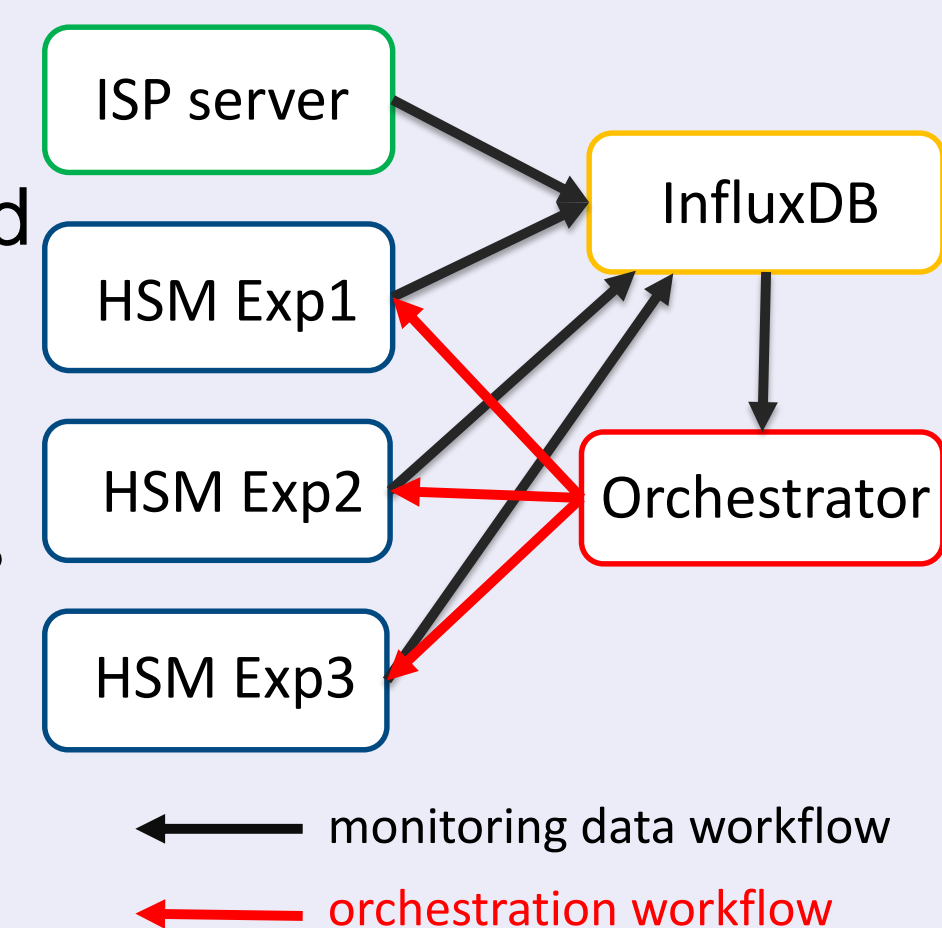
HSM I/O rate

- ✓ Writing
 - ✓ Each tape server capable of hitting the theoretical limit of 1.6GB/s (or 800MB/s) defined by the FC connection, simultaneously for in-bound and out-bound traffic
 - ✓ The real rate depends on the number of drives in use (250MB/s native rate per tape drive)
- ✓ Reading
 - ✓ Lower rate due to non-sequential reading from tape
- ✓ Capacity and demand for transfer rate to tapes will be increasing in the next years
 - ✓ 160 PB by 2021
 - ✓ Planning to buy new library
- ✓ Experiments plan to use tapes as near-line disk
 - ✓ Important increase of recall activity



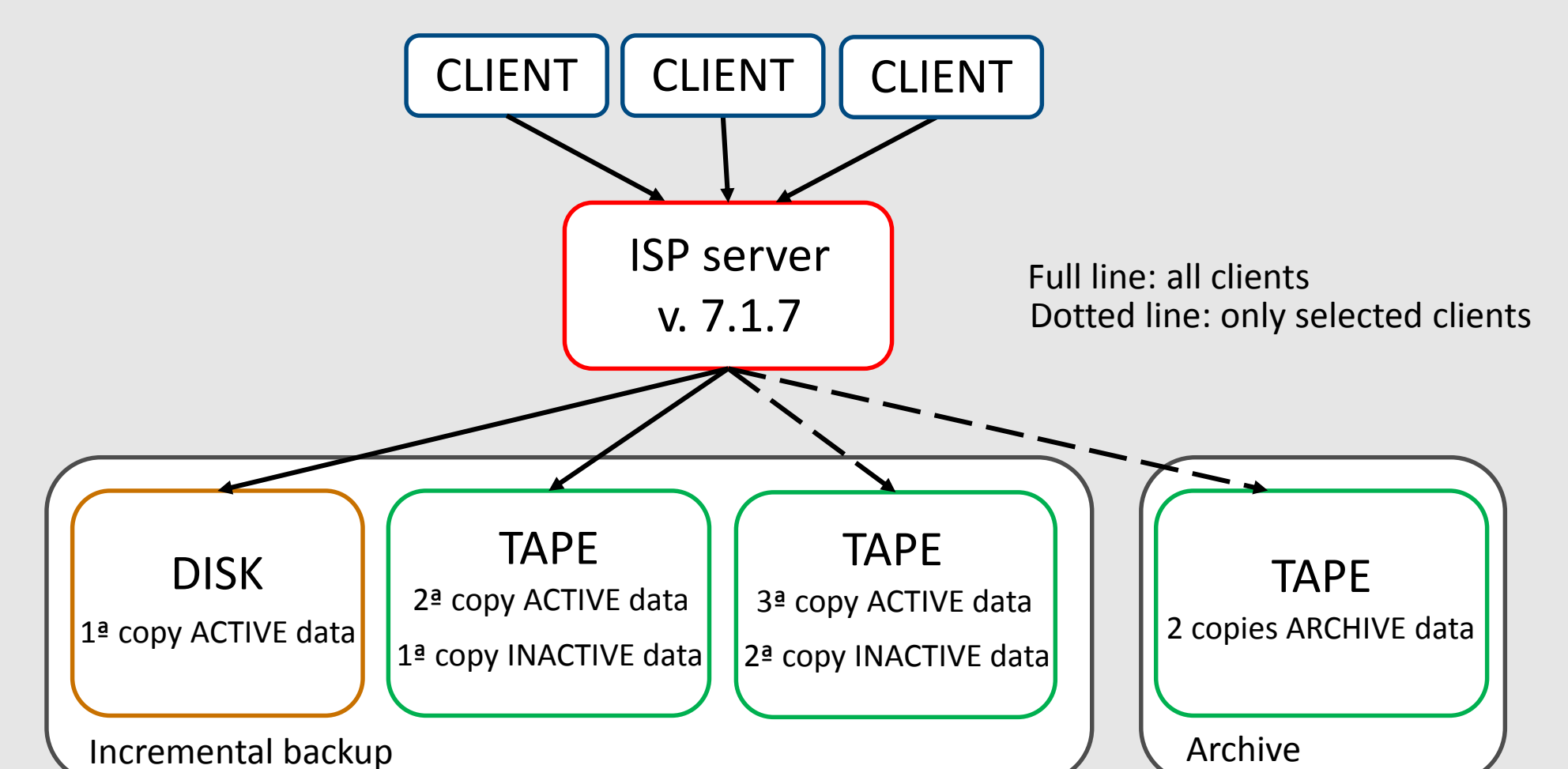
Tape drive sharing

- ✓ Tape drives are shared among experiments
 - ✓ Each experiment can use a maximum number of drives for recall or migration
 - ✓ In case of scheduled massive recall or migration activity administrators can manually allocate more or less threads
 - ✓ In several cases free drives could be used by pending recall threads
- ✓ We are working on a software solution to dynamically allocate drives to experiments
 - ✓ InfluxDB:
 - ✓ knows free drives, number of recall and migration threads running from each HSM and number of pending recall threads
 - ✓ Orchestrator:
 - ✓ performs comparison among pending threads and free drives
 - ✓ can change the number of threads on each HSM server
 - ✓ manages the concurrent access to drives setting a dynamic priority



Backup and recovery service

- ✓ Backup service to protect CNAF services data
 - ✓ Mail servers, web contents, documents, configurations, etc.
 - ✓ Services administered by different CNAF divisions
- ✓ Starting from experience of restore events last year
 - ✓ Backup service has been reconfigured to become more secure and efficient
 - ✓ Data on different media type (disk and tape)
 - ✓ Active data on disk to reduce restore time
 - ✓ Automatic e-mail to notify each CNAF division about its backup jobs status
 - ✓ Periodic restore tests
 - ✓ Collector clients to limit license usage
 - ✓ Exploiting of technologies already used in production for scientific data
 - ✓ Documentation
- ✓ Data amount
 - ✓ 21 TB / 50 millions Active objects
 - ✓ 19 TB / 42 millions Inactive objects
 - ✓ 1.5 TB / 2.2 millions Archive objects
 - ✓ Used space: 21 TB (only Active) disk; 59 TB (Active, Inactive and Archive) tape



- ✓ Incremental backup
 - ✓ Most recent version of data (Active) saved on disk and tape
 - ✓ Old versions of modified or deleted data (Inactive) saved on tape
 - ✓ Different retentions for different data (1 month or 1 year)
 - ✓ Data are sent to server and then simultaneously to disk and tape
- ✓ Archive
 - ✓ Data that will not be changed and need to be saved on tape