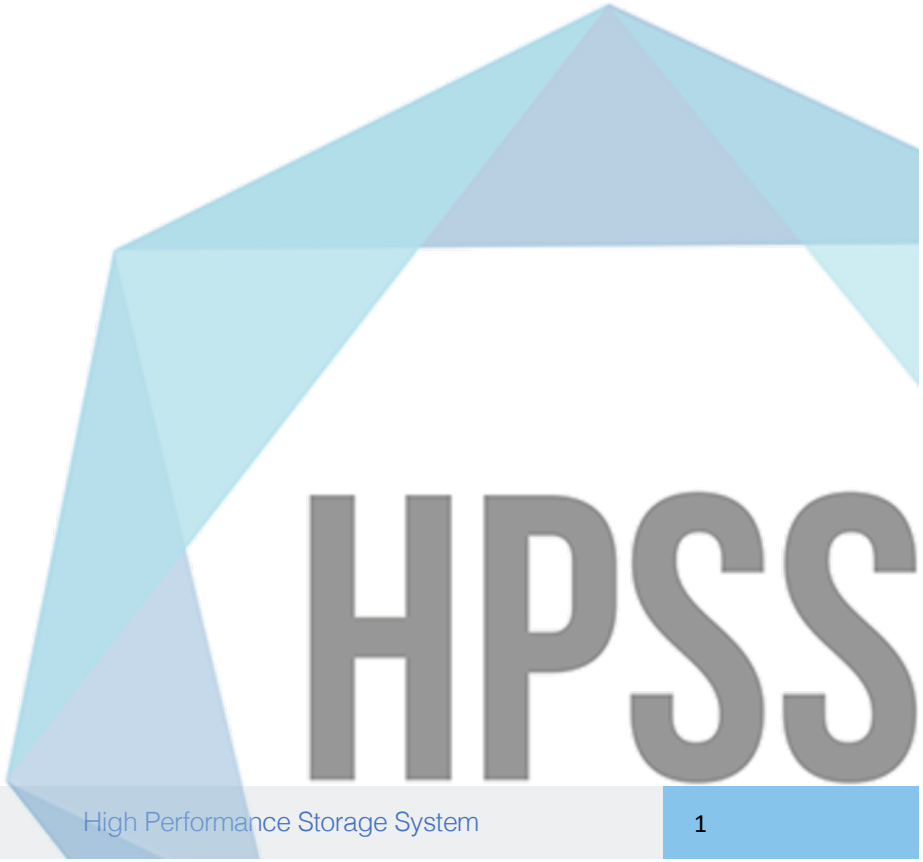
High Performance Storage System

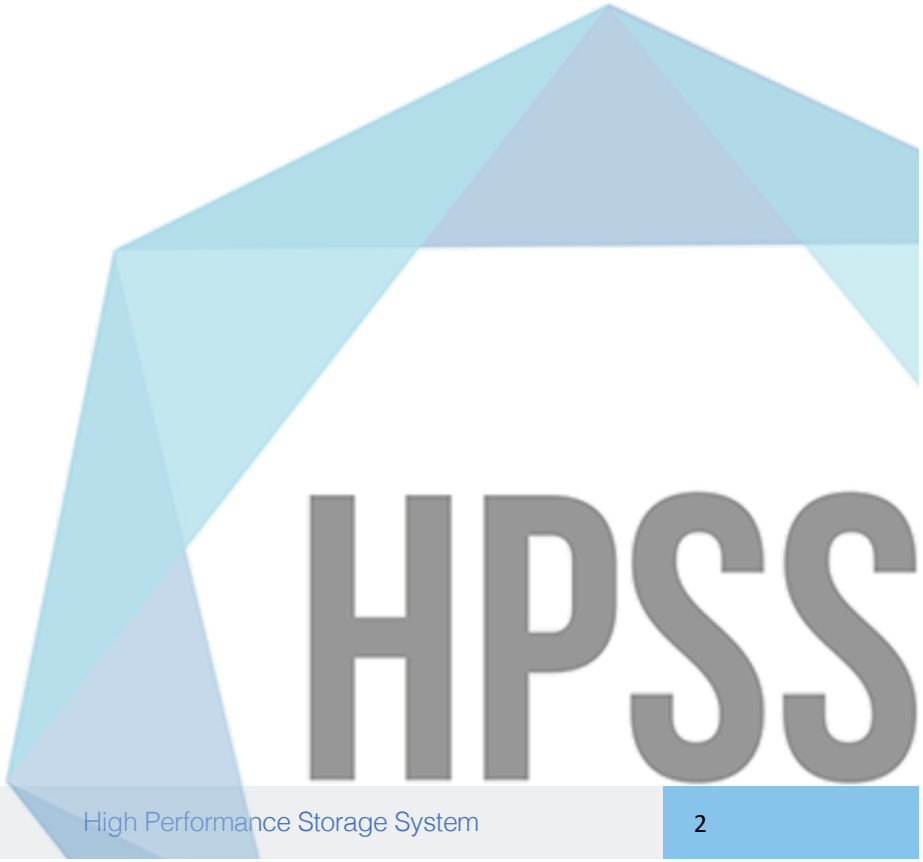
Feb 2026



Jim Gerry
HPSS Architect and Consultant
IBM Consulting

Agenda

- What is HPSS
- HPSS Interfaces
- Incremental Scalability
- Best of breed features
- What's changed since…
- What some folks are doing with HPSS
- Value of on-premises tape

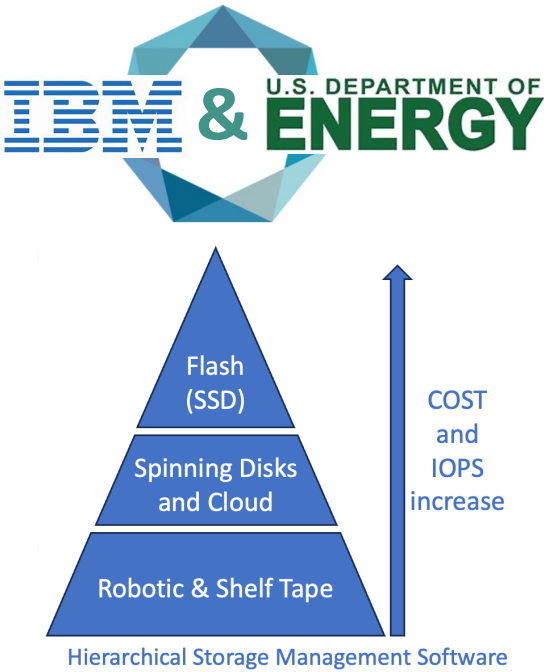


HPSS – a collaboration, storage software, services

- HPSS is a a 33-year-old development collaboration between IBM and five US Department of Energy labs that began in 1992.

- HPSS is storage software that is purpose-built to support the high-performance movement of data, through a single namespace, on-to and off-of long-term pools of storage including:
  - Lower latency flash and disk tiers for active data
  - Economical, near zero-watt, tape tiers for data at rest
  - Cloud storage for redundancy and cloud workloads
  - Local and remote storage spanning. geography

- HPSS is an IBM Consulting service offering
  - Delivery, support and consulting services
  - Not licensed by disk capacity or tape capacity
  - Annual support is the same for 10 petabytes or 10 exabytes



Big savings because HPSS is not licensed by capacity!

		Capacity (PB)	M Files	Since
ECMWF	European Centre for Medium-Range Weather Forecasts	1,450	801	2002
Shared Services Canada	Shared Services Canada	1,158	82	2017
Met Office	United Kingdom Met Office	845	2,011	2009
METEO FRANCE	Meteo France – French Weather and Climate	797	1,183	2015
Brookhaven National Laboratory	Brookhaven National Laboratory	587	312	1998
⋮				
Los Alamos	Los Alamos National Laboratory – Open	17	294	1997
NCMRWF	India's National Centre for Medium Range Weather Forecasting	15	375	2020
EARTHDATA ASDC	NASA Atmospheric Science Data Center	11	51	2018
NOAA NCEI	National Center for Environmental Information	9	137	1998
UTAS	University of Tasmania Research Data Management System	6	100	2020

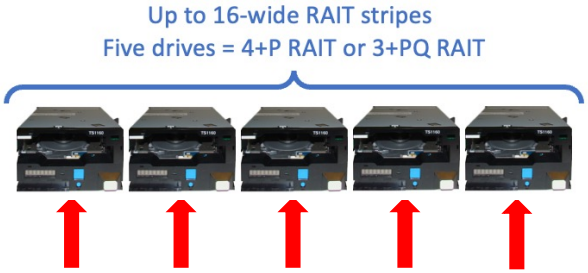
Big savings because HPSS performs and is efficient!

		Capacity (PB)	M Files	Since
Nasjonalbiblioteket National Library of Norway	Nasjonalbiblioteket – National Library of Norway	35	64	2022
IITM Indian Institute of Tropical Meteorology	Indian Institute of Tropical Meteorology	29	441	2020
Sandia National Laboratories	Sandia National Laboratory – Secure	20	196	1997
Sandia National Laboratories	Sandia National Laboratory – Open	20	445	1996
NASA Langley Research Center	NASA Langley Research Center	18	48	1998
Los Alamos National Laboratory	Los Alamos National Laboratory – Open	17	294	1997
NCMRWF National Centre for Medium Range Weather Forecasting	India's National Centre for Medium Range Weather Forecasting	15	375	2020
EARTHDATA ASDC	NASA Atmospheric Science Data Center	11	51	2018
NOAA NCEI	National Center for Environmental Information	9	137	1998
UTAS	University of Tasmania Research Data Management System	6	100	2020

Best of class features

- Best of class for tape data integrity
  - HPSS uses file and block-level checksums
  - Read-after-write verifies data on disk and tape
  - Corrupted data never makes it to tape
  - High-performance and efficient data-re-validation

- Tape stripes with rotating parity to cut redundant tape costs, and to improve data durability and tape library durability with RAIT and RAIL
  - 4+P HPSS RAIT cuts redundant tape costs by 75% over dual-copy
  - 4+P HPSS RAIT stripe @ 1.6 GB/s to write a 1 TB file in 11 minutes



Best of class performance and efficient…

- 90-10 Rule – our experience
  - 90% of files by count take 10% of the capacity
  - 10% of files by count take 90% of the capacity

90-10 Rule

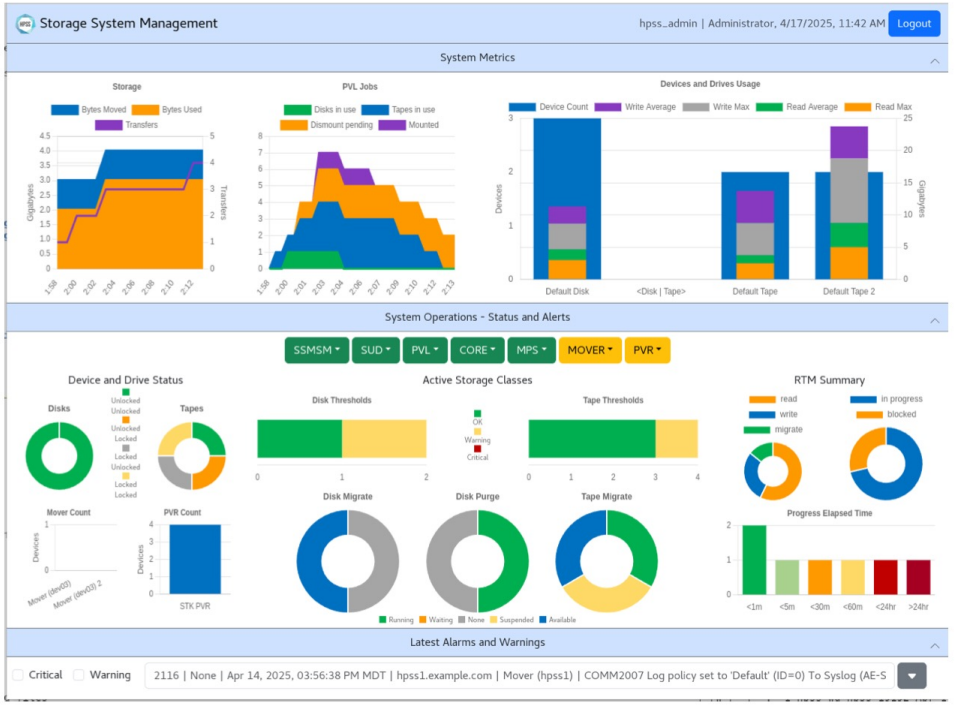
- HPSS achieves near-native tape performance with small and large files
  - HPSS uses fewer tape drives to meet requirements
  - Small file aggregation and buffered tape marks to stream small file tape writes

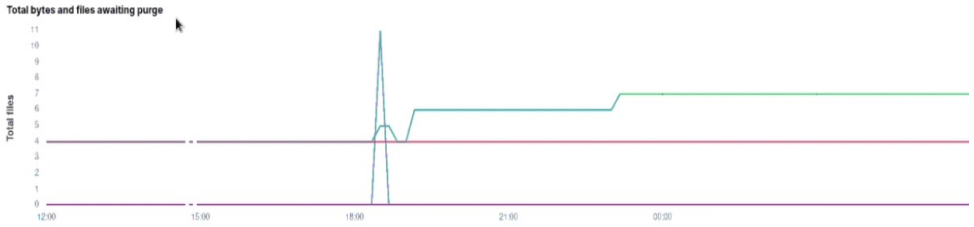
More Data Less Hardware

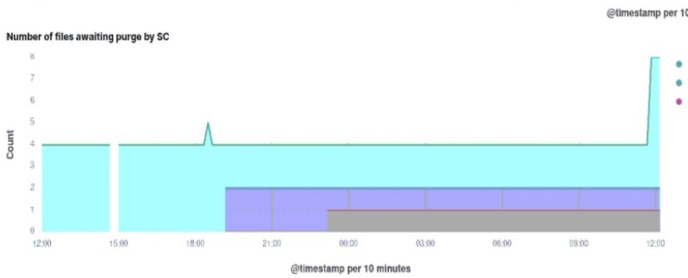
- HPSS enables high-performance tape recalls
  - HPSS automatically groups files on tape by policy
    - Reduces tape mounts and tape seeks on future recalls
  - Full aggregate recall enable streaming tape reads for small files
  - Fully integrated with modern tape drive feature called “RAO” (recommended access ordered)
    - Reduces tape seeks, thereby improving recall efficiency and tape cartridge longevity

Metrics & visualization…

- HPSS tools output data in JSON
- Easily ingested into systems like Kibana, Victoria Metrics, Grafana, etc.
- HPSS has a set of Kibana dashboard templates

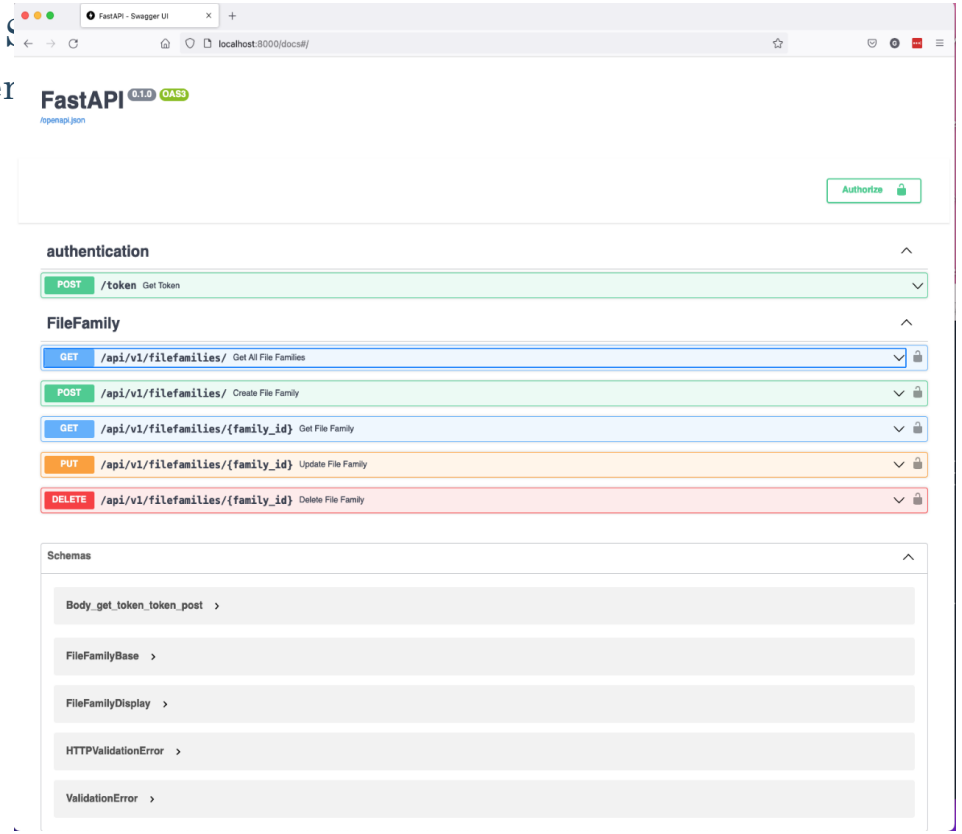






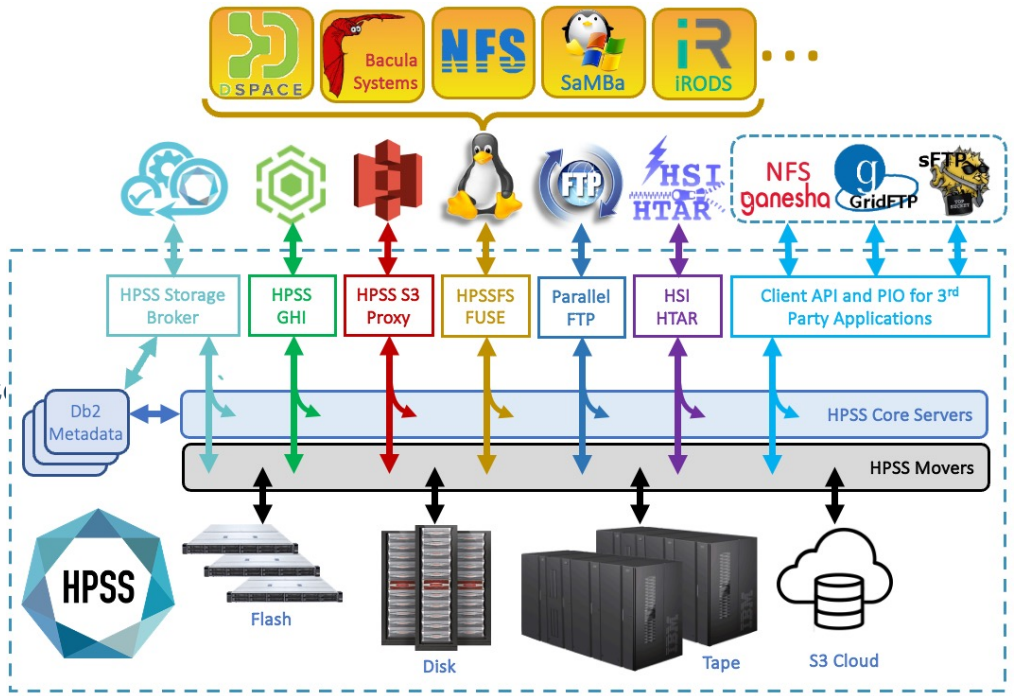
RESTful system management

- Improve ability for sites to programmatically access S
- Enable RESTful access with no client-side deploymer
- Provide a REST interface to interact with SSM
- API Versioning



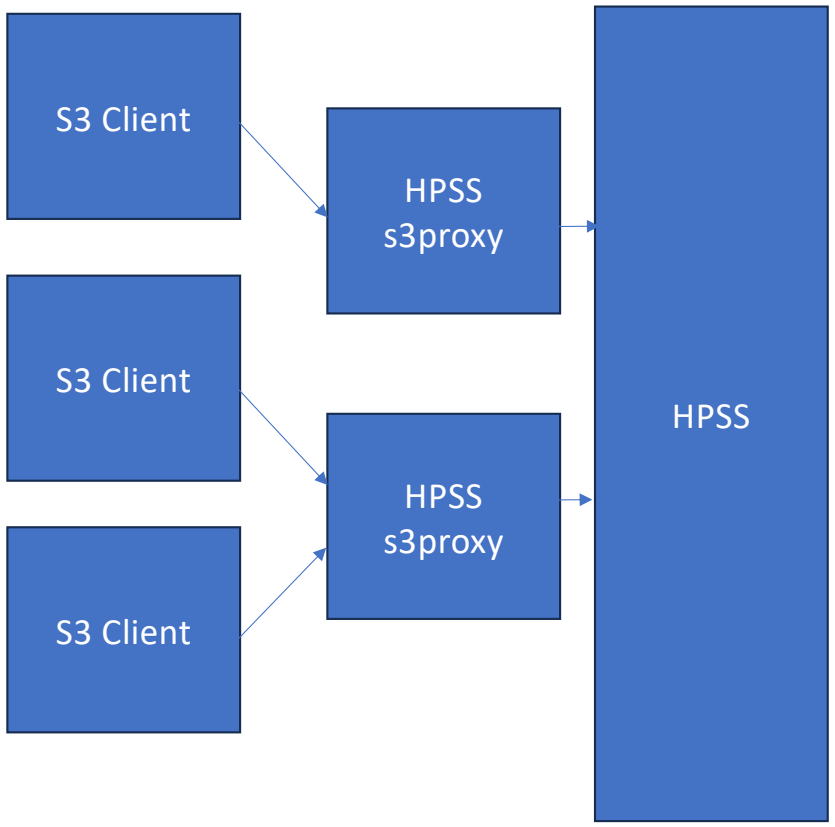
Interfaces for flexibility

- HPSS Storage Broker for datasets container
- HPSS GHI for IBM Storage Scale automation
  - Policy driven space management, and
  - Scale-Out Backup-Restore (SOBAR) support
- HPSS S3 Proxy for on-prem cloud workloads
- HPSSFS FUSE for tape-friendly mount points
  - Export Linux mount point for third party tools:
    - D-Space and Globus for sharing digital reposit
    - Bacula Enterprise for site backups
    - NFSv4 export for non-Linux users
    - Open Samba providing HPSS on Windows
- Standard and Parallel FTP
- HSI and HTAR are HPSS high performance tools
- C and Python client API – for programmers

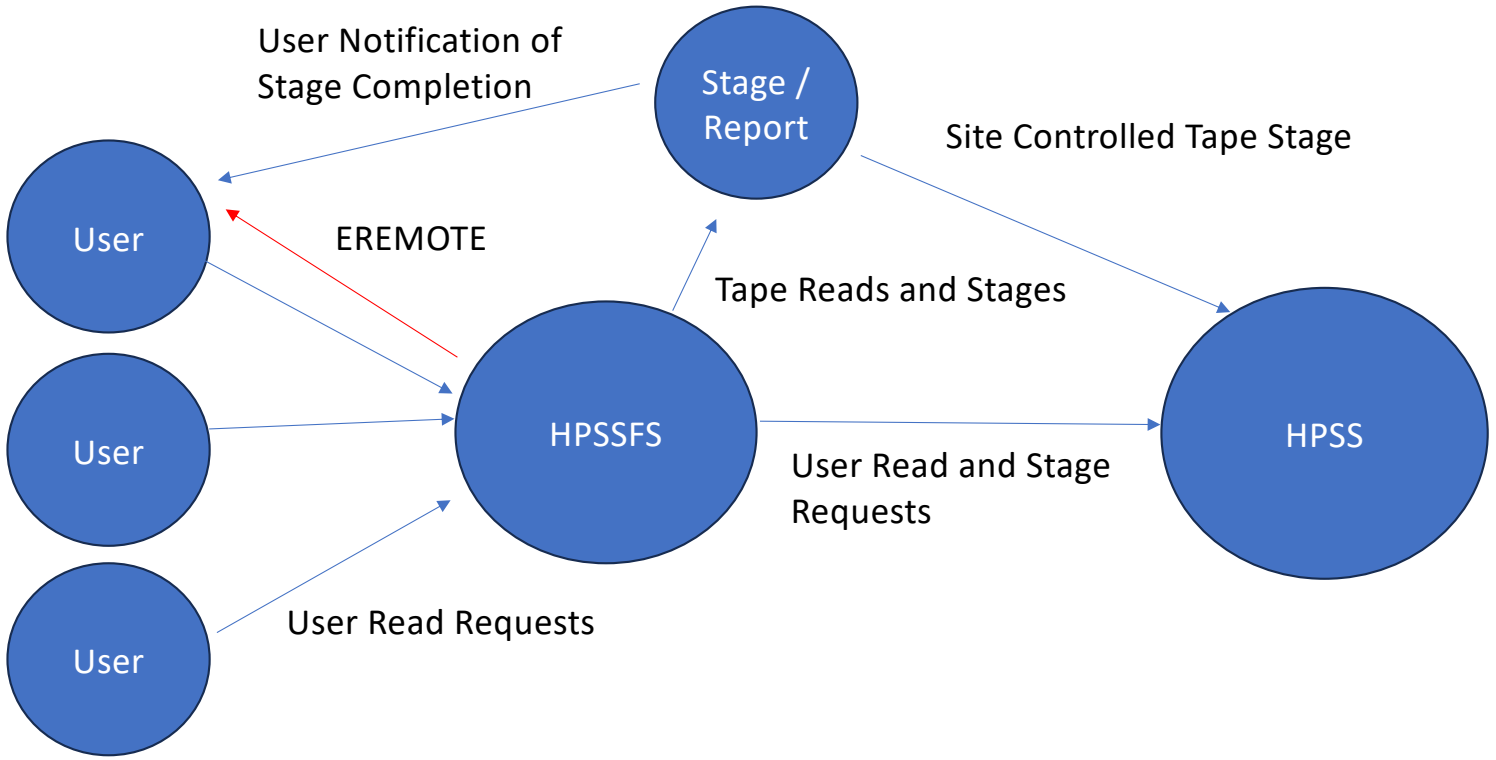


S3 interface

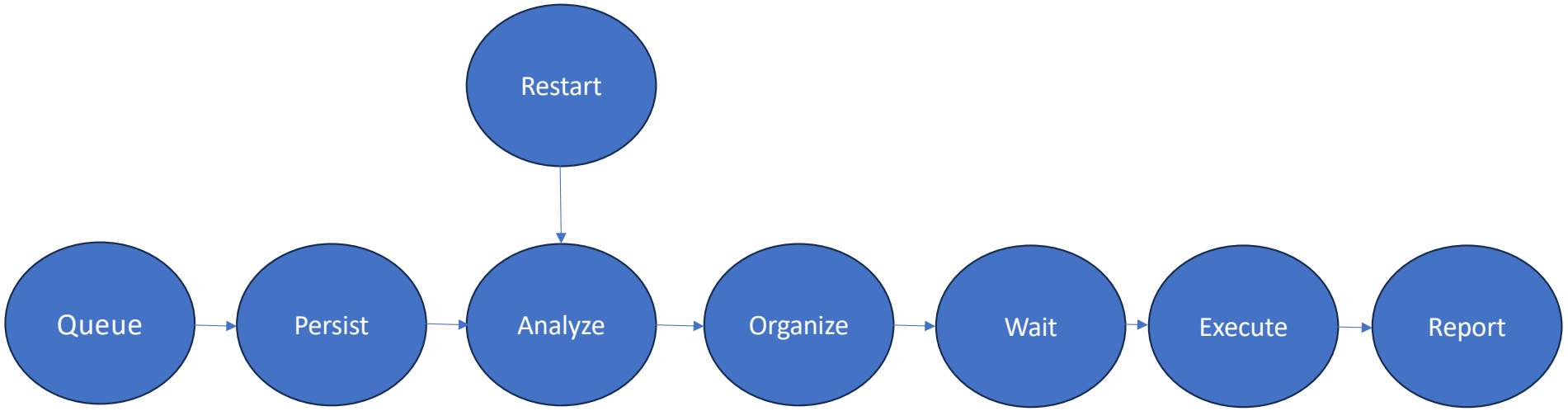
- Access HPSS through an S3 interface
- Use with aws cli, boto3 (python), rclone, s3cmd, etc.
- Optimized for HPSS I/O leveraging HSB data copy code
  - HPSS optimized I/O
  - Optimized checksumming
- Support for glacier restore operations (staging)
- No HPSS code to deploy to clients
- Full interoperability
  - Users coming in through S3 can interact with all of their files in HPSS
  - Users coming in through HPSS can interact with data coming in via S3 interface



HPSSFS FUSE Workflows



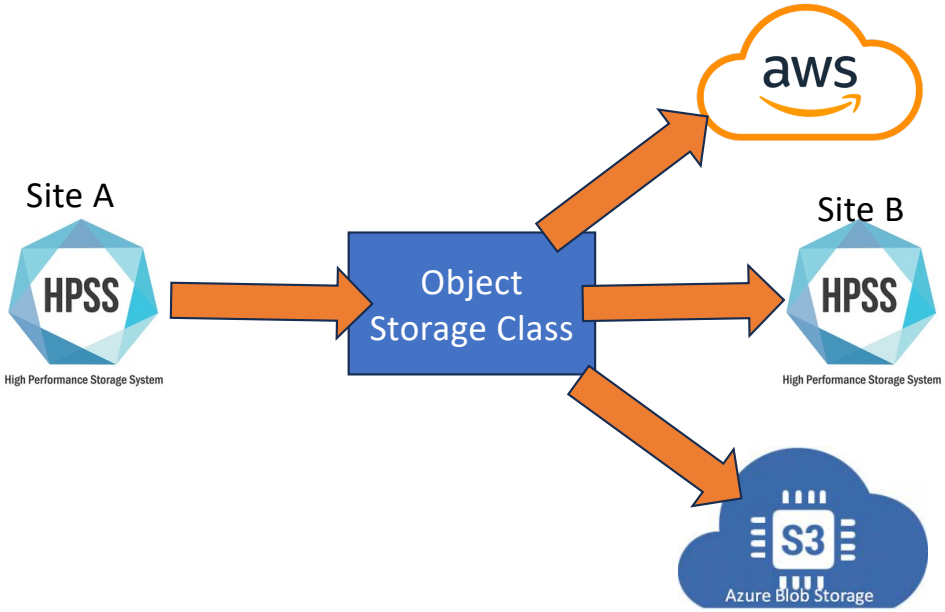
Read and Stage Queuing



- Out of the box tools that use these and enable custom workflows
- Batch client interfaces (submit 1000s at a time)
- Configurable number of active client requests for read and stage
- User requests are globally organized (by tape volume)
- No total limit

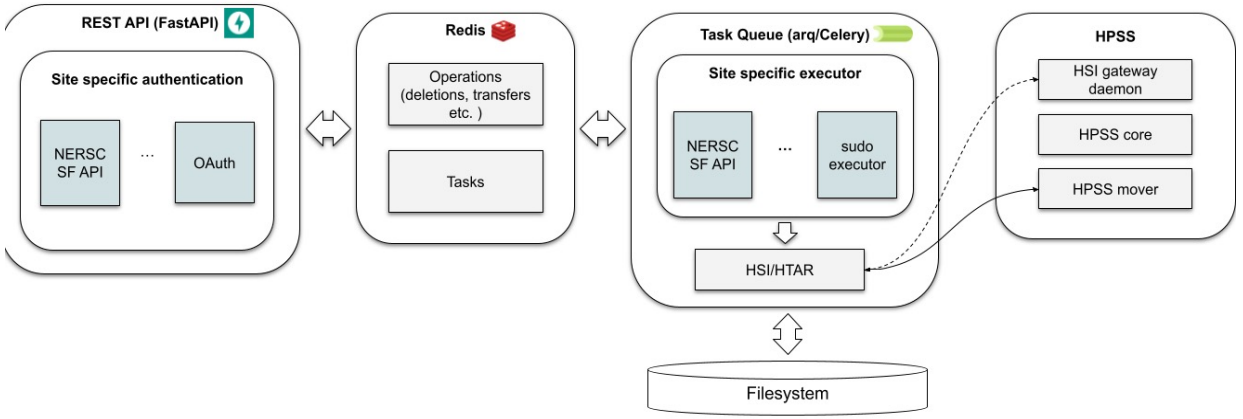
HPSS Object Storage Class

- Migration to object storage is transparent the end user
- One AWS object for each HPSS file, using the same path
- HPSS attributes (e.g. file permissions, timestamps) are reflected in AWS as tags for ease of restore
- Files stored in object storage should be treated as archive / read only copies



RESTful Data Copy Tool (Future)

- Allow users to archive data into HPSS from, or refills of data to high performance campaign stores
- This is different from other interfaces where the user is directly given the data / the data is copied out of HPSS to an area on the local user's system
- No client required
- Asynchronous, front-end queuing



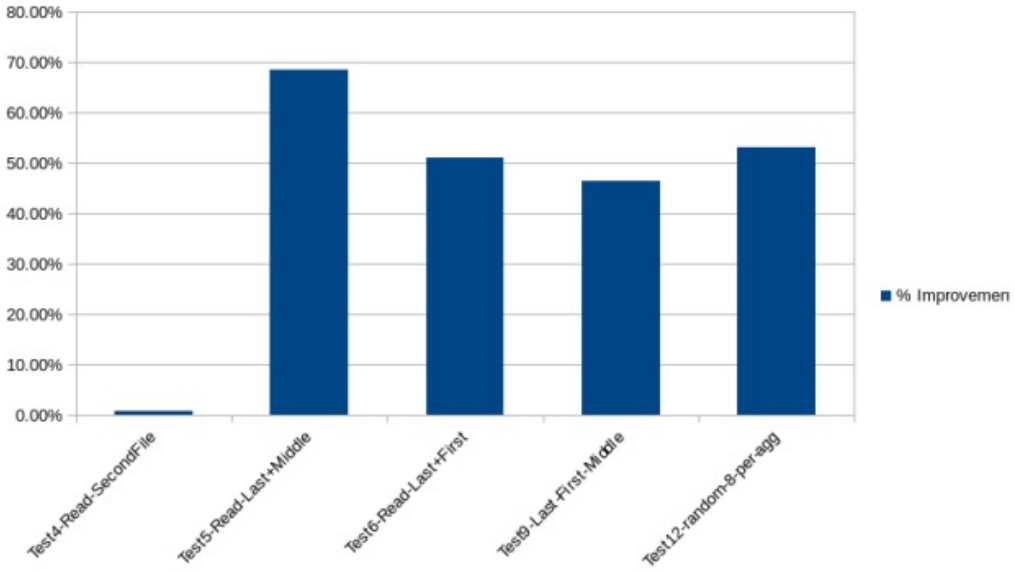
Restricted Access

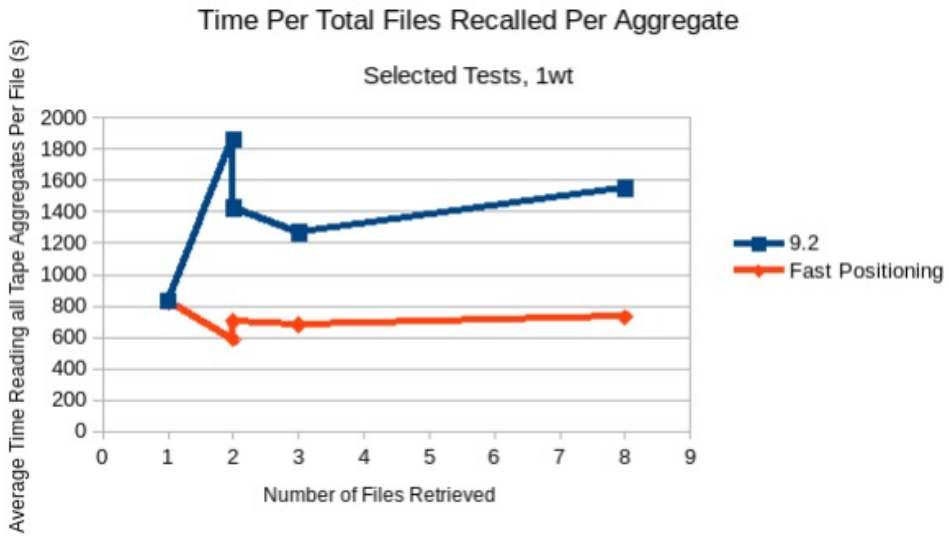
- CR 625 – Restrict Access
  - Make it easy for you to manage fine-grained user controls
  - Administrators can now restrict user access to specific operations – for example, allowing a user to read but not write, or allowing a user to stage but not delete
  - This is planning to be expanded from UID to include other criteria

```
4200 RestrictAll         # don't allow user ID 4200 to perform any HPSS operations
4042 RestrictFileCreates # don't allow user ID 4042 to create any new files (write access to existing files is OK, though)
4242 RestrictStages      # don't allow user ID 4242 to stage any data from tape
4420 RestrictFileCopies, RestrictFileWrites, RestrictFileCreates # don't allow user ID 4420 to create any new data in HPSS
```

Mover Improvements

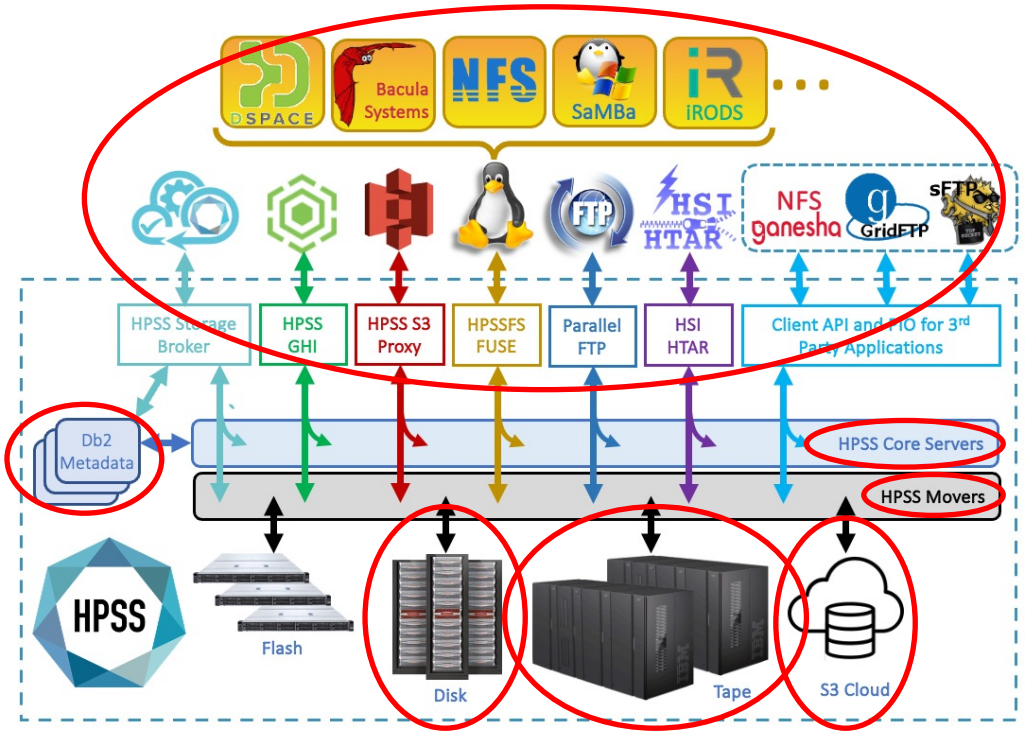
- FAST Locate
  - HPSS will position directly to the data, bypassing header reads





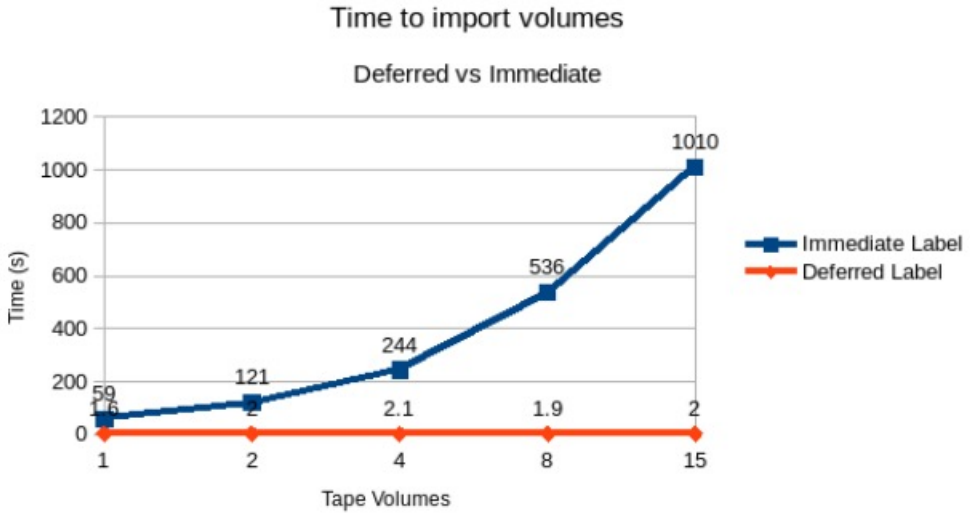
HPSS scales incrementally

- Add HPSS Core Server and Db2 Off-Host Nodes to scale file transaction performance
- Add HPSS Metadata Storage to scale file count capacity and Db2 performance
- Add Disk Cache Storage Units to scale disk cache bandwidth and capacity
- Add Tape Libraries to scale tape capacity, tape drive count and tape mount rate
- Add HPSS Disk Movers to Scale Disk Cache Performance
- Add HPSS Tape Movers to scale tape drive count and tape bandwidth
- Add Client Nodes to scale HPSS client connections and client performance
- HPSS cloud-tiering to send HPSS files to clouds (with AWS Glacier support)



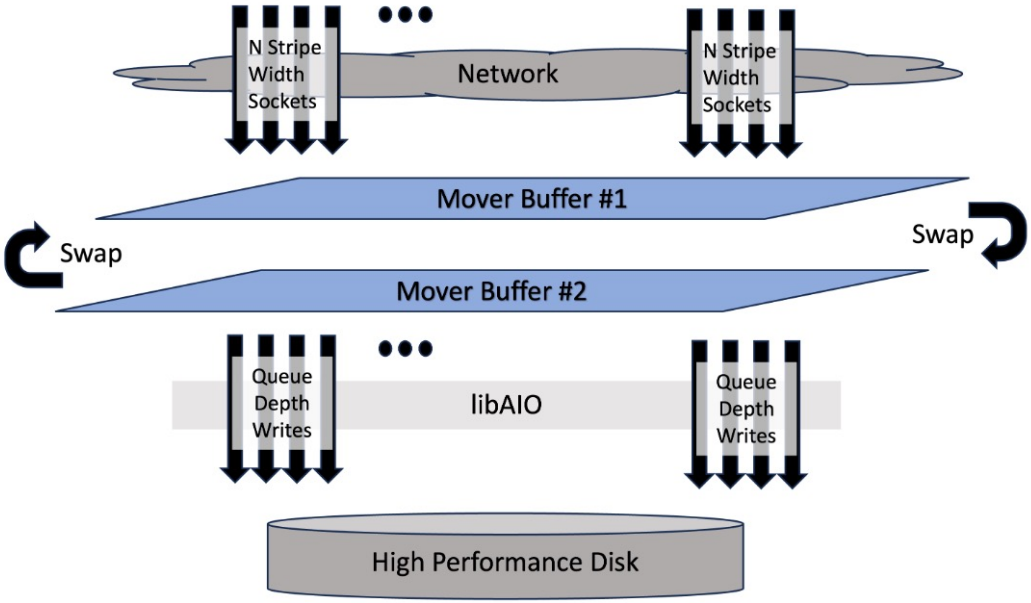
Defer Tape Labeling Test Results

- This was done with JC media and TS1140 drives.
- Deferred tape labeling added an average of 10 seconds to the load time on the first write.
- Deferred tape imports are a metadata only operation that takes very little time.
- Overall savings of **50 seconds** of drive time **per cartridge**.
- HPSS no longer requires babysitting long running import operations.



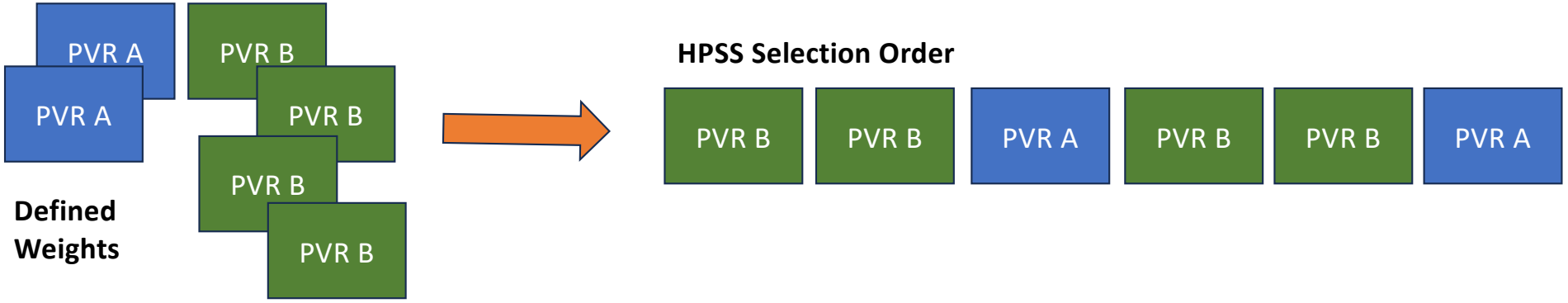
Mover Improvements

- Multi-socket Transfers
  - HPSS clients can request multiple sockets for a data transfer, improving bandwidth availability to disk



Balance New Tape VV Selection

- More on policy-based selection
  - Weighting scheme
  - For a given tape SC, each PVR has a weight
    - Higher weight = chosen more often
    - Lower weight = chosen less often



Value of on-premises tape

- IBM Zurich Research Lab continues to advance tape technology

- At scale, on-premises tape is the cheapest storage.
  - Deep archive cloud is near on-premises spinning disk
  - Idle data demands almost zero watts
  - Tape-workflows enable greater compute workloads

- HPSS software is considered best of class for tape
  - High performance transfers
  - Economical at scale – not licensed by capacity
  - Efficient use of tape hardware
  - Enables the most aggressive tape workloads