HPSS 7.5.2 Update

New Features for the Best of Breed

Disclaimer

- Forward looking information including schedules and future software reflect current planning that may change and should not be taken as commitments by IBM or the other members of the HPSS collaboration.

Summary

- HPSS 7.5.2 is a major software release which started in August 2016 with an 18-month expected release for 1Q2018
- The original release plan incorporated an iterative development process with three defined milestones and system testing being done in parallel with development
- Additional design discussion

7.5.2 Development

- Tracking to 1Q 2018
- More than 250 bug fixes and enhancements
  - 87 enhancement bugs
- Metadata conversion will be more modest than 7.5.1
  - Same QREP process
- 7.5.2 patches will be similar to 7.5.1 patches
  - One or two a year
- A few items have dropped off due to lack of time
  - Lower priority or higher complexity items

7.5.2 Goals

- Address near-term issues while EC/TC/ATWG/Others develop priorities and plan for long-term strategic imperatives
- Deliver a release 18 months after start date
- 7.5.2 content planning focus:
  - Customer burning issues
  - New Business needs
  - TC priorities

7.5.2 Development Plan

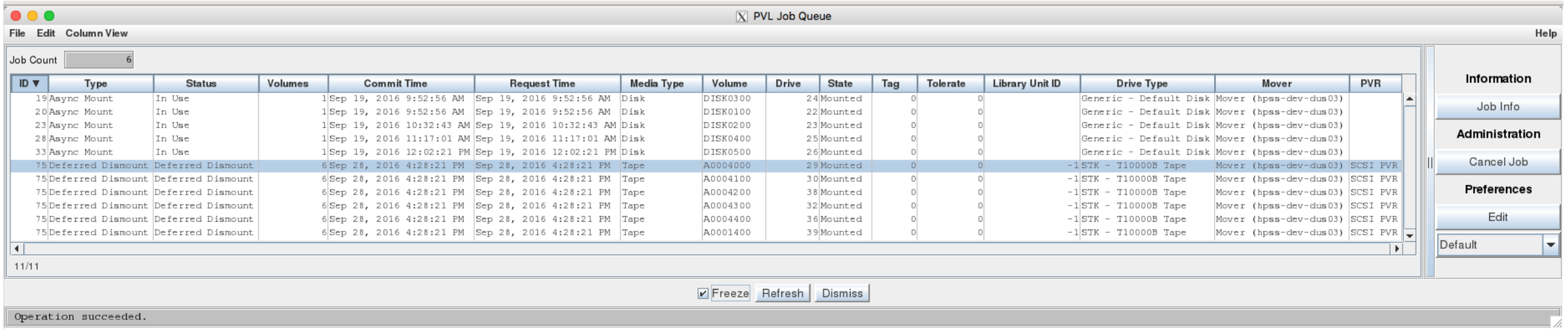
- Learning from the “Long Tail” of 7.5.1
  - 7.5.1 was so large that a “waterfall” test effort caused issues
- Need for test to proceed alongside development
  - Catch problems earlier
  - Get feedback earlier
- Iterative Development Process for Features
  - 4 milestones going back to August 2016
  - Test is already testing completed feature work
  - Shortened release cycle by months

7.5.2 Planned Content

- Use file system as HPSS device (CR 379)
  - Tested with GPFS, ZFS
  - Take advantage of filesystem features
    - Adding more resources online
    - Rebalance data across physical disks online
    - Data replication
    - Data checksumming and forward error correction
    - Data compression and deduplication

7.5.2 Planned Content

- Provide more PVR and VV information in PVL Job Request (Bug 3109)



7.5.2 Planned Content

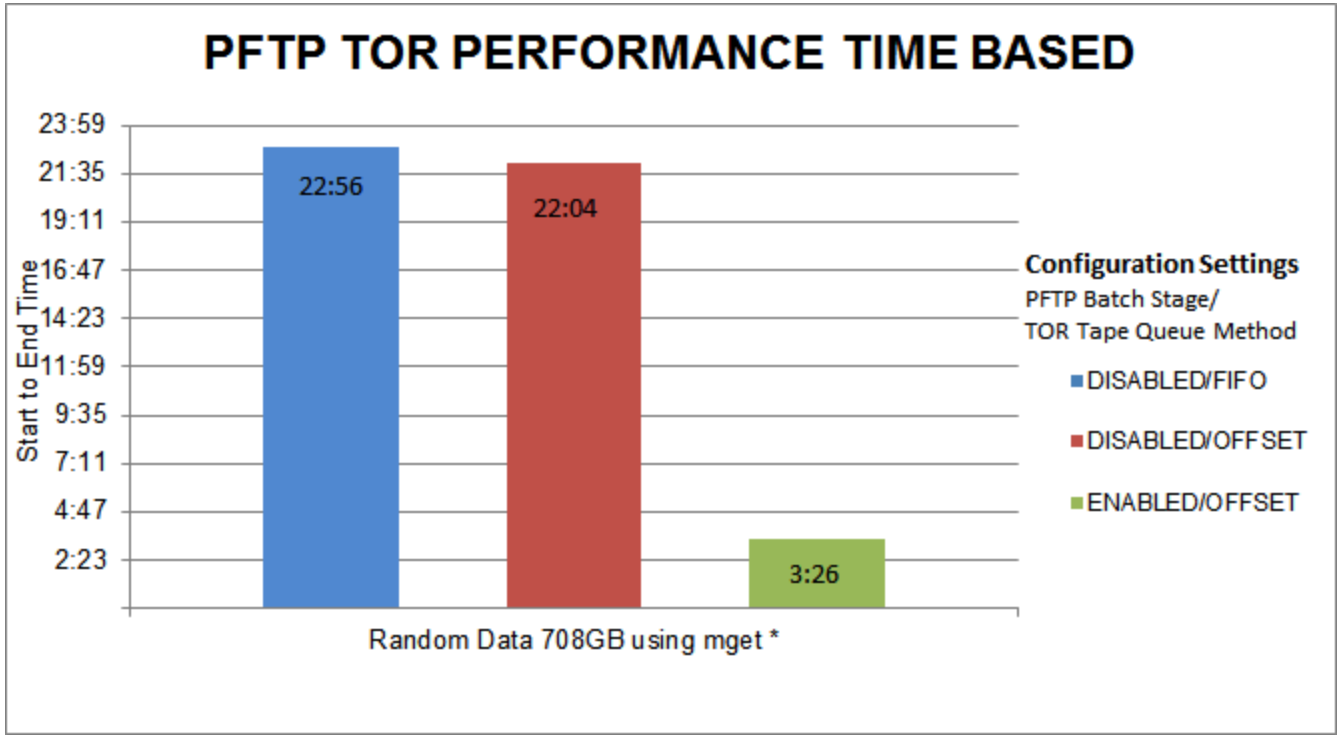
- Improve Recover workflow (Bug 5308)
  - Remove sharp edges from the tool
  - Easier to manage recover status
- New log type: INFO (CR 349)
  - Logging type for “informational” messages
  - Used in other CRs for performance or audit type logging
  - Expected to be more common than EVENT logging
  - Not problem related like a DEBUG or ALARM
  - Does not show up in Alarms and Events

```
Record type=INFO, Event time=2017/09/14 07:58:04 CDT, Severity=NONE
    Subsystem=MOVR, Message#=1049, Error code=0
    Desc name=Mover (mystic), Routine=tp_position_hpss
    PID=2209, Node=mystic.clearlake.ibm.com, User=
    Type=TEXT, Object Class=9, Request Id=0
Completed positioning, device 2, secs 0.018899, 1:18446744073709551615 absaddr=0x2
```

7.5.2 Planned Content

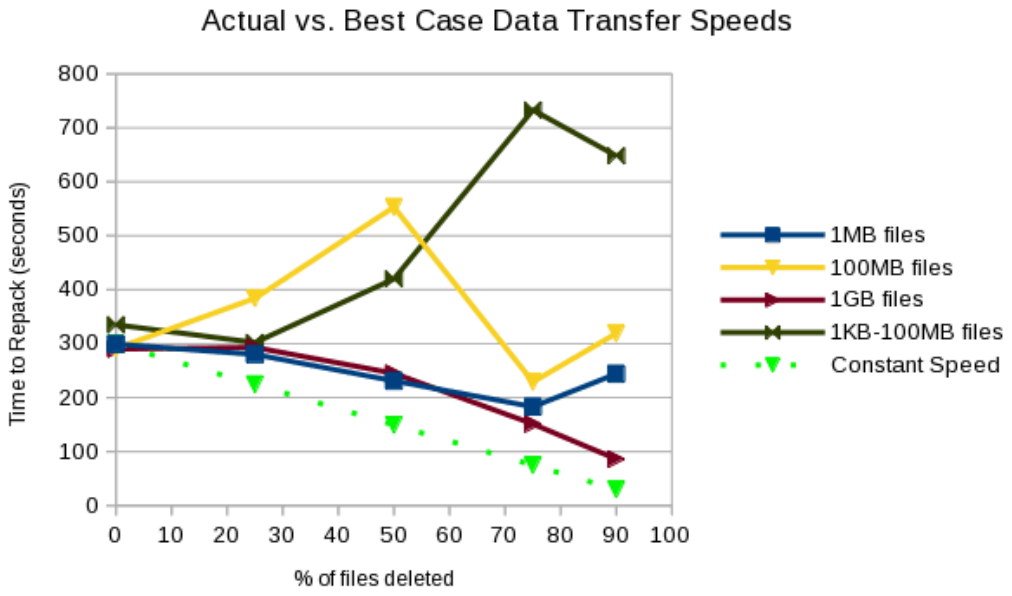
- PFTP support for Tape Order Recall (CR 357)
  - PFTP now pre-stages files it is going to request via mget
  - Defer dismount should be enabled

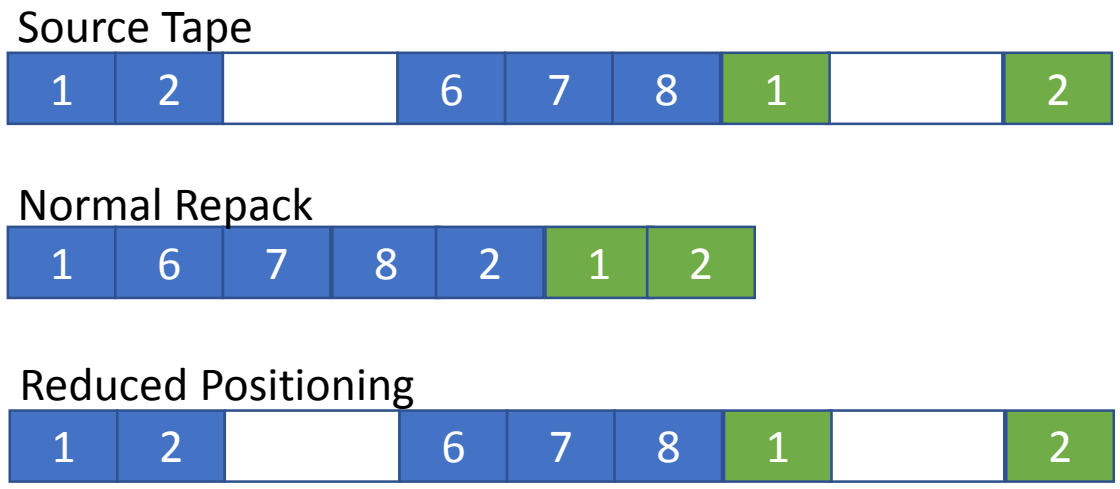
```
# Attempt batch stage on mget
# Depends on the server supporting TOR
# Default = on
; Disable stagebatch on mget
```



7.5.2 Planned Content

- Improve Sparse Repack Performance (CR 397)
  - Do not reorganize or "squeeze " aggregates during repack
  - Up to a 20% increase in performance
  - Must use for Ordered Migration to retain the aggregate organization





7.5.2 Planned Content

- Improve Recover Performance (Bug 5306)
  - If you're coming from 7.4.x, you may see Recover performance improve by 10x or more
  - For 7.5.x users, you may see a 2-3x improvement
  - Recover also now takes advantage of RAO (if available) when recovering files

7.5.2 Planned Content

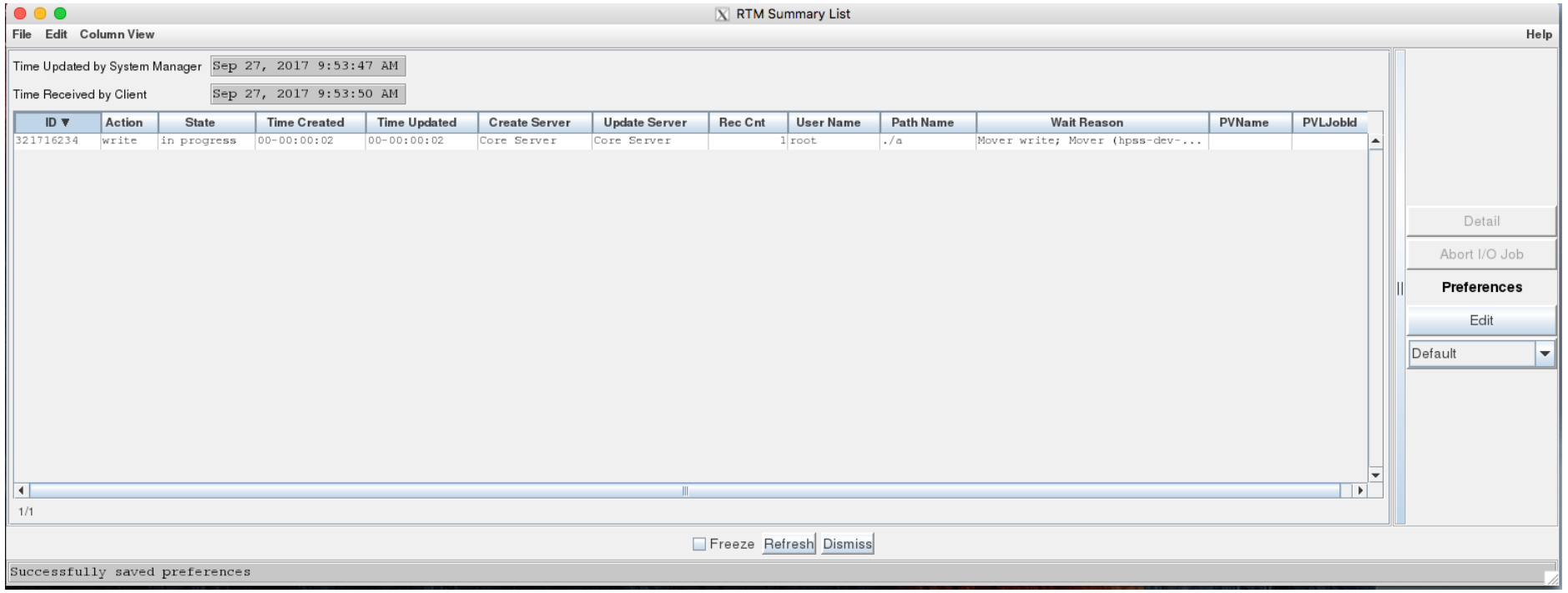
- Syslog-based HPSS logging (CR 375)
  - Take advantage of the capabilities of syslog for HPSS logging
  - See CR 375 HUF presentation for details

  Aug 7 16:24:13.895249 hpss_rait_engine_tcp(HPSS)[21596]::ALARM/MINOR # RAIT Engine (mystic)@mystic@clearlake.ibm.com # RemoteInterfaceSync # -5000 # RAIT1009 # 12345 -- Failed to sync with remote process

- Real-time Disk Ownership (CR 310)
  - Documented procedures for Mover failover

7.5.2 Planned Content

- I/O Abort (CR 183) / Job Control (CR 280)
  - Allow an I/O (read, write, stage, migrate) to be cancelled via SSM (GUI or ADM)
  - Additional information about what a job is doing in the RTM screen
  - hpssadm now supports retrieving rtm detailed information



hpssadm> rtm cancel -id 321716240

Cancelled IO Job 321716240

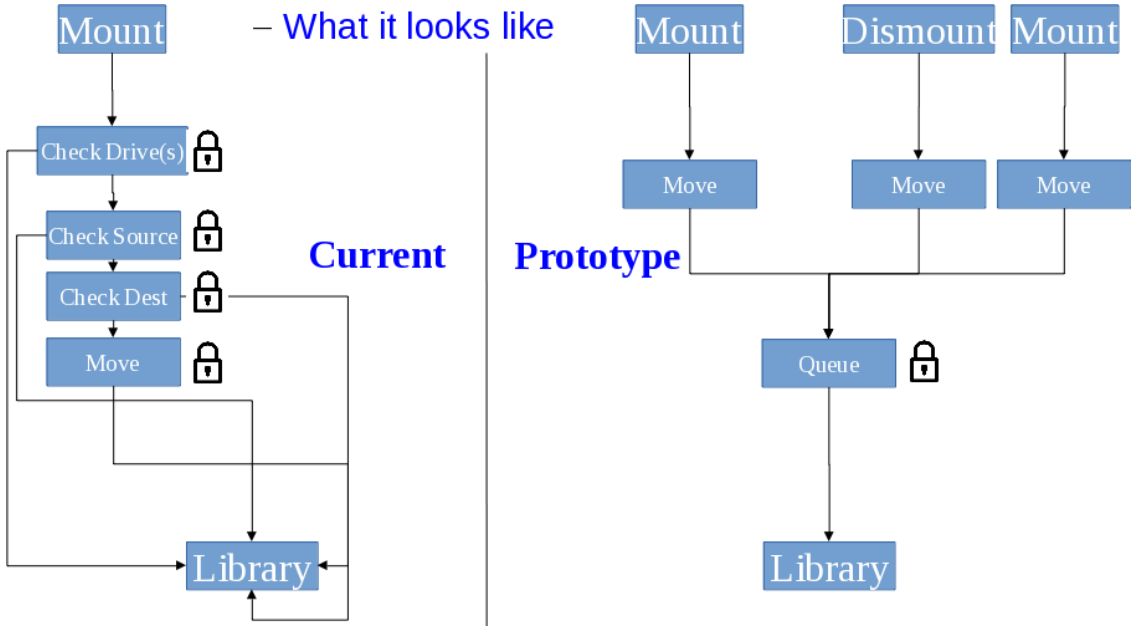
7.5.2 Planned Content

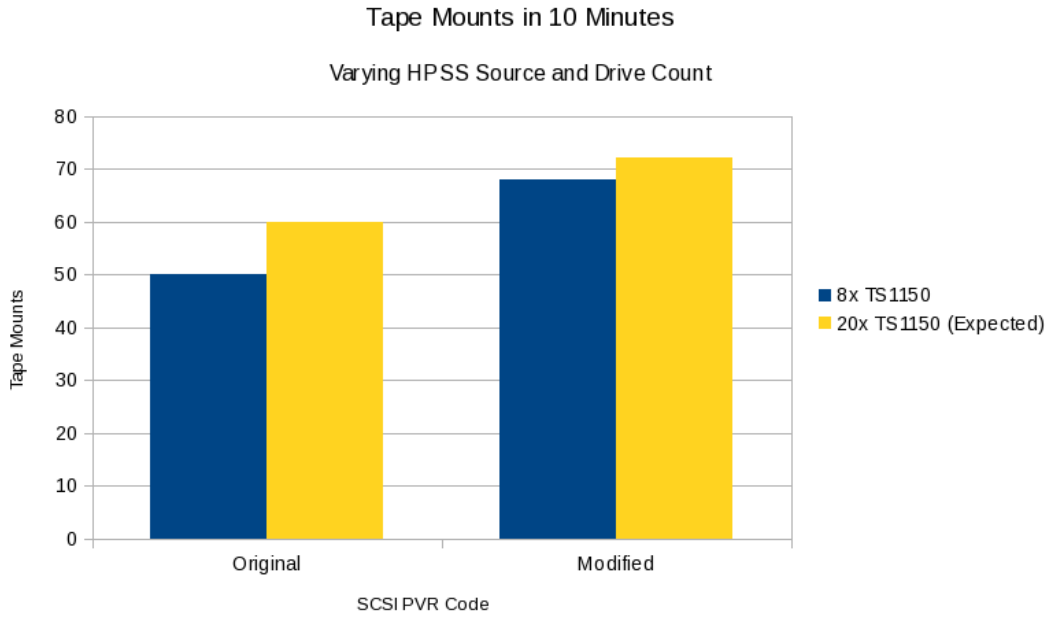
- Log stage performance metrics (CR 320)
  - Much more than stage – INFO logs for reads, writes, stages, migrates
  - Including disk and tape volumes and drive information
  - See CR 375 HUF presentation for details

Stage End : FilesetRoot.2839 : /users/foo/file1 : BFID x'00000001070000000601E713D3F82064BC24A8' : read {D11:D10000, D21:D10001} : wrote {T30:A00001, T31:A00002}

7.5.2 Planned Content

- SCSI PVR Performance (CR 303 / CR 438)
  - Modifications to the mount/dismount code increase performance by 10-20%
  - Able to meet hardware rates
  - Currently a “beta” version
  - Will replace the current SCSI PVR with next release (7.5.3)

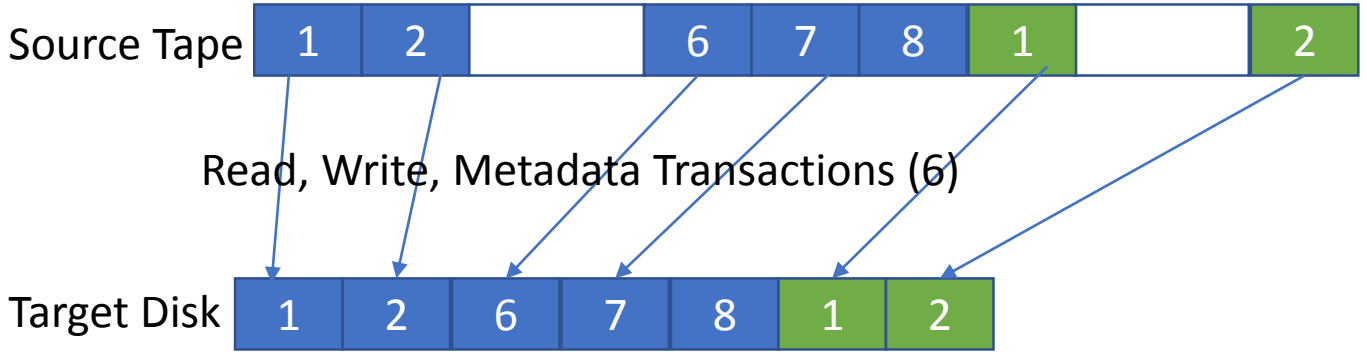




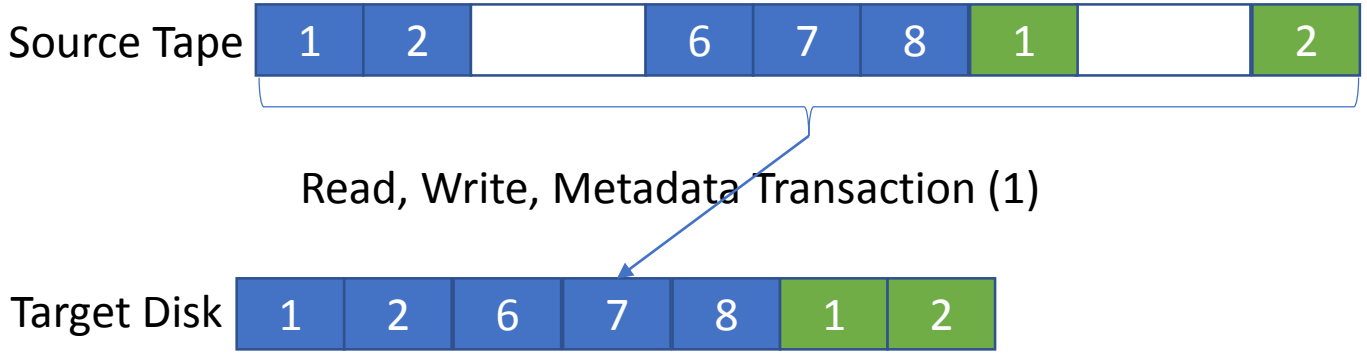
7.5.2 Planned Content

- Full Aggregate Recall (CR 270)
  - COS flag to allow entire aggregate to be read to disk in a minimal number of I/Os
  - Improved aggregate recall performance

7.5.1, reading an aggregate:



7.5.2, reading an aggregate with FAR:

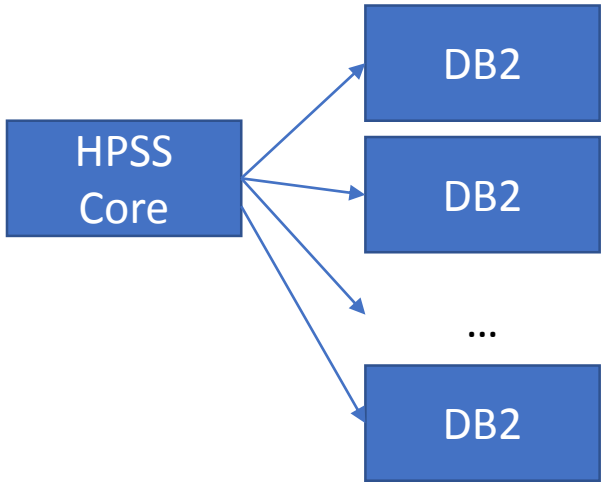


7.5.2 Planned Content

- Ordered Migration (CR 376)
  - Group items within aggregate based upon locality in the HPSS namespace
    - Same parent directory
  - Two flavors:
    - One aggregate = one directory (Read-optimized)
    - One aggregate = multiple directories grouped together (Write-optimized)
- Tape Drive Quotas for Recall (CR 383)
  - Set a number of drives to be made available for Recall
  - Reduce ability of Recalls to encroach on migration resources

7.5.2 Planned Content

- Run DB2 on non-Core System (CR 386)
  - Option to off-load DB2 resources to additional machines
  - Allow us to scale HPSS services up to take advantage of a full single node
  - In 7.5.1, the resources usually work out to 20% HPSS, 80% DB2



7.5.2 Misc. Changes

- Additional queries and metadata operations are partition aware (Performance)
- Client API can now handle more than 4096 open files per instance
- TOR “Queue Filling” optimization (Performance)
- Quaid will stop reporting status for files already staged
- Quaid can now stage a much larger list of files (was 10,000 per instance)
- Support POSIX O_EXCL flag
- hpsssum can now open files without staging them
- Provide a utility for human readable tape scheduler output (tor-summary)