



HPSS Release 4.5 and Release 5.1

Danny Teaff
June 18, 2002

Overview

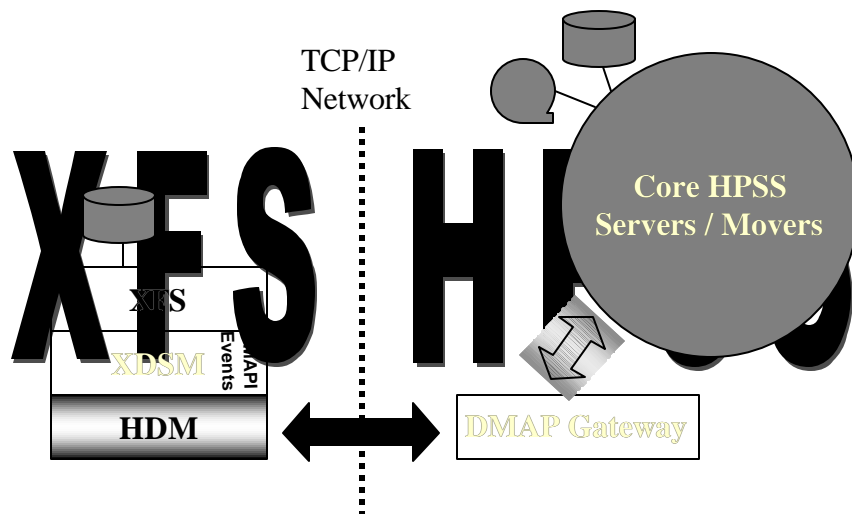
- ✓ Available June, 2002
- ✓ Contents Summary
 - **DMAPI support for XFS.**
 - **Duplicate tape copy support for Tape Hierarchies.**
 - **SCSI LBA support.**
 - **Linux Mover support.**
 - **RAIDZONE OpenNAS support.**
 - **New Device Types.**
 - **NFS performance enhancements.**
 - **RAIT Mirroring support (pending STK release).**

Overview

- ✓ Prerequisites:
 - AIX 5.1 (or 4.3.3) or Solaris 8.
 - RedHat Linux 7.2 for Linux Mover / Client support.
 - DCE 3.2 with PTF set 1.
 - DFS 3.1 with PTF set 4 (if DFS).
 - TXSeries 5.0 for AIX or TX Series 4.3 with PTF set 4 for Solaris.
 - Sammi 4.6.3.5.5 for AIX or 4.7 for Solaris.
 - Java 1.3
- ✓ Conversion from prior releases
 - Convert DMG filesets file and PVL drives metadata.
 - hpss_convert_4v5 provided for conversion from HPSS 4.2 or 4.3.
 - HPSS 4.1.1 customers must first convert to 4.2.

Page 3
© IBM Global Services, 2002

XFS Support



Page 4
© IBM Global Services, 2002

XFS Support

- ✓ Implement an efficient, transparent interface for users to access their HPSS data.
- ✓ Strengths
 - Leverages existing HDM "archive" architecture.
 - XFS available for Linux (no licensing fee).
- ✓ Key requirements
 - Support HPSS access via XFS using DMAPI.
 - XFS / HPSS filesystems shall be accessible via NFS for transparent access.
 - Support archived fileset concept (rename / delete).
 - Support on Redhat Linux.

XFS Support

- Provide migration and purge of data from XFS based on policy.
- Stage data from XFS when data has been purged from XFS
- Support whole file migration.
- Support utilities for the following:
 - Create XFS filesystem / fileset metadata in HPSS (create_fsys, create_fset).
 - Delete XFS fileset metadata in HPSS (delete_fset).
 - Delete all files from HPSS side of XFS fileset (archdelall).
 - Remove "deleted" files older than a specified age from HPSS side (archivedel).
 - Recover files deleted from XFS filesets not yet deleted from HPSS (archiverec).
 - Query DMAP Gateway for stats and list of filesets (dmgetattr).
 - Query DMAP Gateway for attributes of a fileset (getfileset).

XFS Support

- Perform HDM administration (hdm_admin).
- Get DM attributes, regions, and extents of a file (getdmattr).
- Set DM attributes (setdmattr).
- Delete DM attributes (deldmattr).
- Retrieve information from XFS on active sessions and number of tokens (prtsess).
- List the HPSS names of all files in an archived XFS/HPSS fileset (archivelist).
- List the XFS names of all files in an XFS/HPSS fileset (archivedump).
- Compare archive dumps from HPSS and XFS and report inconsistencies (archivecmp).

XFS Support

✓ XFS Pre-Configuration Steps

- **Install XFS kernel patch and user filesets.**
 - Obtain 'vanilla' 2.4.18 Linux kernel (<http://www.kernel.org>).
 - Install XFS 1.1 kernel patch on vanilla 2.4.18 kernel (<http://oss.sgi.com/projects/xfs/>).
 - Apply XFS/HPSS kernel patch (available from HPSS website).
 - Build and boot new kernel.
- **Create DMAPI Device.**
 - XDSM for XFS requires a special device for communication between the kernel and userspace portions of the code. This device must be created using the command below:

```
% mknod /proc/fs/xfs_dmap_i c 10 140
```

XFS Support

- ✓ Configure HDM
 - **Configure HDM files.**
 - There are 5 file types used for configuration of the HDM:
 - **config.dat:** Configures HDM processes.
 - **policy.dat:** Configures migration and purge policies.
 - **gateways.dat:** Gateways allowed to contact HDM.
 - **security.dat:** (DFS Only) Configures remote cell user ID translation.
 - **filesys.dat:** Tracks aggregates and filesets.
 - The normal location for these configuration files is
`/var/hpss/hdm/hdm1`

XFS Support

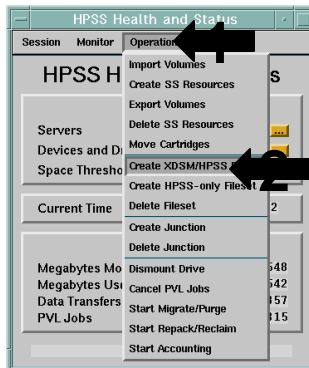
- ✓ Steps to create an HPSS/XFS fileset:
 - **Create a XFS filesystem.**
 - **Create an HDM filesystem entry.**
 - **Create an HPSS fileset.**
 - **Mount the file system.**
 - **Set permissions and owner for the XFS filesystem.**

XFS Support

- ✓ Create an XFS filesystem
 - Create a disk partition, using **fdisk**, **cfdisk**, or **sfdisk**.
For example: `% cfdisk /dev/hdb`
 - Be sure to create a partition of type 83 (Linux).
- ✓ Create an HDM filesystem entry (**create_fsys**).
Syntax: `create_fsys XFS <host> <port> <fname> <path> <media> <option> <queue_size> <migrate> <purge> [key]`
For example: `% create_fsys XFS sp3n01.clearlake.ibm.com 6002 my_fileset /mnt/my_fileset "ide0(3,72)" archive/rename 10000 run wait 123abc`
Note: *This command should be executed on a machine declared as a gateway in gateways.dat. The command sends the request to HDM, which then adds the entry to the filesys.dat file.*

XFS Support

- ✓ Create an HPSS fileset
 - From SSM Health & Status, select "Operations" and then "Create XDSM/HPSS Fileset".
 - 'XDSM' is used here as a generic term to refer to either XFS or DFS managed filesets.



XFS Support

- Fill in the information for your fileset & click "Full Create".

The screenshot shows a window titled "Create XDSM/HPSS Fileset" with the following fields and values:

- Managed Filesystem Type: **XFS**
- Fileset ID: **1079914467 ,, 1022262844**
- Filesystem ID: (empty)
- Filesystem Name: **test_ren**
- HPSS/DMAP TCP Port: **6002**
- HPSS/DMAP TCP Hostname: **sally**
- User ID: **0**
- Group ID: **0**
- Global Mount Point: **/mnt/test_ren**
- Local Mount Point: (empty)
- DMAP Gateway: **DMAP Gateway**
- File Family: (empty)
- Class of Service: (empty)
- Mount Point Name Server: **Name Server 1**

On the right, there is a "Permissions" section with a grid of checkboxes for User, Group, and Other, with columns for Read (r), Write (w), and Execute (x). The "Full Create" button is highlighted with a red arrow and a black circle.

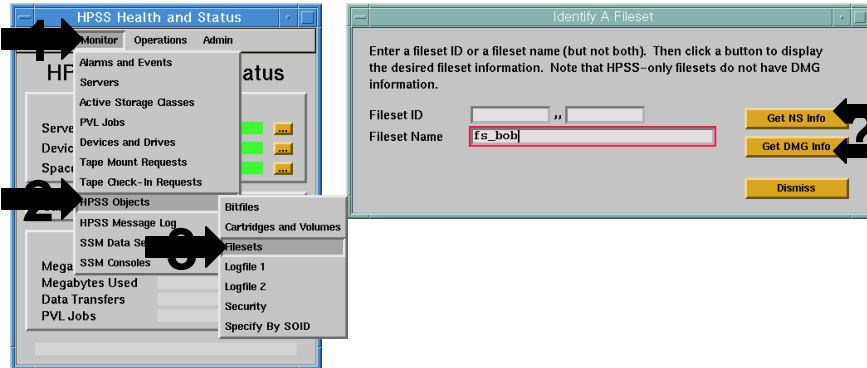
© IBM Global Services, 2002

XFS Support

- ✓ Mount the file system
 - Use the 'mount' command with a few extra options to mount your filesystem:
Syntax: `mount -t xfs -o dmapi <partition> <mountpoint>`
For example: `% mount -t xfs -o dmapi /dev/hdb7 /mnt/test_ren`
- ✓ Set permissions and owner for the XFS filesystem
 - Now that the filesystem is mounted, use the 'chown' and 'chmod' commands to update the permissions on the filesystem's root directory.
For example:
`% cd /mnt`
`% chown hpss:hpss test_ren`
`% chmod 755 test_ren`

XFS Support

- ✓ From SSM Health & Status, select "Monitor", "HPSS Objects", and then "Filesets"



Page 15
© IBM Global Services, 2002

Duplicate Tape Copy Support

- ✓ New migration options supported for tape:
 - **Migrate Volumes**
 - Default option. Same as option supported in the past.
 - Selects virtual volumes that are EOM with the largest free space for migration.
 - Segments are moved laterally to volume in same Storage Class for active segments or down the hierarchy for inactive segments.
 - Useful for managing space without running repack, but does not support multiple copies.
 - **Migrate Volumes and Whole Files**
 - Similar to Migrate Volumes option.
 - Also segments for the file on other tape volumes are migrated.
 - Useful for getting segments of a file on the same volume, but can result in large number of tape mounts.

Page 16
© IBM Global Services, 2002

Duplicate Tape Copy Support

- ✓ Migrate Files
 - File versus volume based migration (similar to disk migration).
 - Individual files selected based on last write times and migration policy.
 - Files sorted by source tape volume prior to migration.
 - Files on source level are not purged.
 - Useful for maintaining a second copy of tape files. Source Storage Class must be manually repacked.
- ✓ Migrate Files and Purge
 - Similar to Migrate Files option.
 - Migrated files are purged from source level as soon as migrated.
 - Read and write times used to select candidate files.
 - Useful for moving inactive files into deep tape archive. Source Storage Class must be manually repacked.

Page 17
© IBM Global Services, 2002

Duplicate Tape Copy Support

The screenshot displays the 'Migration Policy' configuration window, divided into two main sections: 'Basic Policy' and 'Storage Subsystem-Specific Policy'.
Basic Policy:
- Policy ID: [text box]
- Policy Name: Disk Migration (1 thread)
- Last Read Interval: 60 minutes (tape only)
- Last Update Interval: 30 minutes
- Free Space Target: 100 percent
- Request Count: 1
- Runtime Interval: 120 minutes
- Disk Options: Migrate At Warning Threshold, Migrate At Critical Threshold
- Tape Options: Migrate Volumes, Migrate Volumes and Whole Files, Migrate Files, Migrate Files and Purge
- Buttons: Load Existing, Start New
Storage Subsystem-Specific Policy:
- Storage Subsystem: [text box]
- Last Read Interval: [text box] minutes (tape only)
- Last Update Interval: [text box] minutes
- Free Space Target: [text box] percent
- Request Count: [text box]
- Runtime Interval: [text box] minutes
- Disk Options: Migrate At Warning Threshold, Migrate At Critical Threshold
- Tape Options: Migrate Volumes, Migrate Volumes and Whole Files, Migrate Files, Migrate Files and Purge
- Buttons: Load Existing, Start New
At the bottom of the window are buttons for 'Add Basic', 'Delete Basic', 'Update Basic', and 'Dismiss'. A red bracket on the right side of the 'Storage Subsystem-Specific Policy' section is pointed to by a black arrow.

Page 16
© IBM Global Services, 2002

SCSI-2 LBA Positioning

Currently absolute address used to position before 1st byte of a file. Relative addresses are then used to position to a requested tape segment.

- ✓ LBA positioning supported on devices supporting SCSI-2 Logical Block Addresses (LBAs) and the LOCATE command.
- ✓ Option selectable from device / drive configuration window.
- ✓ Destination LBAs calculated based on known LBAs and relative addresses.
- ✓ LOCATE command to position directly to the requested location
- ✓ Provides for faster positioning.

Page 19
© IBM Global Services, 2002

SCSI LBA Positioning

The screenshot shows a window titled "Device And Drive Configuration" with a sub-header "Mover Device and PVL Drive Configuration". The configuration fields are as follows:

- Device/Drive ID: 13
- Device/Drive Type: 9840 Tape
- Mover: Non-DCE Mover (fred)
- PVR: STK PVR
- Controller ID: 13
- Polling Interval: -1 seconds (make negative to disable)
- Device Name: /dev/rmt0
- Drive Address: 0,0,0,1

The "Device Flags" section contains the following options:

- Read Enabled
- Write Enabled
- Removable Media Support
- Locate Support
- NO-DELAY Support
- Write TM(0) to Sync
- Multiple Mover Tasks
- SCSI-2 LBA Positioning (indicated by a black arrow)

Buttons at the bottom include "Start New", "Add", "Delete", "Update", and "Dismiss".

Page 20
© IBM Global Services, 2002

Linux Mover Support

- ✓ Support Mover on Redhat Linux 7.2.
- ✓ Non-DCE Mover is used.
- ✓ System test performed for disk and LTO drives.
Other tape support pending availability of drivers / hardware.

RAIDZONE OpenNAS

- ✓ Support XFS on OpenNAS server through DMAPi interface. HDM installed on RAIDZONE node.
- ✓ Support HPSS Linux Mover on OpenNAS server.
- ✓ RAIDZONE OpenNAS provides:
 - **Switched node disk array technology.**
 - **Runs Linux Kernel and supports XFS.**
 - **Accommodates 5 to 15 hot swappable, 160 GB drives.**
 - **Supports redundancy with rapid recovery and redundant boot partitions.**

New Device Types

- ✓ Provide support for the following devices:
 - **IBM 3590H.**
 - **STK 9940B.**

NFS Performance Enhancements

- ✓ Removed checksum code (conditional compilation).
- ✓ Internal read retries. Hold reads for a short time in case data arrives from HPSS sooner than the client's timeout value.
- ✓ Support read-ahead. Prefetch next block in large files so data is in the datacache when accessed by client
 - (Read next block if actual data in the next block, the current block is uncached, or this is the last block cached for the file and haven't tried prefetch)**
- ✓ Allow data cache to perform concurrent IO streams for a single file.

RAIT Mirroring

- ✓ Support RAIT Mirroring for STK 9840 and 9940.
- ✓ Feature pending formal release from STK.
- ✓ Configuration:
 - **Configure RAIT enabled PVR.**
 - PVR subtype and drive type for 'STK RAIT'.
 - 'Physical drive count' specified.
 - **Configure RAIT drive.**
 - STK -> '9840 RAIT' / '9940 RAIT' drive type.
 - **Create Storage Classes for each RAIT volume type.**
 - STK -> '9840 RAIT' / '9940 RAIT' media type -> 1+1 (data and parity volume count).
 - **Import RAIT volumes as a specific RAIT volume type.**
 - STK -> '9840 RAIT' / '9940 RAIT' media type -> 1+1.
 - Data and parity volume counts.
 - **Create VVs for each imported RAIT volume.**

Page 25
© IBM Global Services, 2002

Release 5.1

- ✓ Target release date: 3/31/2003
- ✓ Major enhancements:
 - **New HPSS administrative graphical interface.**
 - Implemented as Java Client.
 - Demonstration to be provided by Deryl Steinert and Vicky White.
 - **Infrastructure Replacement:**
 - Replace SFS. Utilize DB2 database.
 - Eliminate need for Encina.
 - Reduce reliance on DCE.
 - Merge Name Server, Bitfile Server, Disk and Tape Storage Server, & part of Client API.
 - Eliminate nested and distributed transactions
 - Performance improvements expected.
 - Details to be provided by Dave Fisher.

Page 26
© IBM Global Services, 2002

Release 5.1

✓ Prerequisites:

- **AIX 5.1 or Solaris 8.**
- **RedHat Linux 7.2 for Linux Mover / Client support.**
- **DCE 3.2.**
- **DB2 7.2.**
- **Java 1.3**

Prerequisites are subject to change. Release numbers reflect a minimum level.