# High Performance Storage System `HPSS`

**IBM**®

### *Hierarchical Disk and Tape File System*

High Performance Storage System (HPSS) is hierarchical file system software designed to manage and access many petabytes of data at high data rates. HPSS manages the life cycle of data by moving inactive data to tape and retrieving it the next time it is referenced. HPSS can either present its own disk file system directly to the user, or it can provide a tape pool to GPFS.

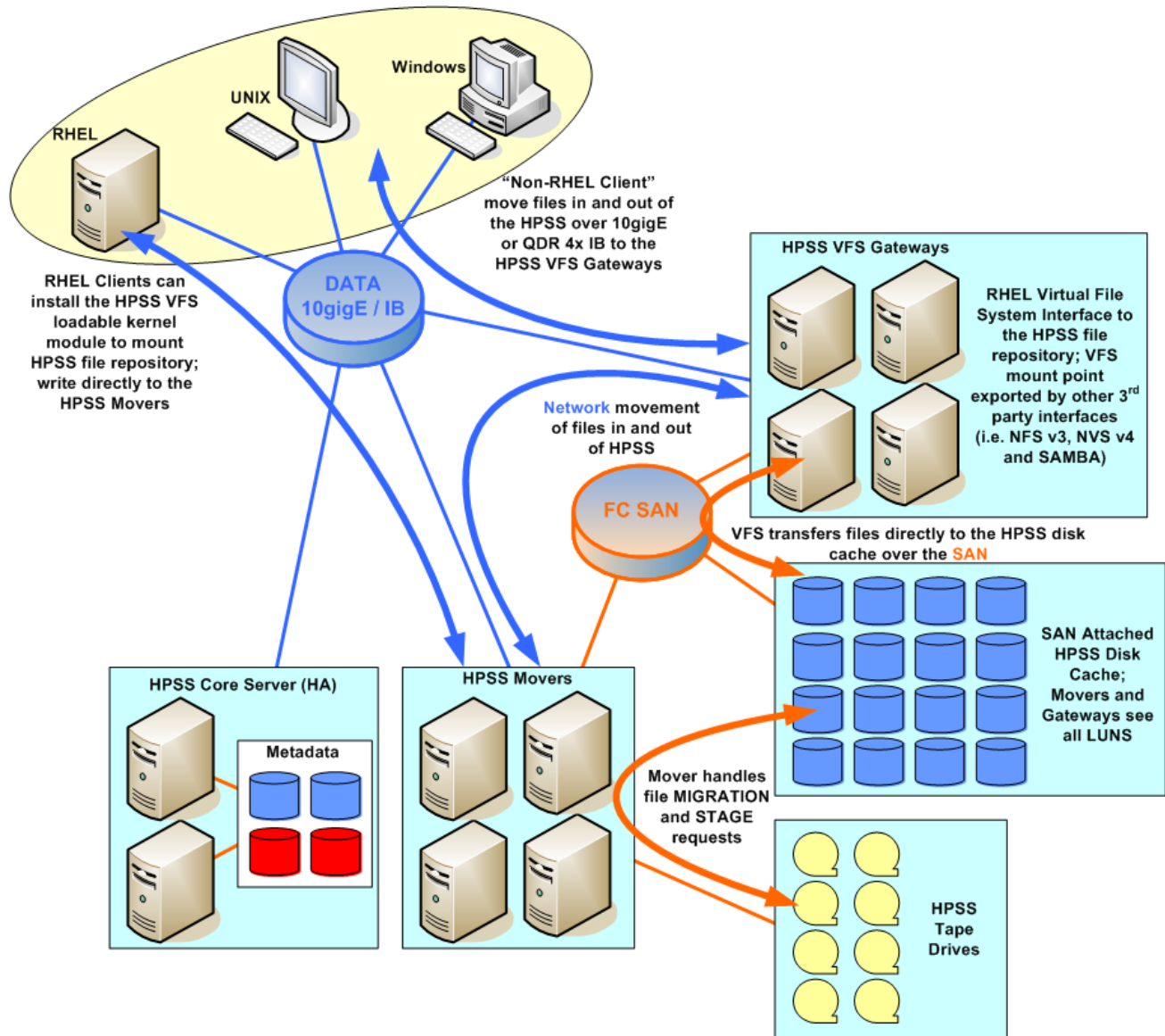### *IBM DB2®- protected Metadata*

HPSS uses IBM DB2 as the metadata library for storing the identity, ownership, location and status of all files and devices. DB2 is the gold standard of database software for reliability, availability, and scalability.

### *Cluster Architecture for Extreme Scalability*

HPSS has a cluster design that combines the power of multiple computer nodes into a single, integrated storage system. By increasing the size of the cluster and by adding disks and tape libraries, HPSS is capable of 100s of millions of files, 10s of petabytes, and data rates of gigabytes per second. No matter how large, the storage system always appears to its clients as a single storage service with a unified common name space.

### *Migration Services*

IBM Global Business Services offers a service to migrate SAM-FS and DXUL users to HPSS. IBM has experience migrating large tape libraries without physically copying the tapes and without continuing to run the legacy software.

## High Data Rates

HPSS enables disks and tapes to be striped to create files that can be accessed at high data rates through parallel I/O operations. With 16-way striping, single file disk data rates of over two gigabytes per second have been achieved.

A single instance of HPSS is capable of concurrently accessing hundreds of tapes for extremely high aggregate data transfers.

HPSS provides aggregation of small files to tape, which increases performance and minimizes the tape space used for small files.

## System and Data Integrity

HPSS **High Availability** is built upon Red Hat Linux cluster services. HPSS supports failover redundancy for HPSS core servers, movers, and gateway nodes.

HPSS provides data integrity checksum utilities that allow a user to make a checksum of file content and place the result in a **User Defined Attribute** of the file for later comparison when the file is retrieved. User-Defined Attributes are stored in and protected by the DB2 metadata server.

## Interfaces

**HPSS API.** HPSS provides a Client API that is a similar to a POSIX standard read-write interface with additional syntax for supporting striping and classes of service. This Client API is the most powerful interface in terms of control and rich functionality. It is the foundation for all other HPSS interfaces.

**HPSS VFS interface for Linux.** Red Hat Linux applications benefit from a true POSIX standard read-write file system interface. This enables HPSS to be mounted as a Linux file system. This interface enables many standard commercial programs that include file I/O to use HPSS as file space, making them into hierarchical disk-tape applications. Indiana University, for example, uses the VFS interface to run Samba, NFS, Secure FTP, and Apache directly on HPSS file space.

**GPFS – HPSS Interface.** HPSS provides hierarchical space management and disaster recovery backup for GPFS. For space management, the GPFS Information Lifecycle Management (ILM) function uses HPSS as a lower tier of disk and/or tape to which files can be migrated when they are inactive, thus freeing valuable GPFS tier 1 disk space. HPSS furthermore uses the migrated files as a source for file system recovery, should it be necessary to rebuild GPFS. This capability saves space on tapes, compared with systems that keep separate copies of files for backup and space management.

## Equipment Supported

The full suite of HPSS software runs on IBM's System p computers with AIX and on IBM System x and System p computers with Red Hat Enterprise Linux.

HPSS supports most disk systems from IBM, DDN, and LSI Logic, and solid state disk systems from Texas Memory Systems. Tape libraries from IBM, Oracle StorageTek, Spectra Logic, QualStar, and Quantum are supported, as are most current IBM, HP, and StorageTek tape drives.

## How HPSS is Offered

HPSS is developed by an ongoing collaboration of five national laboratories of the United States Department of Energy and IBM. This collaboration has been working together since 1992, through seven major releases, and is now in the process of developing the ultimately scalable HPSS 8,

HPSS is licensed and supported by IBM under an agreement between IBM and the Department of Energy. HPSS is licensed, installed, and supported as a service offering of IBM Global Business Services. IBM also offers system engineering services, custom feature development services, and migration services. For U. S. Government organizations, HPSS is listed on NASA's SEWP government-wide acquisition web site, www.sewp.gov.