

StorHouse for Digital Preservation

Ensuring the reliability, longevity, scalability, recoverability, and integrity of digital archives

StorHouse Benefits for Digital Preservation

Government agencies, private enterprises, and scientific, academic, and cultural organizations can entrust their digital content to StorHouse. Benefits of the StorHouse digital preservation solution include:

- **Active business continuance** to ensure the survivability of data by retaining multiple copies in geographically dispersed locations, in multiple devices, and/or on different media through automated backups, duplexing, and replication
- **Storage vendor independence** to incorporate hardware and media choices based on long-term price performance requirements rather than what's available from a particular vendor
- **Scalability to multiple petabytes** with no performance degradation to accommodate growing volumes of digital data and expanding user bases
- **Automated media migration and refreshment** to leverage long-term commodity-based storage and to protect against media degradation and obsolescence
- **Cross-platform, shared access to the digital archive** to provide a consolidated collection of an organization's digital heritage, eliminating duplicate data stores and improving information flow
- **Data integrity and error checks** to ensure authenticity from the time data becomes a record and throughout its life
- **Comprehensive metadata histories and redundant copies** to provide an audit trail of actions performed on a record

The preservation of digital data is more than just storing data long-term. It's about protecting archived data amidst technological advancements, managing escalating data volumes created in a greater variety of formats, and ensuring data authenticity and accessibility over time. This complex environment demands an infrastructure that can continually evolve to integrate new technology, support new formats and types of digital content, respond to the requirements of heterogeneous user communities, and scale from terabytes to petabytes without degrading performance.

How can digital custodians ensure long-term data integrity, accessibility, and persistence while balancing storage costs, emerging technologies, and the rapid rise in the amount, complexity, and demand for digital data?

StorHouse Digital Preservation Solution

The FileTek StorHouse® digital preservation solution combines industry-leading, scalable storage devices and Open System processors with specialized storage management, relational database management, and file system interface software components to manage and preserve massive amounts of digital data. StorHouse technology promotes data longevity by uniquely virtualizing a managed hierarchy of vendor-independent storage devices to present a single view of storage to client applications. The hierarchy can scale without service disruption to incorporate new devices and media types as they become available and can migrate existing data to them automatically.

In addition to virtualizing storage, the StorHouse digital preservation solution supports a consistent and familiar file system interface to the archive. The StorHouse virtual file system handles billions of records and looks just like another network drive, mount point, or share point on the client system. To access records, applications do not need to know where records reside in the archive or how to interface with proprietary APIs or raw storage devices. An authorized request to access a record requires knowledge of only the record ID. Figure 1 illustrates the virtual file system and storage virtualization features of the StorHouse digital preservation solution.

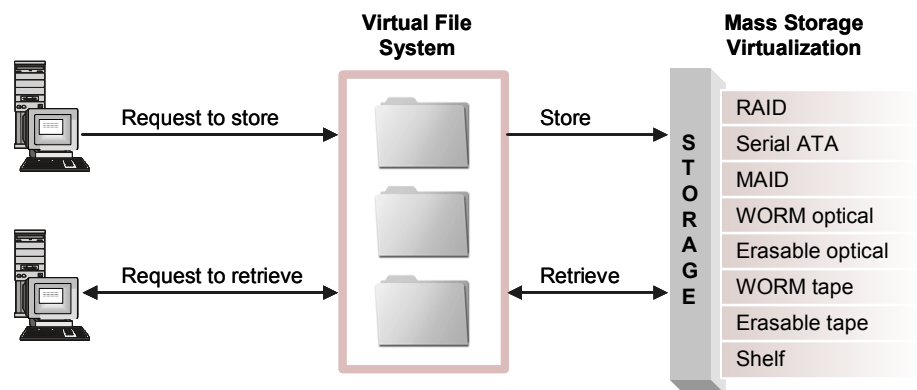


Figure 1 – StorHouse Storage Virtualization and Virtual File System Interface

StorHouse Software

StorHouse/RFS is the virtual file system manager of a StorHouse digital archive. This software manages the flow of data into and out of the archive. Applications read and write records through the virtual file system using standard NTFS, NFS, or CIFS protocol. Authorized users can browse designated collections through the virtual file system using standard operating system tools, such as Windows Explorer or command line tools in UNIX and Linux.

StorHouse/SM is the storage manager of a StorHouse digital archive. This software manages archived data on the storage hierarchy, which may include a mix of the devices listed in Figure 1. It is also responsible for data backup, recovery, migration, replication, and retention.

StorHouse/RM is the relational database manager of a StorHouse digital archive. This software stores, accesses, and manages databases containing file locator data, security information, and other metadata for all archived items. These databases automatically scale to billions of rows as the archive grows.

StorHouse/Control Center is the administrative manager of a StorHouse digital archive. It consists of a graphical user interface for managing the StorHouse product line.



Corporate Headquarters:

FileTek, Inc.

9400 Key West Avenue

Rockville, MD 20850

Phone: 301.251.0600

info@filetek.com

www.filetek.com

International Headquarters:

FileTek Ltd

1 Northumberland Ave.

London WC2N 5BW

Phone: +44 (0) 207.872.5583

intsales@filetek.com

StorHouse Digital Preservation Functions

StorHouse supports the key service functions of a persistent digital archive: ingest, archival storage, data management, access, and administration.

Ingest. An application requests to archive a record to StorHouse by writing it to a directory in the virtual file system. Each record enters the virtual file system with its unique ID. Each directory has assigned policies that all records in the directory inherit, for example, the storage media on primary and mirror systems, the number of copies and geographical locations, and the specifications for retaining, updating, and deleting records. The virtual file system manager aggregates records and creates file locator data to map record IDs to the physical record and to the location of the record within the archive. Additionally, metadata such as security information and standard file system permissions are captured.

Archival Storage. The virtual file system manager passes the aggregated records and metadata to the storage manager with instructions for final archiving. The storage manager writes the records to appropriate media and loads the metadata into an embedded relational database. This is done on both the primary location and at a secondary, or mirror, location if implemented.

Data Management. The storage manager manages the storage hierarchy and ensures all levels of protection for the archived records. This includes migrating records to new technologies, creating backup and archive copies, performing integrity checks, recovering or relocating records, refreshing or retiring media, and monitoring devices. The relational database manager ensures the maintenance, access, and protection of the metadata that describes the archive's holdings. Automated recovery and journaling capabilities ensure all records along with their associated metadata are continuously protected.

Access. Like traditional file systems, applications request a record by supplying the record ID (or file name) to the virtual file system. Designated directories may enable user browsing of archived data, while others may disallow this feature. The virtual file system manager locates, requests, and delivers data through the virtual file system. Security and access control mechanisms associated with archived records are checked within the relational database before the data is delivered.

Administration. Functions of the StorHouse digital archive are administered, controlled, and monitored through a central administrative manager and automated utilities. Comprehensive logging, statistics, error reporting, and alert notifications are generated for each function of a StorHouse digital archive.

For More Information

For more information about the digital preservation solution from FileTek, contact a FileTek sales representative, e-mail info@filetek.com, or visit www.filetek.com.

© 2006 FileTek, Inc. All rights reserved. FileTek and StorHouse are U.S. registered trademarks of FileTek, Inc. Other trademarks included herein are the property of their respective owners. The following U.S. patents protect StorHouse: 4,864,572; 5,247,660; 5,727,197; and 6,049,804.