

The National Archives of the United Kingdom

Harnessing Technology to Preserve a Nation's Digital Content

Case Study

Summary

In 1999, the British government launched an initiative to modernize its departments and streamline business processes by taking advantage of new and emerging technologies. A major part of that initiative mandated the electronic storage and retrieval of all new records by 2004. The National Archives of the United Kingdom knew that it would be responsible for preparing the long-term storage for many of those diverse electronic records, which range from e-mail and their attachments to audio files. The immediate challenge the National Archives faced was how to accomplish this project.

To help meet this challenge, the National Archives turned to FileTek's StorHouse®/RFS relational file storage system software. Integrating StorHouse/RFS with its own data collection and searching system produced by Tessella Support Services plc, the National Archives built a digital data repository sized to initially hold 2 terabytes of records. It is expected that over time, StorHouse/RFS will eventually manage priceless digital records of 100 terabytes or more.

Now the mere mention of a nation's historical archives will no longer only elicit visions of shelves packed with meticulously catalogued papers that preserve a national heritage. The National Archives has changed this image forever by preserving its country's digital history with state-of-the-art technology. It has created one of the world's most unique data repositories. And on March 31, 2003, at 2 p.m., the National Archives released its digital information-rich store to the British public.

Organization Background

Located on the banks of the River Thames in West London, the National Archives is the government department responsible for the safekeeping of state and central legal records. The records range from the Domesday Book (a census commissioned in 1085 by William the Conqueror) to current government papers. Also, as the primary source for locating British historical records, the National Archives administers and indexes over 43,000 record catalogues, containing 150,000 references and 300,000 sub-references.

Information is made available to the public after a designated waiting period, depending on a particular record's classification. Anyone can visit the National Archives reading room and actually see and handle many of the historic items. Certain information is also available online via the Internet.

The Challenge

As a result of the modernizing initiative, the National Archives created digital record standards and helped other government departments implement a records creation and saving process. This proved critical because the National Archives knew that after screening and selecting certain of these digital records, they would be responsible to archive them in perpetuity. "We closely analyzed the best way to meet the new guidelines to comprehend the initiative's full impact on our archiving processes," said David Ryan, Head of Digital Preservation at the National Archives.

The National Archives then began to focus on its main objective: to build a digital repository for the impending influx of government electronic records—such as video clips, website snapshots, and common office-style documents. In March 2002, it requested bids from technology companies to build such an archive while adhering to its specialized provisions. “We had a requirement for a modular configuration so that we could easily change components as technology advances over the coming years,” said David Glover, Systems Manager of Digital Preservation at the National Archives. “The system had to provide absolute secure access to the digital archive. And most important, the archive had to be extremely scalable for future needs.”

The Solution

After close review of various vendor proposals, the National Archives chose Tessella Support Services plc to design and implement the digital archive and FileTek, Inc. to provide its software to manage the storage repository. “We picked FileTek’s StorHouse/RFS for the critical job of managing and accessing stored records because of its reliability, ease of use, support for cost-effective storage media, and ability to scale to 100s of terabytes of records and beyond,” said Ryan.

StorHouse/RFS is built to solve the problems associated with archiving and accessing large quantities of fixed content data, such as those handled by the National Archives. It employs economical multi-level storage media (including high performance/high capacity tape cartridges) in a near-line architecture, allowing for direct record access. StorHouse/RFS provides reliability by managing multiple file copies and automatically prevents tape failures before they occur.

StorHouse/RFS’s scalability is due, in part, because the number of records archived is not limited by particular file system constraints. Instead, StorHouse/RFS uses an internal relational database application to track the location of archived files. “It is difficult to predict how large, in fact, this archive will grow,” said Glover, “but the National Archives must be prepared for any eventuality.”

StorHouse/RFS presents itself to the Tessella application as a virtual drive, i.e., as an NFS, NTFS, or CIFS mount or share point. The application reads and writes files to the virtual drive just as it would to any mapped Windows® or NFS drive. Hardware supporting the digital archive includes a Sun Microsystems™ 280R server and a StorageTek™ tape library with LTO drives with Sony Ultrium media. End-users request records using the Tessella application integrated with StorHouse/RFS to access the archive.

“The National Archives understands that it has built a vital digital records storage system, tomorrow’s equivalent of the traditional paper archive,” said Ryan. “Therefore we approached this project with the utmost level of care. We are making sure that priceless history written today in digital form is protected just as well as the 11th century paper-based Domesday Book.”

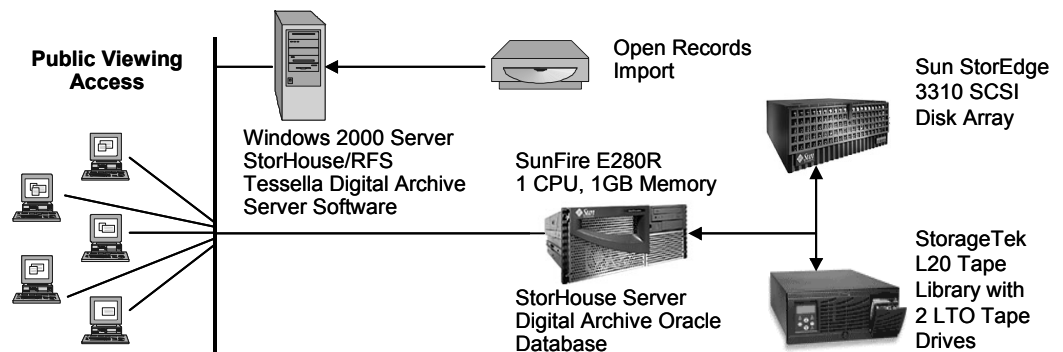
FileTek

Corporate Headquarters:

FileTek, Inc.
9400 Key West Avenue
Rockville, MD 20850
Phone: 301.251.0600
info@filetek.com
www.filetek.com

International Headquarters:

FileTek UK Limited
1 Northumberland Ave.
London WC2N 5BW
Phone: +44 (0) 207.872.5583
intsales@filetek.com
www.filetek.com



The National Archives' Digital Archive Solution

©2003-2005 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