CASE STUDY:
DIGITAL PRESERVATION AT THE
NATIONAL LIBRARY OF NEW ZEALAND
PRESERVATION: A FORWARD-LOOKING MISSION

Sun
microsystems

> "The purpose of the National Library is to enrich the cultural and economic life of New Zealand … by, as appropriate, collecting, preserving, and protecting documents, particularly those relating to New Zealand, and making them accessible for all the people of New Zealand, in a manner consistent with their status as documentary heritage and taonga; …."
> **(Section 7 of the National Library Act)**

> "The problem of preserving digital information for the future is not only, or even primarily, a problem of fine tuning a narrow set of technical variables. It is not a clearly defined problem…Rather, it is a grander problem of organizing ourselves over time and as a society to maneuver effectively in a digital landscape. It is a problem of building…the various systematic supports…that will enable us to tame the anxieties and move our cultural records naturally and confidently into the future."

> **Garrett, J. & Waters, D. (eds). (1996).** Preserving digital information: Report of the Task Force on Archiving of Digital Information. Washington, DC: Commission on Preservation and Access and the Research Libraries Group. http://www.oclc.org/programs/ourwork/past/digpresstudy/final-report.pdf.

With the explosion in the volume of digital assets over the past several decades, libraries and other memory institutions have become increasingly aware of the need to preserve such assets. The National Library of New Zealand, Te Puna Mātauranga o Aotearoa, has identified digital preservation as an important part of its mission and one that requires immediate action. Collaborating with Ex Libris™ Group, Sun, and a peer review group of industry experts, the National Library of New Zealand is blazing the trail in addressing the digital preservation challenge and establishing best practices in this sphere.

The National Library of New Zealand Act 2003 laid the foundations that allowed the National Library to spring into action in the digital preservation frontier. The Act provided the legislative mandate for the library to collect and preserve digital content in ways that will ensure current and future access to New Zealand's documentary heritage.

"The key objective of the Act is to make New Zealand's digital heritage available in perpetuity," says Penny Carnaby, national librarian and chief executive of the National Library of New Zealand. "How will we as a nation be able to look back in 50 or 100 years' time and know what was going on? We do not want to fail the future by leaving a gap in New Zealand's history because we were too slow or unwilling to confront the challenge of digital preservation now."

Following funding approval in 2004, the National Library of New Zealand library launched the National Digital Heritage Archive (NDHA) Programme to address the digital preservation challenge and put in place the organization, strategy, planning, and technology required to make the preservation of New Zealand's digital cultural heritage a reality.

"The strategic importance behind this initiative is a technology and business change which underpins all aspects of the library's responsibilities and services," adds Carnaby. "Both print information and electronic information are being produced at an extraordinary rate. However, electronic data now makes up the majority of all unique, original output, and the gap between the two is widening in favor of digital materials."

The NDHA will serve as the National Library's storehouse for digital materials such as Web sites, CDs, DVDs, images, and digitized copies of print and audiovisual assets that make up its digital heritage collection. The NDHA will preserve these items in their original form and ensure that actions are taken so that they can be viewed, listened to, and explored in the future, even if the original technology has become obsolete.

The National Library of New Zealand has become a Sun Center of Excellence, a premier showcase of the partnership between Sun and educational institutions. The NDHA will serve as an international model for the implementation of digital repositories and preservation management. Through its global reference architecture program, Sun will use the best practices arising from this collaboration to help additional memory institutions worldwide.

In Search of a Solution

Although digital preservation has been a topic of growing interest for national libraries and other memory institutions worldwide, there were few best-practice implementations for New Zealand to draw on when it came to putting in place a fully operational, large-scale preservation system. As a matter of fact, the National Library of New Zealand was one of those institutions that others turned to for advice on the

Java

preservation of digital materials, because the library had already been involved in initiatives such as the development of preservation metadata schemas, a technical metadata extractor, and curator-led Web harvesting.

The first order of business for the NDHA Programme team was to define operational guidelines for a preservation system, with the overarching requirement of securing the integrity and authenticity of digital material deposited at the National Library. Protecting and ensuring accessibility to these materials for perpetuity from the moment they came under library control was also required.

To achieve these goals, a number of key requirements were identified:

- Security: Objects must be stored and managed in a manner that provides the level of redundancy and security needed to ensure that the objects are protected from compromise, whether malicious or accidental.

- Scalability: The system must support the storage and management of millions of objects and scale up in a reliable and cost-effective manner as digital collections grow over time.

- Open computing: The library must have the option to configure, extend, and migrate applications, metadata, and system data across generations of hardware, storage, and software components while maintaining the integrity of the system.

- Compliance: The system must adhere to the Open Archival Information System (OAIS) model and the Audit Checklist for the Certification of Trusted Digital Repositories (TRAC), as well as other related standards, such as Preservation Metadata: Implementation Strategies (PREMIS) and Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).

The team spent over a year defining a detailed requirements document, which they are now happy to share with other organizations that are looking to put a preservation system in place.

Buy, Build, or Partner?

Although several memory organizations in different parts of the world had experimented with limited-scope preservation projects, pilots, and trials, the National Library of New Zealand, having examined the available commercial and open-source digital repository management systems, was unable to identify a working solution that would meet the extensive requirements of a large-scale, comprehensive preservation program such as the NDHA. Building a custom system was therefore another option evaluated by the library staff.

"Our experience and that of other institutions showed that custom software developed for a mission-critical application tended to become unaffordable over time, often leading institutions to opt for a commercial software solution. We concluded that a custom-built solution—with high up-front and ongoing enhancement costs—would be high risk for our organization," says Graham Coe, strategic advisor of innovation at the National Library. "Our preferred option for mitigating the risk was to contribute to a vendor's development cost in exchange for the commercial rights to use the product."

The library decided to take a leadership role in developing a digital archive and preservation management system in partnership with commercial vendors—Ex Libris and Sun—and in consultation with preservation experts from peer organizations around the globe. The solution developed by this partnership would be the first-ever commercially viable system of its kind for organizations wishing to preserve valuable digital collections and provide perpetual access to them.

One of the first steps was to establish an international peer review group (PRG) to serve as an independent consulting resource for the partnership. The group consists of recognized thought leaders and innovators, with institutional expertise in digital preservation and permanent access, from the international library and academic communities. Members represent such

distinguished institutions as the British Library, Cornell University Library, the Getty Research Institute, the National Library of Finland, the National Library of the Netherlands, the National Library of China, National Library Singapore, the University of Glasgow, and Yale University.

As a result of the extensive research that the National Library of New Zealand carried out and the input that it received from the PRG, the library chose to develop, in partnership with Ex Libris Group, a standards-based, commercial digital archive and preservation system that would run on Sun hardware. "We knew that we needed to engage both software and hardware partners who shared our vision for the importance of preserving a nation's digital memory in perpetuity and telling a nation's stories," explains Carnaby, of the National Library.

Assembling the Parts

The Digital Preservation System that the National Library of New Zealand is putting in place is based on the six high-level functions described in the OAIS reference model:

- Ingest
- Storage
- Data management
- Administration
- Preservation planning
- Access

StorageTek 9900

- Crossbar switch-based architecture provides the highest levels of performance, connectivity, and availability.

- Virtualization of internal and external heterogeneous storage in a single system improves flexibility and reduces TCO.

- Centralized resource management puts storage management, monitoring, and tuning under one pane of glass.

- The design helps ensure that there is no single point of failure, provides preemptive problem resolution, and supports continuous operation.

- The system scales up to 330 TB capacity with 300 GB HDDs with extreme performance at 81 GB/sec aggregate system bandwidth (68 GB/sec data bandwidth; 13 GB/sec control bandwidth) and fibre channel connectivity up to 192 redundant pairs.

Hardware and Storage Architecture

Building on an excellent, long relationship, the National Library of New Zealand turned to Sun as the hardware and storage provider most suitable to support the security, scalability, and open-architecture requirements defined for the system. "Sun has a clear leadership position in the library market," notes Coe, of the National Library. "It was a natural choice given Sun's support for modular, open, flexible, and sustainable system architecture."

The initial-state NDHA architecture includes 18 Sun Solaris™ servers for the development, testing, training, and production environments of the preservation system. The NDHA utilizes best-practice multi-tier storage architecture that helps ensure that data is stored on the least expensive, most reliable media possible while meeting the access requirements specific to each data segment. The Sun StorageTek™ Storage Archive Manager (SAM) software provides the core functionality for the storage archive architecture, featuring a flexible intelligent storage environment with policy-based data classification that enables placement across multi-tiered storage devices ranging from high-speed disks to low-cost disks or tapes.
As a self-protecting file system with continuous file integrity checks, the SAM software provides centralized metadata and simplified data management. The tiered storage architecture also simplifies data migration and technology upgrades.

Infinite Archive System L700 (formerly StorageTek CIS2)

- Management of massive amounts of data on tiered storage with multiple media types is made affordable as it is based on business-oriented policies that match the storage device to data requirements.
- SAM-QFS offers automated data migration/archival and continuous backup protection.
- Embedded SAM-QFS functionality in a preconfigured appliance reduces risk and time to deployment.

"Sun's goal is to meet the evolving needs of an international community that is grappling with storage and data management, workflow, and high-level architecture trends for preservation and archiving. In addition to providing modular, open, flexible, and sustainable system architecture, we aim to build intelligence into the hardware and storage so that the library can focus on preservation."

Art Pasquinelli
Education Market Strategist
Global Education & Research
Sun

"The Digital Preservation System is a key element of the Ex Libris product development strategy and a cornerstone of our digital product suite. Ex Libris is proud to be collaborating with the National Library of New Zealand and the peer review group in the design and development of an end-to-end digital preservation solution that will enable institutions to maintain the usability and integrity of digital resources in perpetuity. Building on the important groundwork done by the National Library of New Zealand in defining digital preservation best practices, this revolutionary commercial system will provide a comprehensive solution to the preservation needs of memory institutions worldwide."

Yaniv Levi
Preservation product manager
Ex Libris

NDHA Storage Architecture
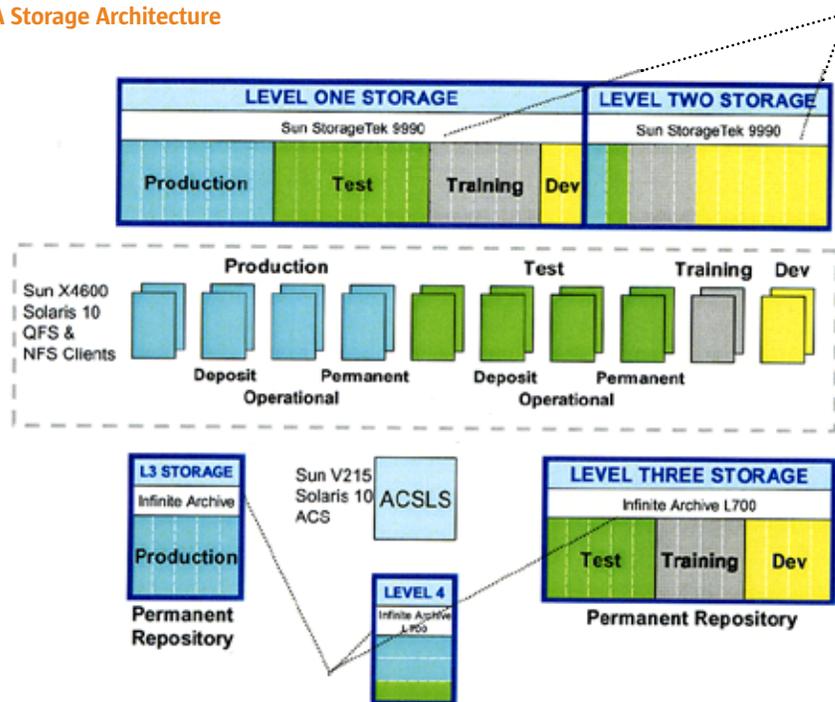
Through its collaboration with Sun, the National Library of New Zealand has access to an evolving open-computing, standards-based architecture that is intended to meet the needs of large-scale, long-term preservation programs such as the NDHA. This architecture is designed to provide:

- **A ready-to-run environment** with preinstalled hardware and software that have been tested end to end
- **Inherent business continuity and disaster recovery** with automatic recovery and recall of data from any storage tier, local or remote
- **Low total cost of ownership (TCO)** through the elimination of the overhead of separate backup solutions
- **Compliance with legal retention requirements** with always-online archiving for rapid access to all data and deep archive options with WORM and checksum capabilities
- **Extensibility and flexibility** to address many different environments

Software Architecture

Partnering with the National Library of New Zealand proved to be a perfect match for the Ex Libris product development strategy, which is based on three major principles:

- The creation of solutions that meet evolving customer and industry needs
- Close collaboration with customers and partners on the design, development, and enhancement of products at all stages of their life cycle
- The development of open platforms that are based on industry standards and enable users to customize functionality, develop their own extensions or applications, and easily integrate the product with other systems in their information architecture

In partnership with the National Library of New Zealand and with the guidance of the peer review group, Ex Libris began developing the digital preservation system that would serve as the core of the NDHA software architecture. The system follows the OIAS model, featuring modular

architecture that provides versatility in hardware and storage configuration and supports the scalability and fault tolerance requirements of this large-scale preservation initiative.

"The NDHA will offer a ground-breaking solution to address the global need for digital preservation technologies," says Chief Executive Carnaby. "We're excited about working with Ex Libris Group to the benefit of the international library community."

NDHA Software Architecture
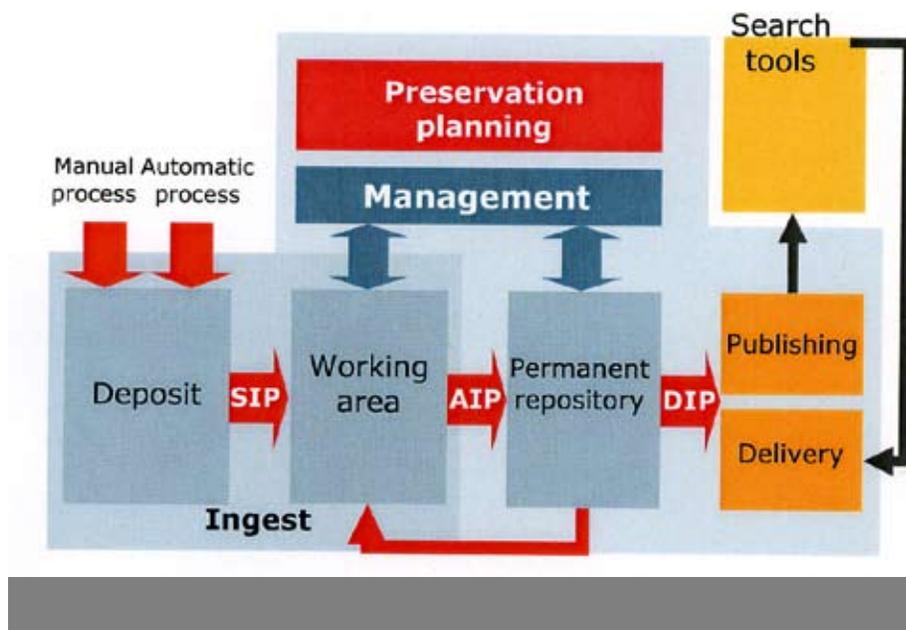
- Ingest: The deposit module and working area module facilitate the ingestion of digital objects into the system.

- Deposit: a module for the orderly deposit of materials by a wide range of depositor types, from manual deposits by an occasional depositor to mass automatic deposits by commercial and academic publishers

- Working area: an interim zone in which the characterization, processing, enrichment, and approval of materials take place prior to their entry into the permanent repository

- Permanent repository: Materials are stored in this repository after processing. The repository is designed to optimize security with minimal dependence on the underlying software.

- Management: This module supports the ongoing management of repository objects, including:

- Verification (for example, virus checks and checksum)

- Modification (for example, adding manifestation and editing) of individual objects, a group of objects, and their associated metadata

- Preservation planning: This module enables institutions to perform risk analysis on a selected set of objects and manage the preservation process for those objects.

- Access: Through its two components—publishing and delivery—this module disseminates information stored in the repository.

- Publishing: By replicating the data set or using OAI-PMH, this component creates dissemination information packages (DIPs) that external systems (such as search engines and resource discovery solutions) can use to access data stored in the permanent repository.

- Delivery: This module provides end users with access to the preserved digital entities via built-in viewers or the integration of third-party applications (such as a video streaming server).

- Administration: This module supports administrative functions, including:

- Configuration

- User management

- Monitoring

- Reporting

Paving the Way Toward Preservation

At a time when many organizations continue to debate the preservation challenge, the National Library of New Zealand has already demonstrated exemplary leadership in charting a course of action toward a practical solution that can serve as a model for large-scale, long-term digital preservation programs and systems.

The National Library of New Zealand, Sun, Ex Libris, and the peer review group have collaborated to make this vision a reality. As a result of the development of this open-system solution, others will have the opportunity to benefit from the New Zealand initiative. The National Library of New Zealand is currently sharing its business requirements, functional specifications, and other material with libraries and institutions seeking to implement a digital preservation solution. Through open source, the National Library is also making available software applications developed by the NDHA Programme. Institutions around the world are taking advantage of use cases developed by the New Zealand team to promote like initiatives in their regions.

"If each nation were to take on the responsibility for saving its national collections, the combined archive would provide people in the future with a complete digital heritage picture. This is an opportunity for national libraries to take a leadership position in this direction," says Steve Knight, associate director of the National Digital Library at the National Library of New Zealand. "All of this can be brought about effectively by a preservation network, which would make the vision of an all encompassing international memory viable."