

Repository and Preservation Storage Architecture

Keith Rajecki

Sun Microsystems, Inc.
15 Network Circle
Menlo Park, CA 94025 USA

keith.rajecki@sun.com

Abstract

While the Open Archive Information System (OAIS) model has become the de facto standard for preservation archives, the design and implementation of a repository or reliable long term archive lacks adopted technology standards and design best practices. This paper is intended to provide guidelines and recommendations for standards implementation and best practices for a viable, cost effective, and reliable repository and preservation storage architecture. This architecture is based on a combination of open source and commercially supported software and systems.

Although several operating systems currently exist, the logical choice for an archive storage system is an open source operating system, of which there are two primary choices today: Linux and Solaris. There are many varieties of Linux available and supported by nearly all system manufacturers. The Solaris Operating System is freely downloadable from Sun Microsystems. Many variants of the Linux operating system and Solaris are available with support on a fee base.

The Hierarchical Storage System, or HSM, is a key software element of the archive. The HSM provides one of the key components that contributes to reliability by through data integrity checks and automated file migration. The HSM provides the ability to automate making multiples copies of files, auditing files for errors based on checksum, rejecting bad copies of files and making new copies based on the results of those audits. The HSM also provides the ability to read in an older file format and write-out a new file format thus migrating the format and application information required to ensure archival integrity of the stored content. The automation of these functions provides for improved performance and reduced operating costs.

The Sun StorageTek Storage Archive Manager (SAM) software provides the core functionality of the recommended preservation storage architecture. SAM provides policy based data classification and placement across a multitude of storage devices from high speed disk, low cost disk, or tape. SAM also simplifies data management by providing centralized meta-data. SAM is a self-protecting file system with continuous file integrity checks.

The digital content archive provides the content repository (or digital vault) within Sun's award-winning Digital Asset Management Reference Architecture (DAM RA). DAM RA

enables digital workflow and the content archive provides permanent access to digital content files.

With SAM software, the files are stored, tracked, and retrieved based on the archival requirements. Files are seamlessly and transparently available to other services. SAM software creates virtually limitless capacity. Its scalability allows for continual growth throughout the archive with support for all data types. The policy based SAM software stores and manages data for compliance and non-compliance archives using a tiered storage approach with integrated disk and tape into a seamless storage solution, SAM software simplifies the archive storage. Allows you to automate data management policies based on file attributes. You can manage data according to the storage and access requirements of each user on the system and decide how data is grouped, copied, and accessed based on the needs of the application and the users. Helps you maximize return on investments by storing data on the media type appropriate for the life cycle of the data and simplifying system administration.

Sun Open Storage solutions provide the systems built with an open architecture using industry-standard hardware and open-source software. This open architecture allows the most flexible selection of the hardware and software components to best meet storage requirements. In a closed storage environment, all the components of a closed system must come from the vendor. Customers are locked into buying disk drives, controllers, and proprietary software features from a single vendor at premium prices and typically cannot add their own drives or software to improve functionality or reduce the cost of the closed system. Long term preservation is directly dependant on the long term viability of the software components. Open source solutions offer the most viable long term option with open access and community based development and support.

Repositories and Preservation Storage Architecture

The Repository and Preservation Storage Architecture illustrates the integration of Sun software into the implementation of digital repositories and preservation archiving software on Sun systems. This architecture delivers extreme levels of availability and offers proven enterprise-class scalability. The architecture includes specific recommendations for hardware and software components that can help improve manageability, operational performance and efficient use of storage

infrastructure.

Guidelines and Recommendations on Building a Digital Repository and Preservation Archive

The first step to building a repository and preservation storage architecture is the assessment of the business processes and defining the goals of your repository and preservation archive. Incorporating the business processes into your architectural design is crucial to the overall success of the long term archive. Documenting your organizations policies and procedures including data types, length of archive, access methods, maintenance activities, and technical specifications will increase the probability your archive architecture will meet the business requirements.

A reliable long term archive is also dependant on the software components being open and supporting interoperability. Storing, searching, and retrieving data is not sufficient criteria for a successful long term archive. A long term archive should incorporate open source standards based software to ensure future support.

The overall storage system architecture addresses the physical storage components and processes for long-term preservation. Key components to address when architecting your long-term archive are security, storage, and application interoperability. The security layer focuses on the data access in order to ensure integrity and privacy. Storage addresses the placement of the objects within the various hardware components based on retention policies. Application interoperability is the systems and applications ability to be backward compatible as well as the ability to support expanded system functionality.

When designing your repository or preservation archive system it is important to understand the needs of the users of the system. Users are not limited to those who will be accessing the repository or archive looking for objects, but includes those who will be ingesting objects as well. Your users may consist of students, faculty, researcher, or event the general public. Each of which may have different access needs. These needs will influence the server requirements of your access tier as well as the performance requirements of your search and data retrieval. You must be able to define your acceptable levels of retrieval response times in order to ensure your objects are being stored on the most appropriate storage device. High speed disk systems will provide you with faster data access compared to tape library that may need to search and mount media prior to retrieval.

Funding is also an important consideration when planning your repository or preservation archive system. You must consider the operating and upgrade cycles of your architecture in addition to the initial acquisition costs. This will prevent you from implementing a solution that is either too costly to maintain or requires drastic re-architecture as a result of the growth of the repository. This architecture takes advantage of low

cost storage combined with open standards that lower your total cost of ownership.

This architecture supports a wide variety of content types. When planning your repository or preservation archive, you should consider the various content types you will be required to support. You may want to begin evaluating and planning different preservation policies for different content types. Not all content has the same preservation requirements or value. Flexibility of the tiered storage architecture allows you to expand and contract your individual storage tiers independently as your content storage requirements evolve. Here are a few examples of some of the content type you may be consider digitizing, ingesting, and preserving in your repository:

- Manuscripts
- Books
- Newspapers
- Music, Interviews, and Video
- Web Documents and Content
- Scientific and Research Data
- Government Documents
- Images
- eJournals
- Maps

In addition to understanding your digital object types, you also want to consider the size of those objects as well as the total size of the repository. This will also allow you to forecast the growth rate of your digital repository in terms of the number of objects, object size, replication of objects, and total storage capacity. You will also want to establish and adhere to standard file formats when storing your digital objects such as tiff, jpg, or txt. It will be important that these file formats can be read by the applications that are available in the future when they are accessed from the repository or archive..

Repository Solutions

The term repository is widely debated by some. For the purposes of this solution architecture, repository refers to the system by which objects are stored for preservation archiving. There are a number of viable repository solutions available that provide the capability to store, manage, re-use and curate digital materials. Repository solutions support a multitude of functions and can be internally developed or extended. These repository solutions were highlighted for their ability to integrate into a tiered storage architecture and their support for interoperability. The repositories must be sustainable and supportable in order for the underlying storage system to operate.

Fedora

Fedora is developed by the Fedora Commons non-profit organization as a platform for providing sustainable technologies to create, manage, publish, share and preserve digital content as a basis for intellectual,

organizational, scientific and cultural heritage. Fedora is open source software built around a robust integrated repository-centered platform that enables the storage, access and management of virtually any kind of digital content. Content in Fedora can easily be accessed from the Web or by almost any software applications using available extensible application programming interfaces (API's). The connections between content items can be captured and stored in Fedora as semantic relationships describing both the linkage and its meaning.

Fedora is the first open source repository designed to work as part of an extensible framework of service components. This allows you to seamlessly incorporate Fedora into your organization's existing infrastructure. This extensible framework also allows Fedora to support trusted, secure organizational repository needs while supporting rapidly changing Web services applications. Fedora's standards-based framework can incorporate the latest technology while keeping the content safe and accessible. Using this framework, you can easily add innovative technologies as services or plug-ins without compromising the trusted core.

DSpace

DSpace is an open source digital repository system that allows researchers to capture, store, index, preserve and redistribute digital data in virtually any format. More than 300 institutions worldwide use DSpace as their digital repository. DSpace provides organizations with an easy to use end-to-end solution for managing and providing permanent access to their digital works. DSpace was originally developed as a joint effort between MIT Libraries and Hewlett-Packard (HP). It is freely available to all commercial and non-commercial organizations under the BSD open source license. DSpace is designed to work out of the box and yet it also provides the flexibility to be easily customized to meet an institution's unique needs. DSpace Manakin provides a modular user interface layer, enabling institutions to design a unique look-and-feel that can be different for each community, collection and item across the repository. Manakin also allows the user interface to extend outside of DSpace into an existing Web presence. DSpace supports multiple types of storage devices through a lightweight storage API. The storage layer currently provides support for local file systems, Storage Resource Broker (SRB), Amazon S3, or Sun SAM/QFS. New storage devices or approaches can be quickly integrated using the existing storage API's.

EPrints

EPrints is an open source software package for building open access repositories that are compliant with the Open Archives Initiative Protocol for Metadata Harvesting. It shares many of the features commonly seen in Document Management systems, but is primarily used for institutional repositories and scientific journals. EPrints was developed at the University of Southampton School of Electronics and Computer Science and is released under a GPL license. EPrints is a Web and command-line application based on

the LAMP architecture but has been ported and optimized for Solaris. Version 3 of the software introduced a (Perl-based) plugin architecture for importing and exporting data, as well as converting objects (for search engine indexing) and user interface widgets.

VTLS Inc. Vital

VITAL is a commercial institutional repository solution from VTLS Inc. designed for universities, libraries, museums, archives and information centers. Built on Fedora™, this software is designed to simplify the development of digital object repositories and to provide seamless online search and retrieval of information for administrative staff, contributing faculty and end-users. VITAL provides all types of institutions a way to broaden access to valuable resources that were once only available at a single location and to a finite number of patrons. By eliminating the traditional limitations information seekers encounter, this technology grants access to materials for all authorized end-users, from professional researchers to recreational learners. Vital is a perfect solution for organizations looking for a commercially supported alternative to open source applications.

Storage Architecture Components

Whether you are building a repository for managing institutional content, to preserve historical material, to store data for business compliance, or meet evolving business needs, a tiered storage architecture can provide you with the most reliable and cost effective solution. If architected incorrectly, ingest, searching, and preservation can be time consuming and costly. Traditional tape only archival methods simply can not meet the access requirements of many of today's repositories and long term archives. Likewise, storing all the data on disk requires greater administration and is more costly. The proposed architecture provides a proven solution with a balance between disk and tape storage hardware to support long term archiving.

Storage Archive Manager (SAM/QFS)

The Sun StorageTek Storage Archive Manager (SAM) software provides the core functionality of the recommended archive storage architecture. SAM provides policy based data classification and placement across a multitude of tiered storage devices from high speed disk, low cost disk, or tape. SAM also simplifies data management by providing centralized metadata. SAM is a self-protecting file system with continuous file integrity checks.

Sun Storage Archive Manager addresses compliance by applying policies to files, copying and moving files based on those policies and maintaining audit information on files. SAM indexes files for searchability and writes multiple copies to specific media based on the compliance retention policies.

Designed to help address the most stringent requirements

for electronic storage media retention and protection, Sun StorageTek Compliance Archiving Software provides compliance-enabling features for authenticity, integrity, ready access, and security.

Key Benefits of StorageTek Compliance Archiving Software

- Enforces retention policies at the storage level
- Software-controlled disks implement non-rewritable and non-erasable files
- Offers a cost-effective alternative to optical or tape archives
- Provides flexible Fibre Channel or SATA configurations

StorageTek Compliance Archiving software implements write-once read-many (WORM) files that are nonrewritable and nonerasable. Robust security features such as audit logs, user authentication, and access controls, combine to help safeguard the integrity of the digital information. In addition, the critical metadata attributes cannot be changed.

The Hierarchical Storage System, or HSM, is a key software element of the archive. The HSM provides one of the key components that contributes to reliability through data integrity checks and automated file migration. The HSM provides the ability to automate making multiples copies of files, auditing files for errors based on checksum, rejecting bad copies of files and making new copies based on the results of those audits. The HSM also provides the ability to read in an older file format and write-out a new file format thus migrating the format and application information required to ensure archival integrity of the stored content. The automation of these functions provides for improved performance and reduced operating costs.

Sun Fire X4500 Server

The Sun Fire X4500 Server provides a single platform for both applications and data, with enterprise server reliability features and extremely high data throughput rates. The integration of server and storage technologies, makes this an ideal platform for an inexpensive clustered storage tier. The Sun Fire X4500 Server delivers the remarkable performance of a four-way x64 server and the highest storage density available, with up to 48 TB in 4U of rack space. This system also delivers incredibly high data throughput for about half the cost of traditional solutions.

Sun Customer Ready Infinite Archive System

The Sun Customer Ready Infinite Archive System provides a pre-installed and configured storage solution for digital repository and preservation archiving. The Infinite Archive solution scales easily providing petabyte scalability. The Sun Customer Ready Infinite Archive System provides a three tier storage system consisting of the following components.

- Working Data Set, Online, on fast Fibre Channel (FC) Storage (Sun StorageTek 6140 Array)

- First Level Archive, Midline, high capacity SATA storage (Sun StorageTek 6140 Array)
- Second Level Archive, Nearline, high-performance tape storage (Sun StorageTek SL500 Modular Library System)
- Remote Archive provides a further level of archiving, with remote off-site storage of archived tapes

The Infinite Archive System takes advantage of Sun SAM/QFS software to manage the placement and retention of the data to ensure the most cost effective use of your storage resources.

Sun StorageTek 6140 array

The Sun StorageTek 6140 array is the perfect blend of performance, high availability, and reliability. The StorageTek 6140 array architecture scales to 112TB per system including the non-disruptive addition of capacity and volumes, RAID and segment size migration, and switched technology with point-to-point connections. All components in the array's data path are redundant and eliminate any single point of failure. If one component fails, the StorageTek 6140 array automatically fails-over to the alternate component, ensuring continuous uptime and uninterrupted data availability. Every component in the StorageTek 6140 array (from the disk drives to the midplane) is hot-swappable. Hot spares in every storage tray of the StorageTek 6140 array ensures high availability. Hot-spare drives can be allocated from unused drives and are always powered up and available as a spare to any virtual disk in any tray. Each array controller has two power supplies, each with its own battery backup system providing redundant power.

The StorageTek 6140 array easily adapts to change without disrupting existing applications. Compatible storage modules enable non-disruptive system upgrades and data-in-place migration of assets. The compatible and common array management across the entire Sun StorageTek Series 6000 product line protects your investment in management tools, training, and forklift upgrades.

Sun CoolThreads Servers

Sun systems with CoolThreads technology deliver breakthrough performance with dramatic space and power efficiency. Sun CoolThreads Servers are powered by the UltraSPARC T2 or T2 Plus processor, the industry's first "system on a chip" packing the most cores and threads of any general-purpose processor available. These unique servers offer energy efficiency and high performance for vertical and horizontal scalability. The Sun SPARC Enterprise T5140 and T5240 servers utilize the UltraSPARC® T2 Plus processor, which adds multisoocket capabilities to the successful UltraSPARC T2 processor. These servers are ideal for meeting the demands of ingest, web services, and metadata management.

Sun StorageTek Modular Library System

The Sun StorageTek Modular Library Systems are the

most scalable solutions on the market with up to 56 petabytes and 70,000 tape slots. This makes them the ideal platform for tape archives for off-line or dark archives. The Sun StorageTek Modular Library is complemented by the Sun StorageTek VTL Plus or Sun StorageTek VTL Value virtual solutions, which integrate seamlessly with physical tape. As a result, you gain a no-compromise solution that balances the performance, reliability, and ease of management of VTL to enable tape consolidation with the low cost, cartridge removability, and long-term retention capabilities. This tiered storage solution is managed by policies on the VTL, so the overall solution reduces your labor costs for virtual and physical tape management.

The StorageTek Modular Library Systems provide greater levels of reliability ensuring access to your data. The robotic mechanism maintains reliability regardless of the number of expansion modules and helps to increase the stability and predictability of backups. Redundant, hot-swappable components, such as power supplies and fans, minimize disruption. An advanced digital vision system automatically calibrates the library to reduce wear and tear on the cartridge, drive, and robot. Dynamic worldwide naming and firmware code uploads eliminate single points of failure.

Sun Identity Management Suite

The Sun Identity Management Suite is a key component to ensuring the security and data integrity of the digital repository and preservation archiving solution. Identity Manager provides a comprehensive user provisioning and identity auditing for efficiently and securely managing identity profiles and permissions while Sun Directory Server Enterprise Edition provides a secure, highly available, scalable, and easy-to-manage directory infrastructure that effectively manages identities in this growing and dynamic environment.

Solaris

Although several operating systems currently exist, the logical choice for an archive storage system is an open source operating system, of which there are two primary choices today: Linux and Solaris. There are many varieties of Linux available and supported by nearly all system manufacturers.

The Solaris Operating System is freely downloadable from Sun Microsystems and provides a number of technical advantages from file system support to security and supportability. The Solaris ZFS offers a dramatic advance in data management with an innovative approach to data integrity, performance improvements, and integration of file system and volume management capabilities. Solaris Dynamic Tracing (DTrace) allows you to analyze, debug, and optimize your systems and applications.

The Solaris OS also offers binary compatibility within each Sun server line, whether based on UltraSPARC®, AMD Opteron, or Intel Xeon processors. As a result, all Sun servers running the Solaris 10 OS provide powerful

features that can help reduce cost, complexity, and risk. Many variants of the Linux operating system and Solaris are available with support on a fee base.

Conclusion

A tiered storage architecture provides the most cost effective solution for object repositories and long-term archives while supporting scalability. The extent at which those storage tiers are deployed is dependant on the access patterns and archival policies. Although this architecture is not intended to cover all business requirements, it can be applied in a modular approach to address specific business requirements where one or more tiers may not be feasible due to business or technical requirements.

Repository and Preservation Storage Architecture Key Benefits

The architecture identifies key system components and processes that are required to achieve high service levels and scalability. It provides the following major benefits to educational institutions:

- Higher service levels — The architecture is designed to optimize service levels with redundant components and automated failover using storage virtualization and cluster technologies.
- Reduced cost — Virtualization technologies enable consolidated solutions with higher resource utilization and tiered storage helps customers avoid overprovisioning or underprovisioning their systems. Best practices for management can also reduce the cost of maintaining the solution environment.
- Faster time to delivery — Accelerates deployment by providing proven and tested configurations with simplified installation to be up and running almost immediately.
- Reduced risk — Validated hardware and software configurations greatly reduce the risk of unforeseen problems in a production implementation.

References

Consultative Committee for Space Data Systems (2002). "Reference Model for an Open Archival Information System (OAIS)". CCSDS 650.0-R-1 – Blue Book. Available at: <http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>

Repository and Preservation Storage Architecture

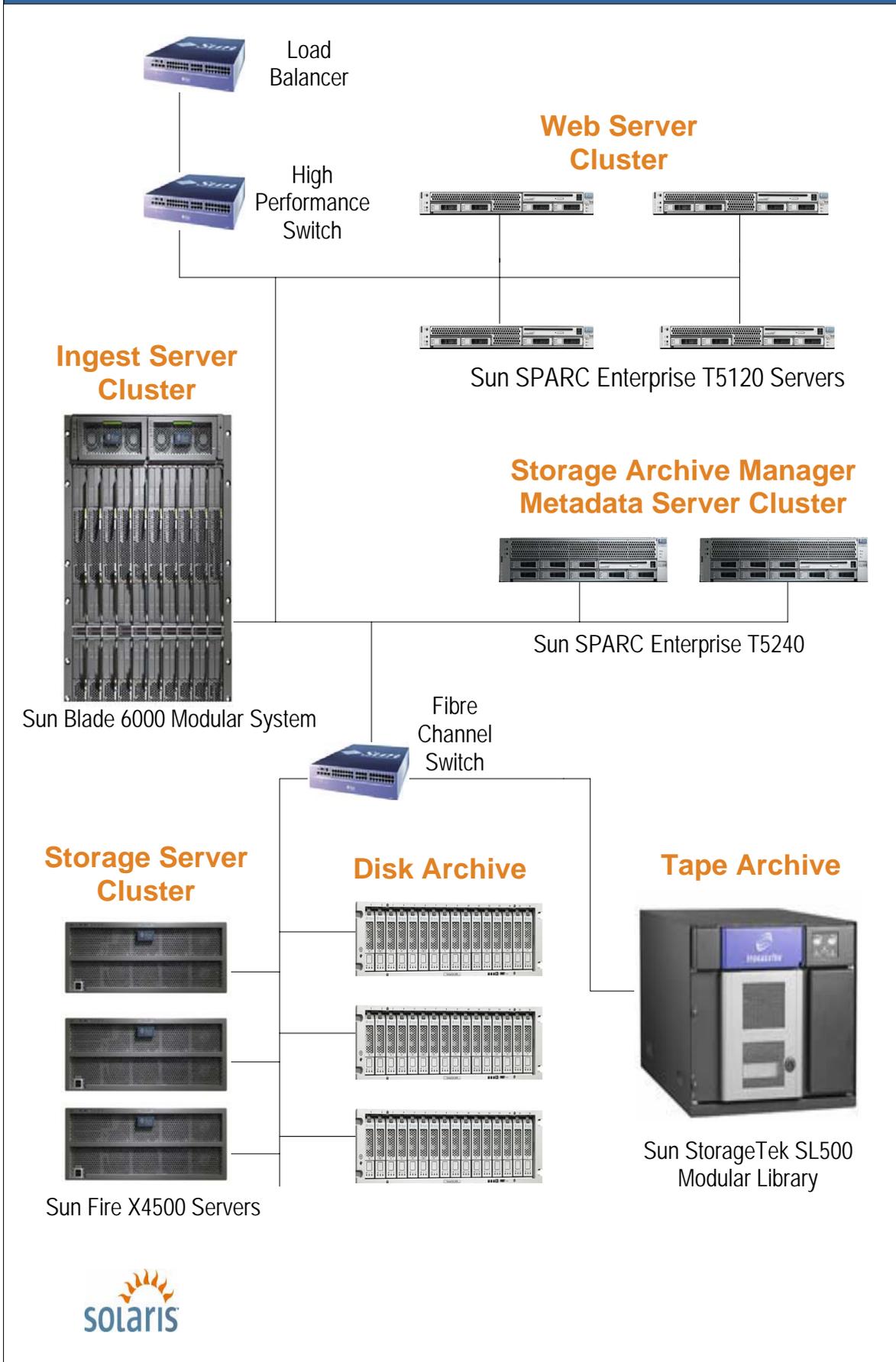


Figure 1: Digital Repository and Preservation Architecture