

Enterprise-Wide Digital Repository and Archive

Breakthrough open solution based on the Sun Open Archive Framework



Highlights

- Eliminate risk of propriety repositories and archives – this next generation solution based on the Sun Open Archive Framework leverages industry-leading open source content management and repository applications, tested and optimized for Sun Open Storage.
- Fully utilize enterprise-wide digital information while reducing cost and complexity of preserving it through its lifecycle.
- Easily add rich metadata to describe and find digital assets.
- Create collaborative information spaces for different work groups or projects where relationships between items can be captured and preserved – relationships can be as important as the data itself.
- Manage digital rights and authentication.
- Scale with simple, cost-effective expansion. Start with a single server and two terabyte storage system and grow to multi-server, multi petabyte systems. All using the same application software stack, and without the extra expense of license fees.
- Comprehensive support services for the solution are available today through Sun and its partner network.



Repositories are critical to the management of enterprise-wide digital information and to unlocking the value of the information created and stored. Whether saving digital information for on-going collaborative use or for long-term preservation, one of the most challenging aspects of repository and preservation archive solutions is that data must survive multiple generations of technology and complete turnover in the staff who manage the data. Open source solutions provide greater assurances that the digital information can be preserved and accessed regardless of technology upgrades or staffing issues.

The Challenge

As people are creating and managing content on a scale never imagined before, IT managers are looking for more efficient ways to manage their costs, reduce the risk of damaged or lost data and store the data so that it can be readily accessed and shared as needed. It is becoming clear that traditional methods of managing data are not sufficient for today's needs. Organizations are looking for easy to deploy, scalable repository solutions that can offer:

- Efficient methods to protect data and preserve its integrity over time
- Ways to easily find, share and reuse information to increase its business value
- Open data formats and technologies and use of industry standards to enable data and applications to survive technology transitions, vendor consolidation or obsolescence, and personnel changes
- Economic sustainability; ways to control costs and complexity as systems scale

A New Approach to Scalable Repositories

Together, Sun and its partners bring a breakthrough in innovation for digital content. Institutions can now more fully utilize their digital information while reducing the cost and complexity of preserving it through its lifecycle.

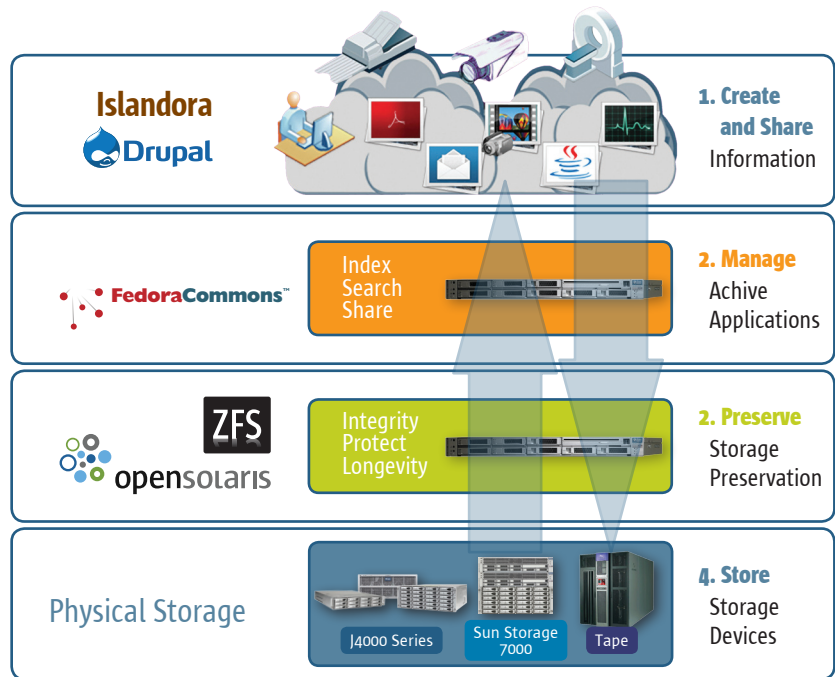
Future-Proof

The Sun Repository and Archive solution leverages the Sun Open Archive Framework. The Framework decouples the storage preservation aspects from the storage platform so that Sun and its partners can architect an archive solution that best meets customer requirements now – and into the future. The Framework includes a broad selection of Sun's standards-based, Open Storage devices, and also includes tape where it makes sense. For the Sun Repository and Archive solution, Sun engineers and its partners have optimized the applications for the Sun Storage 7000 Unified Storage Systems and the J4000 Series – providing customer choices for performance, capacity and cost based on their data size, ingest/retrieval patterns and scale requirements. The web-based user experience, data management and application-level preservation features of the Sun Repository and Archive solution are provided by open source partners and their powerful applications.

More than the sum of its parts

The result is a solution which is more than the sum of its parts. Drupal with Islandora provides an easy, powerful way to create customized Web sites with an organization's own unique content and branding and offers

Sun's Open Archive Framework brings together the components required to provide the highest level of data availability and preservation in addition to future-proofing data from technology changes over time.



Leveraging the Sun Open Archive Framework

fine control over access to collections and individual data assets. Adding the Fedora Repository provides durability to the content while also enabling seamless sharing of content with other applications. The Sun Open Archive Framework's Preservation Software layer adds robust storage protection and data handling combined with powerful management tools, while Sun Open Storage delivers the most cost effective and easily deployed storage available. Together these components get customers up and running fast with the assurance they will be able to grow and evolve the system gracefully, protecting investments.

Drupal is a popular open source social publishing system that blends web content management and social media capabilities – such as blogs and community discussion forums. It provides an ideal, easy-to-use user interface to the Fedora Repository when combined with the Islandora module.

Islandora is an open source module for the

Drupal web content management system written by the University of Prince Edward Island (UPEI) to allow Drupal to act as a web-based front end to the digital repository and preservation platform from Fedora Commons. Islandora uses content models to determine what MIME types are allowed to be ingested into Fedora Repository, and what to do with the object on ingest. For instance a PDF content model may tell the module to create a thumbnail and ingest the thumbnail as a datastream along with the actual PDF datastream. The module also enables viewing and management of Fedora Repository objects, including functions such as ingesting, purging, adding datastreams, searching and browsing by collection.

Fedora Repository (or Flexible Extensible Digital Object Repository Architecture) is an open source storage repository and preservation platform that uses a service-oriented architecture to enable the creation of innovative, collaborative information spaces. It is designed for the longevity and integrity of any kind of

digital content, and also offers the ability to inter-relate (link) such content from different sources. Fedora Repository was originally created at Cornell University under research grants from NSF and DARPA. This research evolved into a successful open source project jointly managed by Cornell and University of Virginia. Today, development of Fedora Repository continues in the context of the Fedora Commons non-profit organization.

Sustainable and Reliable

In today's economic climate, it's more important than ever that businesses deploy solutions that are cost-effective. They must be simple to order, fast to deploy, and run – day in and day out, without interruption. The Sun Open Storage building blocks available as storage devices for this solution are based on cost efficient industry standard components and a robust open source software stack built from the enterprise hardened OpenSolaris™ OS and Zettabyte File System (ZFS), and supported by a passionate, global community of developers. While many other

vendors charge license fees for protocols and data services such as CIFS, NFS, HTTP, and Replication, these and other software features are included in the price of Sun Open Storage Systems. No additional software licenses are required. Sun Open Storage Systems also offer further economic value by reducing energy consumption and data center space.

The Fedora Repository has been optimized to run on top of the Sun Storage 7000 Unified Storage Systems and J4000 series as well as Solid State Disk technology to enhance performance as part of a Hybrid Storage Pool. Fedora Repository is built with a flexible framework that allows easy extensibility not just to new data formats but also to new storage platforms and devices.

Customers benefit from an infrastructure that will evolve gracefully in order for the digital assets to outlive the technology that initially creates, persists and uses it – providing long-term sustainability and ability to non-disruptively move digital assets through technology changes over time – into the future. The Repository and Archive solution provides customers the choice of Sun Open Storage devices so that they can eliminate their dependence on proprietary systems, saving up to 90% on the cost, up to 70% less energy, 50% less space and with minimal IT staff versus alternative offerings.

Advanced preservation features

Sun Storage Preservation Software

The storage preservation layer pioneered by Sun and based on OpenSolaris, ZFS, and SAM, provides all the necessary storage preservation, policy and persistence required

of archived digital assets and is de-coupled from the physical storage. Features include data integrity verification and repair (19X9s), checksumming and protection (Raid, snapshot, clone), analytics, predictive self-healing sensors, policy based data migration and simplified management. Based on open software and standards, the customer is not locked into a proprietary API or a specific vendor. As hardware technology changes, digital assets are easily migrated from old technology to new while maintaining complete integrity and persistence. This differentiates this Sun Open Archive Framework solution from traditional architectures where the storage preservation is locked into the hardware itself.

Fedora Repository features

Fedora Commons provides an advanced platform for deploying preservation-enabled repository systems that can seamlessly be integrated with existing applications and IT infrastructure. All content and the metadata essential for making sense of content is kept in ordinary files which can be replicated to remote locations for added safety. The metadata is stored as XML which is open and accessible for the foreseeable future. In the event of a disaster, the entire Fedora Repository can be rebuilt from those files eliminating the risk of any single point of failure. As content formats change, Fedora provides support for migrating files to new formats, optionally keeping the previous versions for security. To help manage formats, Fedora is not limited to MIME types but can use format identifiers from any format registry or scheme at any desired precision. Fedora has an advanced capability that enables adding customized services to content types that

hides implementation details and provides a stable way for applications to access the content—even as it is migrated to new formats—and easing the introduction of new technologies over time. Changes to content or metadata are recorded in audit trails, also in files, and are combined with cryptographic checksums to help ensure authenticity. With an eye to the future, Fedora enables plugging in viewers and emulators so that obsolete formats can be accessed even if the original software is no longer available.

Services and support

Sun and its partners' deep technical expertise and proven implementation and migration tools makes it easy to deploy Sun Open Storage and the open source Islandora/Drupal and Fedora Repository software stack. They provide the expertise to help ensure that your systems are running at an optimal level right from the start.

Open source, Open to Innovation

Many of these innovative solutions are built on open standards, so you can easily extract archived data for future use or move it between devices when you add new technologies to your infrastructure.

Secure? You bet. A US Federal Government agency has found open source software to be 7 times more secure than commercially distributed software.

Learn More

Sun Open Archive Framework:

<http://www.sun.com/openarchive>