AUTOMATED DATA RETENTION WITH DELL EMC ISILON SMARTLOCK

ABSTRACT

Dell EMC Isilon SmartLock protects critical data against accidental, malicious or premature deletion or alteration. Whether you need to safeguard intellectual property, preserve digital work products, or comply with corporate governance or regulatory requirements for data retention, SmartLock delivers security with precision, control and simplicity.

October 2016

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Dell EMC® Isilon® SmartLock™ software is a reliable and secure data protection and retention capability that protects critical data from unauthorized alteration. Protecting critical data from accidental deletion or alteration is a key business imperative for most organizations today. Loss of financial data or business records is a critical need for many organizations today to help ensure that they meet necessary regulatory and governance requirements. SmartLock is a software-based approach to Write Once Read Many (WORM) data protection.

AUDIENCE

This white paper is intended for IT Managers, storage administrators, and security administrators.

INTRODUCTION

The most valuable outputs of modern companies and organizations are usually electronic: data and digital work products created on computers and stored on disk. For example, the manufacturing of physical goods is based on an electronic design, a finished movie usually consists of one big file, an architectural design may only ever exist on disk, and medical treatment planning is dictated by an electronic health record or electronic medical image. These product designs, movies, building plans, x-rays and other digital elements need to be protected. In many cases, how it is protected and for how long will be determined by company policy or regulatory oversight.

Adherence to retention rules is most easily and reliably met through the use of automation. Automated retention systems set the retention time of data based on user requirements, and hold the protected data unchanged for the required time.

Retention systems can be implemented in either hardware or software. Hardware implementations are dedicated retention systems rather than general purpose storage. This typically means they carry a price premium and also that your staff needs to be trained on how to manage the additional storage infrastructure. Software implementations vary widely in manageability, flexibility and granularity with some requiring large capacities of storage to be dedicated to retention for the long-term, or even permanently.

Ideally, an automated data retention solution is simple to implement and manage, highly reliable, flexible enough to support multiple use cases without having to invest in dedicated hardware, and easily scalable to meet your needs today and for the foreseeable future.

This paper describes how Dell EMC Isilon SmartLock can be used to automate data retention to protect data efficiently and securely.

AUTOMATED DATA RETENTION: THE Dell EMC ISILON APPROACH

ONE CLUSTER FOR BOTH GENERAL DATA AND RETAINED DATA

Dell EMC Isilon SmartLock software delivers secure, powerful data retention in a simple, general purpose NAS architecture that scales to petabytes. SmartLock is very flexible and does not require application integration, specialized hardware or dedicated minimum storage capacity.

Because it is integrated with the Dell EMC Isilon OneFS® file system, SmartLock is designed to seamlessly work with OneFS core capabilities and other key storage functions such as snapshots, replication, provisioning, backup and restore, deduplication. and virtual environments. Integration with OneFS also means your retention environment can grow seamlessly, with just one file system to manage, from terabytes to over 68 petabytes in the same cluster.

SmartLock is implemented at the directory level in the file system, which means that with one cluster, you can store both general data and data with retention requirements, regardless of your capacity requirements or growth patterns. As capacity requirements increase, the entire cluster grows as a whole – you do not need to provision general data capacity separately from retained data capacity.

THE BENEFITS OF SCALE-OUT ARCHITECTURE FOR DATA RETENTION

Most software-based retention solutions that work with traditional NAS systems require you to dedicate, at a minimum, an entire storage volume for data retention. Typically, this means a minimum capacity investment just to get started and an incremental capacity investment every time you want to grow. Volume based approaches also have maximum size limitations, forcing you to split data across multiple management points as you grow, increasing complexity and administration overhead.

The Dell EMC Isilon OneFS operating system combines the capabilities of traditional volume management, file system and RAID data protection into a flexible and powerful software layer that is space efficient and simple to manage – even at scale. Because OneFS does not rely on underlying RAID structures like volume groups, data management with OneFS is very flexible and granular. This means you can implement automatic data retention with any amount of data – from one file to billions of files.

SMARTLOCK OPERATING MODES

In OneFS 7.0 and later, there are two SmartLock operation modes available to the cluster: SmartLock compliance mode and SmartLock enterprise mode. SmartLock enterprise mode is the default SmartLock operation mode. SmartLock compliance mode was introduced in OneFS 7.0. It enables you to protect your data in compliance with the regulations defined by U.S. Securities and Exchange Commission (SEC) rule 17a-4.

SmartLock enterprise mode provides the ability to create write-once read-many (WORM) directories on the cluster, but the WORM implementation does not meet the requirements of SEC rule 17a-4. You can use SmartLock enterprise mode if you want to protect files from accidental modification or deletion.

HOW SMARTLOCK WORKS

Once a directory is created and marked as a SmartLock directory, files committed in this directory will be kept immutable until their retention time expires; they cannot be deleted, moved or changed. Retention dates on files are set by the administrator, and can be extended but not shortened. When a file retention policy expires (i.e. passes its retention date), it becomes a normal file which can then be moved or deleted as required, allowing you to reclaim that storage capacity for general purpose or retained data.

MIXING SMARTLOCK AND NON-SMARTLOCK DATA

Any empty directory under the OneFS file system can be designated as a SmartLock directory. Starting in OneFS 8.0, a directory need not be empty before designating it as an ‘Enterprise’ SmartLock directory. This means you can mix SmartLock and normal directories on the same cluster. Once a directory is designated as a SmartLock directory it is ready to protect files placed there. Any subdirectories in a SmartLock directory are automatically protected by SmartLock and inherit all the settings of the parent directory.

Note: In OneFS 8.0 a directory still needs to be empty if it needs to be designated as a ‘Compliance’ SmartLock directory.

COMMITTING FILES AND SETTING RETENTION DATES

Files placed in a SmartLock directory are changeable or moveable until they are committed. This allows time for administrative changes and appends before the file becomes immutable. Files can be committed to a SmartLock directory locally or over a network (via NFS or SMB/CIFS).

File retention times can be set in two ways – on a per file basis or using the directory default setting. If you need to have a specific file committed for a specific amount of time, you set that file’s retention date individually. If, however, as is more typical, you want any data committed to a SmartLock directory retained for the same length of time, you simply set the directory default to the desired period. In this scenario, any file committed to that directory without a unique retention date will take that directory default.

Even after the retention date has expired, files protected with SmartLock cannot be altered while they remain in a SmartLock directory – they must be moved to a non-SmartLock directory before they can be changed. SmartLock directories themselves cannot be deleted while data is stored in them. Only a completely empty SmartLock directory can be deleted.

LITIGATION HOLD

In some situations, such as legal discovery actions, it is necessary to guarantee data remain protected until that specific situation is resolved. In these cases, a directory level override of retention dates is available (the “The Override Retention Date“) which automatically extends the retention date of any file to or passed the expected resolution time. This functionality will only extend retention times – files whose retention times were already set beyond the scope of the override are not affected.

In Table 1 below, the Directory Default Retention Offset is one year for both Example 1 and Example 2. Therefore, any file committed to that directory that does not have a specific expiry date, as in Example 1, will automatically get a one year expiry date from the date it is committed. This means that the SmartLock protection of a file committed on Jan 1, 2012, for Example 1, will be January 1, 2013, based on the one year default setting. In Examples 2 and 3, a specific retention date of February 1, 2013 has been specified for a file. In these cases, the specific file retention date takes precedent over any Directory Default Retention Offset period.

In Example 4, the company receives a one year litigation hold on all data pertaining to a case under investigation on December 31, 2012. The Override Retention Date setting at the directory level is used and all data in that SmartLock directory is automatically protected through a minimum of December 31, 2013.

	Example 1 No File Retention Date Set	Example 2 File Retention Date > Directory Offset	Example 3 Directory Offset > File Retention Date	Example 4 Litigation Hold Added
File Retention Date	None	February 1, 2013	February 1, 2013	February 1, 2013
Directory Offset Retention Date	1 Year	1 Year	2 Years	1 Year
File Committed Date	January 1, 2012	January 1, 2012	January 1, 2012	January 1, 2012
				1 Year Litigation Hold Added Dec 31, 2012
Expiration Date	January 1, 2013	February 1, 2013	February 1, 2013	February 1, 2013

Table 1: File Retention and Litigation Hold Examples

PRIVILEGED FUNCTIONS - DELETING, MOVING OR CHANGING SMARTLOCK DATA

General users and applications cannot change the data or metadata of SmartLock committed files, or move them or delete them. If a general user or application needs to change SmartLock protected data, they need first to copy it to a non-SmartLock directory and then make their changes to the copy. The SmartLock protected original cannot be altered by anyone except a privileged user.

For OneFS versions prior to OneFS 8.0 a privileged user is someone who has root access to the system and can delete SmartLock protected files. Privileged deletes can only be performed locally, not over the network, which adds an additional layer of control of privileged functions. The privileged user exists only in the Enterprise version of SmartLock. Starting with OneFS 8.0 a privilege delete can be performed by non-root users via the RBAC role “ISI_PRIV_IFS_WORM_DELETE “.

INTEGRATION WITH OTHER ONEFS CAPABILITIES

Because it is integrated with the Isilon OneFS file system, SmartLock seamlessly integrates with OneFS core capabilities and add-on software for snapshots, replication, provisioning, backup and restore, virtual environments and other key functions described below.

SNAPSHOTIQ

Unlike traditional NAS environments which have many restrictions, Dell EMC Isilon SnapshotIQ™ provides an unlimited number of snapshots within a single pool of Isilon storage and up to 1,024 snapshots within any single directory. Unique snapshot policies to be set at the cluster, directory, and sub-directory level giving you the control to protect exactly what you want.

Data in a SmartLock directory can be snapshotted using SnapshotIQ. Such snapshots can be treated as normal data, or moved into a SmartLock directory and committed for retention.

For more information on SnapshotIQ, please see: http://www.emc.com/storage/isilon/snapshotiq.htm.

ARCHIVING, DATA BACK UP AND RECOVERY

Isilon clusters can also be backed up by most major archiving and data backup products via Network Data Management Protocol (NDMP). Data backup also can be parallelized leveraging multiple nodes to maximize performance. SmartLock retention settings are retained through the backup and restore processes so you do not need to recommit files after a restore.

For more information on Isilon Backup and a list of certified products please see: http://www.emc.com/data-protection/archive-big-data.htm.

DATA REPLICATION FOR RELIABLE DISASTER RECOVERY

Data on Isilon clusters can be replicated using Dell EMC Isilon SyncIQ™. SyncIQ delivers extremely fast replication of millions of files and terabytes of data over the WAN and LAN. SyncIQ also enables replication jobs to be parallelized and evenly distributed across all sending and receiving nodes in the Isilon IQ storage clusters. SyncIQ can be used to replicate SmartLock protected or normal data for business continuity, disaster recovery, disk-to-disk backup or remote archive.

The following table illustrates the compatibility of SyncIQ between different types of SmartLock directories on the source and target of SyncIQ.

Source Directory Type	Target Directory Type	SyncIQ Source-To-Target Compatibility
Non-WORM	Non-WORM	Yes
Non-WORM	Enterprise SmartLock	Yes
Non-WORM	Compliance SmartLock	No
Enterprise SmartLock	Non-WORM	Yes*
Enterprise SmartLock	Enterprise SmartLock	Yes
Enterprise SmartLock	Compliance SmartLock	No
Compliance SmartLock	Non-WORM	No
Compliance SmartLock	Enterprise SmartLock	No
Compliance SmartLock	Compliance SmartLock	Yes

* Replication type is allowed, however retention will not be enforced

For more detail on SyncIQ, please see: http://www.emc.com/storage/isilon/synciq.htm.

SYNCIQ FAILOVER AND FAILBACK WITH SMARTLOCK

SyncIQ replicates the data asynchronously from the Primary Isilon cluster to one or more clusters. This creates multiple copies of data. This can be used in case of a failure on the primary Isilon cluster. The process of activating a secondary copy for read-write purposes is called Failover. This is usually done when the primary Isilon cluster becomes unavailable or is taken down for maintenance. When the original primary cluster is back online, a reverse sync can be established using SyncIQ. If required, the original primary can be made the read-write copy. This process is called Failback.

In OneFS 8.0 both Failover and Failback using SyncIQ could be done only for ‘Enterprise Mode’ SmartLock directories.

OneFS 8.0.1 introduces support for SyncIQ failover and failback with ‘Compliance Mode’ SmartLock directories, delivering automated disaster recovery for financial services SEC-17a4 regulatory compliance.

This means Isilon customers can fail over from one Isilon cluster running in compliance mode to another if, for example, a cluster becomes unavailable. The customers can then fail back to a primary cluster if the primary cluster becomes available again. Customers can revert failover if they decide that the failover was unnecessary, or if they failed over for testing purposes.

USE CASES

As described above, automated retention systems are used to protect critical data from accidental, premature or malicious alteration or deletion. Let's take a closer look at some of the use cases – some which apply across multiple workflows and some which are specific to certain industries.

COMPLYING WITH CORPORATE GOVERNANCE

Frequently companies have put governance requirements in place to comply with standard industry practices, government regulations such as ISO or Sarbanes-Oxley, or to create specific security or information management standards.

These regulations frequently require the retention of data for specific timelines. Retention timeline requirements can vary widely by department, function or even data type, so without an automated data retention system, adhering to such guidelines would be very difficult; resource intensive and risky.

With the ability of SmartLock to match retention requirements to any of a number of parameters including data type, location, and owner, corporate governance requirements can be met easily over time regardless of how they change.

MANUFACTURING: RETAINING REFERENCE AND CURRENT DESIGN DATA

Most manufacturing design processes draw on current work, historic designs and a library of reference data. Of these, historic designs – representing past and currently-manufactured products which are the company's current revenue stream – are frequently the company's most critical asset. Historic designs are not only a reference for current development; they also constitute the company's intellectual property portfolio and are the foundations of patent creation and protection as well as liability defense.

Keeping all necessary data together is beneficial for applications and end user access. Mixing it in the same cluster also significantly reduces the overhead of purchasing, provisioning and maintaining a separate storage solution.

Because SmartLock is integrated with OneFS and its Dell EMC Isilon SmartPools® automated tiering capability, current work, historic designs and reference data can all be stored on the same cluster. As shown in Figure 1, each data type can be placed on the most appropriate storage type to ensure proper performance and access. Each data type can also have a different protection configuration to balance cost of storage versus data criticality.

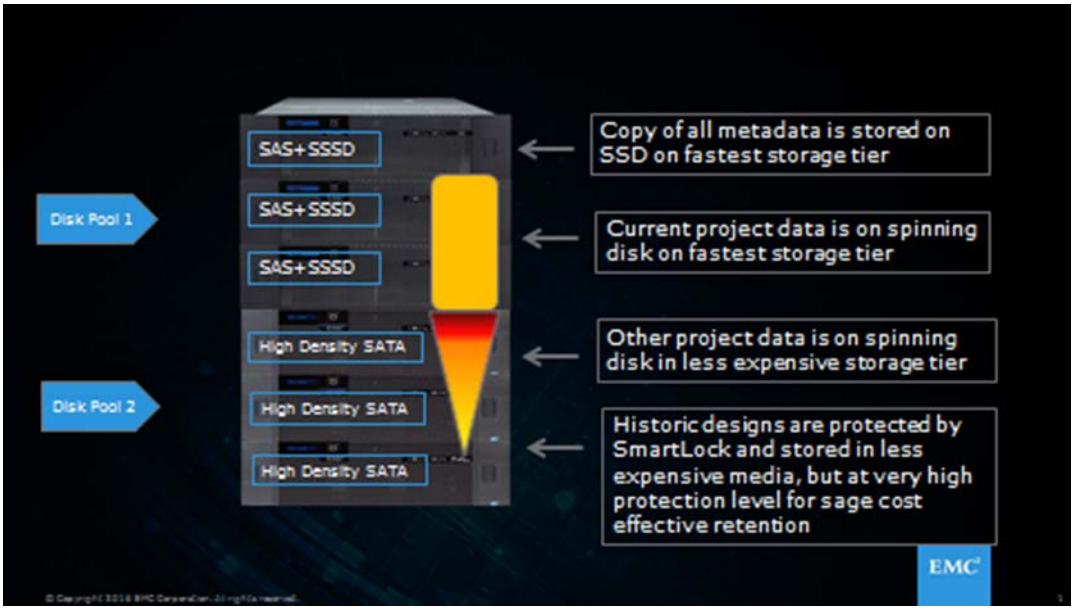

Figure 1: Automated Data Retention for Manufacturing Design Data

FEATURE FILMS: LOCKING DOWN FINAL CONTENT IN A PRODUCTION ENVIRONMENT

A modern feature film with a large number of special effects or one that is in 3D can easily require over a petabyte of storage while work is in progress. A finished feature film is in the multiple terabytes and can be stored as one or many files, depending on the creator's distribution requirements, all of which are quite large.

Most production facilities work on multiple projects at a time. Some, like sequels, reference previous films, drawing on the same creative. Once complete, movie files need to remain unchanged for very long periods of time, or, as some producers say, "Forever". If the company does not want to purchase and maintain a completely separate facility, they need to accommodate the finished films in their production storage environment. Production environments tend to operate at great speed with many users having significant system access.

With SmartLock, finished films can be specifically protected within an otherwise general purpose storage environment – keeping the film company's greatest assets intact regardless of current production activity.

GAMING: LIMITING COMPLEX FRAUD IN CASINOS

The gaming industry attracts a lot of attention from potential thieves due to its concentration of cash and high volume of financial. Theft attempts are especially high in casinos, offering a physical location as a tempting target. To address this threat, many firms have installed extensive video surveillance systems to monitor gaming activities.

Industry studies show that an astounding percentage of casino theft is carried out with the involvement of casino employees (67% in the State of Nevada[1]). The more sophisticated attempts have included wiping of surveillance data, which is crucial in prosecuting casino based theft. To protect this data, many casinos are beginning to invest in automated data retention technologies.

The integration of SmartLock with OneFS enables a casino to begin with a few terabytes of surveillance data and then grow that same cluster to 68 petabytes before needing to install a second cluster. A casino can also chose how long it retains data without worrying about running out of room. If that casino is attached to a resort or hotel, the surveillance system for the entire property – hotel public areas, casino, parking lots, etc. – can be combined into one cluster with retention set only on the casino floor data.

CONCLUSION

The need to retain data to meet specific requirements concerning data immutability and longevity is rapidly growing due to a number of factors including increased corporate oversight and regulatory compliance requirements, and evolving data protection requirements.

Dell EMC Isilon SmartLock is an automated retention system that removes much of the risk and complexity associated with retaining data for specific time periods. SmartLock is designed to provide enterprises with a highly flexible system and an easy-to-use system to protect their data against accidental, malicious or premature alteration or deletion.

Because it is integrated with the Isilon OneFS file system, SmartLock also works seamlessly other key storage functions including data backup, archiving and disaster recovery. Integration with OneFS also means your retention environment can grow seamlessly, with just one file system to manage, from terabytes to over 20 petabytes in the same cluster. Additionally, customers can now use SyncIQ failover and failback with SmartLock Enterprise as well as Compliance directories to perform disaster recovery. This makes Dell EMC Isilon the first company to offer the failover and failback capability in SmartLock Compliance mode.

For more information, see http://www.emc.com/storage/isilon/smartlock.htm.

[1] Casino Enterprise Management Magazine, December 2010