



Integration of IBM Storage Protect for Data Retention with IBM Storage Scale immutable filesets

Overview and configuration guidance

Version 1.7 (June 2026)

Contents

Introduction	3
Solution overview	3
How it works	4
Special case: Reclamation	5
Configuration	6
Prerequisite.....	6
Creating IBM Storage Scale Immutable fileset	6
Creating and configuring the storage pool	7
Creating and configuring nodes, domains and archive copy groups	8
Further guidance.....	8
Administration	9
Appendix	9
References	9
Disclaimer.....	10

Acknowledgement:

Thanks to Haizhu Liu (IBM Storage Scale development), Colin Dawson (IBM Storage Protect development) for the thorough technical implementation. Thanks to Harley Pucket, Del Hoobler (IBM Storage Protect) and Carl Zetie (IBM Storage Scale) for driving the changes into the products. And thanks to Diem Nguyen (IBM Storage Protect test team) for creating and executing the test suites.

Introduction

In the backdrop of exponentially growing data accelerated by new IT trends - such as cloud, analytics, Internet of Things and social media - digital archiving is becoming increasingly important. The ever-growing volumes of data along with stringent regulatory and industry compliance requirements are the main drivers for long-term archival solutions. Just using backups to fulfill the requirements for archiving is not enough.

One key business challenge for archiving is the need to comply with laws and regulations. Failures to do so can result in huge fines. The first and perhaps hardest step is to identify laws and regulations that are relevant for certain types of data. Once identified this data needs to be archived in accordance with the requirements of subject laws and regulations. One of the key requirements imposed by laws and regulations is to store archive data in an immutable fashion that prevents deletion and changes during the defined retention period. In other words, the data must be stored in a write-once-read-many fashion (WORM).

IBM provides two software defined storage offerings with WORM storage capabilities: IBM® Storage Protect™ for Data Retention (SP-DR) and IBM Storage Scale™ immutable filesets. These two offerings can be combined to provide an end-to-end WORM storage solution for archiving. This paper describes this solution and provides configuration and operational guidance.

Solution overview

As shown in figure 1 the solution combines two software defined storage solutions: IBM Storage Protect for Data Retention and IBM Storage Scale immutable filesets:

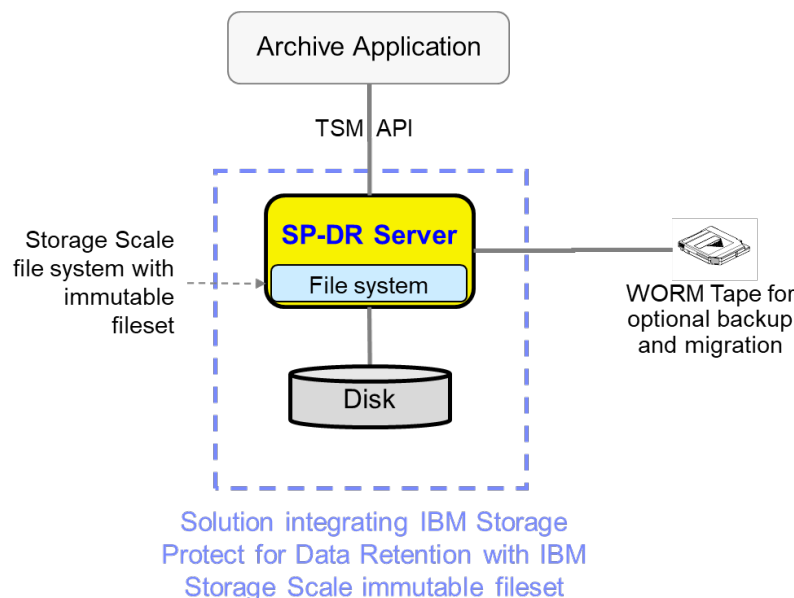


Figure 1: High level architecture of the solution combining IBM Storage Protect for Data Retention with IBM Storage Scale immutable filesets

IBM Storage Protect for Data Retention (formerly IBM System Storage Archive Manager, SSAM) provides storage software to help organizations meet legal and regulatory requirements for archiving and retrieving data [1]. IBM Storage Protect for Data Retention is a special version of IBM Storage Protect that is made for archiving of data in a WORM fashion. IBM Storage Protect for Data Retention receives archived objects from applications, associates a retention period and stores objects in so-called storage pools that reside in file systems.

IBM Storage Scale is a scalable parallel file system that can be used for many purposes. IBM Storage Scale also allows configuring immutable partitions in a file system that are called immutable filesets [3]. From a user perspective a fileset is a directory within a IBM Storage Scale file system. Immutable filesets allow managing immutable and append-only files, like the SnapLock® method created by NetApp Inc. With the SnapLock method files can be set to immutable or append-only for a given retention time using standard file system commands. During the retention time such files cannot be deleted or modified. When the retention time has expired these files can be deleted but still not modified.

IBM Storage Protect for Data Retention can store the archived data in immutable filesets of an IBM Storage Scale file system and manage the retention periods. IBM Storage Scale assures that immutable files stored in an immutable fileset cannot be deleted. The advantage is that the data is protected on a storage device level and not only on the level of IBM Storage Protect for Data Retention. In the next section we will explain how this works.

Both IBM Storage Protect for Data Retention and IBM Storage Scale with immutable filesets have been assessed for compliance according to different laws and regulations [2] and [4]. Combining these two products in one solution provides WORM protection for archived data from end to end.

How it works

As shown in figure 1, applications connect to IBM Storage Protect for Data Retention through an Application Programming Interface (API) – the so called TSM API - to archive, retrieve and query for archived objects. While archiving an object obtains a retention period configured in the IBM Storage Protect for Data Retention server. An application cannot modify or delete an object if its retention period has not expired. Upon expiration of the retention period the application can delete the object but not modify its content.

IBM Storage Protect for Data Retention stores the archived objects in storage pools residing in file systems. In fact, IBM Storage Protect for Data Retention aggregates multiple archived objects into storage pool volumes that are stored as files in the storage pool file system. Thus, a storage pool-volume is a large file in the file system that holds many objects.

IBM Storage Protect for Data Retention supports different types of file systems for its storage pool volumes including file systems that prevent modification and deletion of files during the retention period. This capability is provided by IBM Storage Scale immutable filesets where files can be set to append-only or immutable with a retention period associated. When a file is set to append-only, then data can be appended but existing content cannot be modified or deleted. Immutable files cannot be appended, modified or deleted during the retention period.

Retention periods, append-only and immutable states of files in a IBM Storage Scale immutable fileset are controlled by IBM Storage Protect for Data Retention. When IBM Storage Protect for Data Retention allocates a new storage pool-volume in a IBM Storage Scale immutable fileset it creates an empty file in append-only mode. Archived objects are appended to the volume, and the retention period of the associated file is set to the longest expiration date of any object stored in the volume. A storage pool-volume has a fixed maximum size and when it is full then IBM Storage Protect for Data Retention makes the associated file immutable (WORM protected) and updates its retention period in accordance with the longest expiration date of any object stored in the volume.

IBM Storage Scale immutable filesets ensure that the underlying files of storage pool volumes cannot be deleted during the retention period. If a volume approaches the end of the retention period, then IBM Storage Protect for Data Retention reclaims the volume and deletes it.

The integration of IBM Storage Protect for Data Retention with IBM Storage Scale immutable fileset requires that the IBM Storage Protect for Data Retention server mounts the IBM Storage Scale file system containing the immutable filesets directly. Mounting the IBM Storage Scale immutable fileset via NFS or SMB does not work. In addition, it requires the use of IBM Storage Protect storage pools of the FILE device class.

Important Note: The integration of IBM Storage Protect for Data Retention with IBM Storage Scale immutable filesets requires the use of storage pools configured with the *FILE* device class. The storage pool data format must be set to *nonblock*. In addition, the IBM Storage Protect for Data Retention server must mount the IBM Storage Scale directly and not via NFS or SMB.

Special case: Reclamation

Reclamation is a process in the IBM Storage Protect server to manage the storage capacity used by storage pool volumes. A storage pool-volume is a container-file in the file system that contains many archived objects. If archived objects expire over time they still occupy space in the storage pool volume. This decreases the utilization of storage pool volumes and hence of the underlying file system. The reclamation process takes care of this by copying archive objects that have not expired to a new storage pool-volume and if the old storage pool-volume does not contain any non-expired data it is deleted from the file system. This way the storage utilization is increased because expired objects are eliminated. This is also useful when regulations require that data must be deleted at the end of the retention period.

The reclamation process for storage pool volumes stored in an immutable fileset is different because the storage pool-volume cannot be deleted from the file system until its retention period has expired. So even if almost all archived objects in a storage pool-volume have expired, the reclamation process will not start until the retention period of the last archived object has expired. For this purpose, IBM Storage Protect for Data Retention associates a reclamation period with such storage pool volumes that is 30 days before the final retention period of the volume expires. Within this reclamation period IBM Storage Protect for Data Retention will reclaim the volume.

There are cases where the actual retention period of a storage pool-volume cannot be determined because it contains archived objects that are bound to event-based management classes or have a legal hold. Such objects have an indefinite retention until an event is sent by the application via the API or the

legal hold has been released. Upon receiving the event the definite retention period of the object is calculated. Thus, a storage pool-volume can contain objects with indefinite retention periods. However, the retention period of the storage pool-volume cannot be indefinite because the volume could never be deleted. For this reason, IBM Storage Protect for Data Retention checks the actual retention period of all objects in each volume when it runs reclamation during the reclamation period. If the amount of space that could be reclaimed is higher than the reclamation threshold it will copy non-expired objects to a new storage pool-volume and once the retention period of the old volume has expired, it will delete the old volume. The retention period of the new volume is the maximum of the longest retention period of any object stored in the volume or the time specified with the parameter RETENTIONEXTENSION.

Configuration

In this section we describe how to configure the solution in figure 1 (see section [Solution overview](#)). First, we highlight some prerequisites, afterwards we explain the creation of an immutable fileset in IBM Storage Scale and the configuration of storage pools in IBM Storage Protect for Data Retention that use the immutability features provided by IBM Storage Scale.

Prerequisite

The server running the IBM Storage Protect for Data Retention software must be a member of the IBM Storage Scale cluster providing the immutable fileset. Alternatively, the server can be part of a separate IBM Storage Scale cluster mounting the IBM Storage Scale file system with the immutable fileset remotely through the IBM Storage Scale remote mount capability.

The version of IBM Storage Scale must be 5.0.1 or above. The version of IBM Storage Protect for Data Retention must be 8.1.8 or above.

The IBM Storage Protect for Data Retention software is installed and functioning.

The IBM Storage Scale cluster must be active, the node running IBM Storage Protect for Data Retention must be active and must have the file system mounted that provides the immutable fileset. In the guidance below the file system name is denoted by *fsname* and the path where the file system is mounted is denoted by *fspace*.

Creating IBM Storage Scale Immutable fileset

To create an immutable fileset in a IBM Storage Scale file system use the following command:

```
# mmcrfileset fsname fsetname [--inode-space new]
```

- *fsname* is the name of the file system
- *fsetname* is the name of the fileset
- `--inode-space` is optional. When set to `new` this will create an independent fileset. However, it does not have to be an independent fileset, so this parameter can also be omitted.

Now set the IAM-mode “compliant-plus” for this fileset. This mode includes some further adoptions of the SnapLock® method and must be used for the integration with IBM Storage Protect for Data Retention.

```
# mmchfileset fspath fsetname --iam-mode compliant-plus
```

Important Note: The integration of IBM Storage Protect for Data Retention with IBM Storage Scale immutable filesets only works with the fileset configured in compliant-plus mode.

Now link the fileset into the file system path that is later used as the storage pool directory by IBM Storage Protect for Data Retention.

```
# mmlinkfileset fspath fsetname -J junctionpath
```

-J *junctionpath* is the path of the fileset. This can be the file system path plus another directory name.

Finally check the fileset definition:

```
# mmlsfileset fspath fsetname --iam-mode
```

Creating and configuring the storage pool

Once the immutable fileset is created and linked within the IBM Storage Scale file system the IBM Storage Protect for Data Retention storage pool can be configured.

The first step is to create a device class of type FILE with the directory pointing to the immutable fileset in IBM Storage Scale:

```
SSAM> define devc devcname devt=FILE mountlimit=num maxcap=cap  
directory=junctionpath
```

- *devcname* is the name of the device class that must be unique in the server
- *devt* is the type of the device class and must be set to FILE.
- *mountlimit* denotes the maximum number of sessions that can concurrently access this device class. The actual number depends on the applications and how many sessions these create in parallel.
- *maxcap* is the maximum size of storage pool volumes created in this device class
- *directory* is the path of the immutable fileset that was configured with the command `mmlinkfileset` (see section [Creating IBM Storage Scale Immutable fileset](#)).

Note: the device class type must be set to FILE (*devt*=FILE). Container pools are not supported.

The next step is to create a storage pool using this device class:

```
SSAM> define stg poolname devcname pooltype=primary crcdata=yes  
maxscratch=num dataformat=nonblock | block RECLAMATIONTYPE=SNAPLOCK
```

- *poolname* is the name of the new storage pool
- *devcname* is the name of the device class created before. The device class must be of *devtype*=FILE.

- `pooltype` denotes the pool type. In this case it is a primary pool.
- `crcdata` specifies whether additional checksum should be calculated and stored with the data. This is recommended for IBM Storage Protect for Data Retention storage pools
- `maxscratch` defines the maximum number of storage pool volumes that this pool can allocate. It should be matched to the number of volumes defined with the `define stg` command and the size of the file system hosting the immutable fileset.
- `dataformat` denotes the format in which the data is stored in storage pool volumes. The `dataformat` must be set to `nonblock`.
- `reclamationtype=Snaplock` indicates that the storage pool volumes should be retention managed using the SnapLock semantics. This parameter makes the storage pool aware of the WORM capabilities of the underlying file system.

Important Notes: The device class name must refer to a device class of the device type *FILE*. The `dataformat` of the storage pool must be set to *nonblock*. If the `dataformat` is set to `block`, then write errors will occur during write operations to volumes stored in an immutable fileset. The affected volumes will be set to *Full* even though they have not used the entire capacity of the volume.

During the definition of the storage pool IBM Storage Protect for Data Retention performs some tests on the immutable fileset provided by IBM Storage Scale. If the storage pool definition fails, examine the return code and check that the immutable fileset in IBM Storage Scale has been created with IAM-mode “compliant-plus”.

Creating and configuring nodes, domains and archive copy groups

Once the storage pool with reclamation type SNAPLOCK has been created define domain, policy set, management class and archive copy groups.

There are two types of retention for archive copy groups: event-based retention and chronological retention. With **event-based** retention the application has the ability controlling the expiration of an object. When an object is stored in an event-based management class and copy group its retention period is indefinite. An event sent via the TSM API triggers a finite retention period according to the setting of the copy group parameters RETMIN and RETVER.

With **chronological** retention the object expiration time is fixed, according to the copy group parameter RETVER. The selection of the retention type for an archive copy group depends on the application.

Lastly assign nodes to the domain created before.

Further guidance

The IBM Storage Protect for Data Retention server has a special parameter RETENTIONEXTENSION that defines the amount of time (in days) to be added to the storage pool-volume retention period. This amount of time is added if the retention period of the storage pool-volume has expired but not all archived objects stored in the volume have expired. This can occur if archived objects are associated with event-based retention or have legal holds. With event-based retention or legal hold the archived object expires when the application sends an event or release through the API. If there has been no event or release sent for an object its retention period is unlimited. However, the retention period of a

storage pool-volume is never set to unlimited because it cannot be shortened. Instead, the retention period of a storage pool-volume is set to a finite period, usually derived from the maximum value of RETMIN and RETVER that is defined in the archive copy group. If this finite retention period has expired but not all objects in the volume have expired, then IBM Storage Protect for Data Retention adds the time specified by the parameter RETENTIONEXTENSION to the retention period of the storage pool-volume.

The default setting of this parameter is 365 days. This is a reasonable default. To adjust this parameter, use the command: `SETOPT RETENTIONEXTENSION days`.

Administration

Monitoring of IBM Storage Protect for Data Retention and IBM Storage Scale is done independently. It is recommended to configure event notifications. This can be done for IBM Storage Scale using the GUI and for IBM Storage Protect for Data Retention using operations center.

If more capacity is needed for the file system hosting the immutable fileset, then additional capacity can be added concurrently to the IBM Storage Scale file system. It is recommended to restripe the data blocks on disk using the appropriate IBM Storage Scale commands (`mmadddisk` or `mmrestripefs`).

Software updates to the solution are non-concurrent to the IBM Storage Protect for Data Retention server software. When upgrading the IBM Storage Scale software then the IBM Storage Scale and IBM Storage Protect for Data Retention software must be stopped on this server. Likewise, software upgrades for IBM Storage Protect for Data Retention require to stop the IBM Storage Protect for Data Retention software.

The storage pool used to store data in the immutable fileset can be backed up by IBM Storage Protect for Data Retention, using the `backup stgp` command. Do not use the IBM Storage Scale backup function (`mmbackup`) to backup IBM Storage Protect for Data Retention storage pool volumes, this may result in inconsistencies. Likewise, archived objects stored in storage pools in immutable filesets can be migrated to other storage media like tapes. Again, use the IBM Storage Protect for Data Retention functionality (`migrate stgp` and `NEXTPOOL` configuration) and not the IBM Storage Scale ILM functions. When using tapes for backup and migration it is recommended to use IBM WORM tapes.

Appendix

References

[1] IBM Storage Protect for Data Retention introduction
<https://www.ibm.com/support/pages/node/3125835>

[2] IBM Storage Protect for Data Retention version 8.1 assessment report
https://kpmg.bryter.io/s/wEstKqupRqOpUAW27LPLHw/atlas_automation_report_service_01_en?id=a5bbaca-c60b-4b43-9a68-4641699ad6f3&var=a

[3] Introduction and configuration guidance for IBM Storage Scale Immutable fileset
<http://www.redbooks.ibm.com/abstracts/redp5507.html>

[4] IBM Storage Scale version 5.0 assessment report
https://kpmg.bryter.io/s/wEstKqUpRqOpUAW27LPLHw/atlas_automation_report_service_01_en?id=dfb70983-f63d-410d-94eb-2566b06918e1&var=a

Disclaimer

This document reflects the understanding of the author regarding questions asked about archiving solutions with IBM hardware and software. This document is presented “As-Is” and IBM does not assume responsibility for the statements expressed herein. It reflects the opinions of the author. These opinions are based on several years of joint work with the IBM Systems group. If you have questions about the contents of this document, please direct them to the Author (nils_haustein@de.ibm.com).

This document provides guidance for certain configuration and operational aspects. IBM does not guarantee that this guidance complies with laws or regulation. To obtain a compliance assessment an independent auditor must be engaged by the client. IBM cannot be made liable for any findings or violations of laws and regulations. The software nature of the solution may allow malicious hackers to exploit the system and elevate privileges of users.

The guidance given herein does not imply warranty that the commands given, or their intention satisfies the purpose. IBM cannot be made liable upon damage caused by any of the commands or guidance.

The following terms are trademarks or registered trademarks of the IBM Corporation in the United States or other countries or both: IBM, IBM Storage Scale and IBM Storage Protect.

SnapLock® is a registered trademark of NetApp Inc. in the United States and other countries.

Microsoft® Windows® is a registered trademark of Microsoft Corporation in the United States and other countries.

Other company, product, and service names may be trademarks or service marks of others.