IBM Spectrum Storage

Deploying a Solution that Combines IBM Spectrum Protect for Space Management with IBM Spectrum Scale

Document version 1.4

Nils Haustein (EMEA Storage Competence Center)
Dominic Müller-Wicke (IBM Spectrum Protect Development)

Contents

1 Introduction

You can deploy a data-protection solution that combines IBM Spectrum® Protect for Space Management with IBM Spectrum Scale. Follow the instructions to install, configure, and manage the solution.

IBM Spectrum Scale is a clustered parallel file system that can be used to store files on the most appropriate storage tier based on Information Lifecycle Management (ILM) policies. IBM Spectrum Protect for Space Management can be used as a storage tier within IBM Spectrum Scale. From the storage tier, you can migrate files to IBM Spectrum Protect servers. Thus, the integration of IBM Spectrum Protect for Space Management with IBM Spectrum Scale makes it possible to migrate files from internal IBM Spectrum Scale storage tiers to IBM Spectrum Protect servers, where the files can be stored on tape.

After briefly describing the key ILM concepts in IBM Spectrum Scale and the integration with IBM Spectrum Protect for Space Management, we provide guidance for installing and configuring IBM Spectrum Protect for Space Management in an IBM Spectrum Scale cluster. Subsequently, we provide guidance for operations that can be performed with IBM Spectrum Protect for Space Management. (We will not focus on ILM policies in detail; for more information about ILM, see Item 17 in the References.)

1.1 Target audience and scope

This publication is intended for IBM Spectrum Scale administrators with a basic understanding of IBM Spectrum Protect. High-level instructions are provided for installation, configuration, and solution management, but these instructions are intended only as an example of how to set up and run the system. The authors do not guarantee that this implementation will be appropriate for your system environment. Results can vary, depending on many factors. For official documentation, see the IBM Spectrum Protect for Space Management product documentation in IBM Knowledge Center.

1.2 Conventions

The following conventions are used in command syntax examples:

- `[options]` represents options that can be specified for the command. For more information about available options, see the command reference.
- Italicized words and phrases represent variables. For example, *`server_port`* represents a value that you enter for the server port.
- Updates in scripts are represented by bold, italic text.

1.3 IBM Spectrum Scale

When you configure IBM Spectrum Scale, you can implement storage tiers as storage pools within a file system as shown in Figure 1.

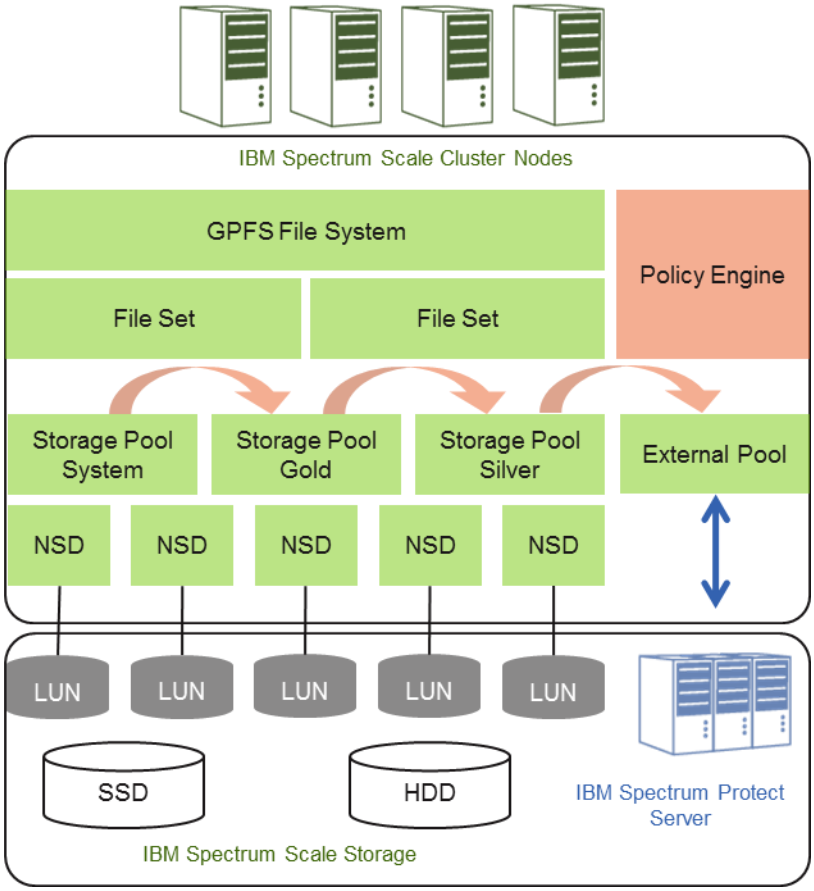

Figure 1. Basic concept of an IBM Spectrum Scale file system and pools

The IBM Spectrum Scale *file system* represents a global name space over all IBM Spectrum Scale cluster nodes that are members of the cluster. The file system is comprised of file system *pools*. The example in Figure 1 shows three file system pools named system, silver, and external. Any IBM Spectrum Scale file system has at least one pool that is named "system." The file system is comprised of data (such as files) and metadata (such as inodes) that are stored in pools. *Inodes* are data structures that include file metadata such as time stamps, attributes, and pointers to the data blocks of the file. Each file and directory has one inode. Metadata can be stored *only* in the "system" pool, while data can be stored in any pool.

IBM Spectrum Scale includes two kinds of file system pools:

- An *internal pool* is a collection of disks (hard disks or solid-state drives, SSDs) with similar properties that are managed together. There must be a minimum of one internal pool (the system pool) and there can be up to eight internal pools per file system. Placement and migration can be done on internal pools.
- An *external pool* is an external storage device that is represented and managed by an interface script. An external pool does not require disk storage; the pool can also be based on tape or other storage devices. Storing and retrieving data (files) in an

external pool is managed by the interface script. IBM Spectrum Protect for Space Management provides an interface script and can be configured as external pool.

An *NSD* is the representation of a storage logical unit number (LUN) or volume provided by a storage system. An NSD is dedicated to a pool. All NSDs in one internal pool are comprised of storage LUNs with similar characteristics. Referring to Figure 1, the NSDs of the system pool are configured on LUNs from SSD drives, and the NSDs of the silver pool are configured on LUNs from Nearline SAS (NL-SAS) drives. Each file system pool has its own input/output (I/O) characteristics.

Also shown in Figure 1 are *file sets*. A file set is a logical subtree of the file system that is managed as a logical partition within the file system. Snapshots and quotas can be configured on a fileset level. In the context of ILM, file sets can be used to control placement and migration. A file set is identified by the fileset name and the associated path name within the file system. Placement and migration of files in a file set can be controlled based on the fileset name or path name in conjunction with other file attributes such as file size, file type, access time, etc.

The *IBM Spectrum Scale policy engine* can identify files based on programmable criteria and conditions (policies and rules) and manage their lifecycle. Management of files includes:

- Placement of files when they are created
- Migration of files during the lifecycle
- Listing files
- Encrypting files
- Deleting files

In this document, we focus on policies for migration in combination with IBM Spectrum Protect for Space Management. A policy is a set of rules, such as migration rules. A *migration rule* can describe the migration of all or a subset of files from one file system pool to another file system pool. The selection of files for migration is based on file attributes. Both file system pools must belong to the same file system. The file system pool can be an internal pool or an external pool. A migration rule is run by the policy engine (`mmapplypolicy` command or callback).

1.4 IBM Spectrum Protect for Space Management

As previously described, IBM Spectrum Protect for Space Management integrates with IBM Spectrum Scale by providing an external pool for migration. In fact, the IBM Spectrum Scale policy engine invokes the IBM Spectrum Protect for Space Management interface script that manages migration of files from internal pools to IBM Spectrum Protect and vice versa (see Interface script). IBM Spectrum Scale file systems can be selectively enabled for space management.

Figure 2 shows the general architecture of IBM Spectrum Protect for Space Management integrated with IBM Spectrum Scale.

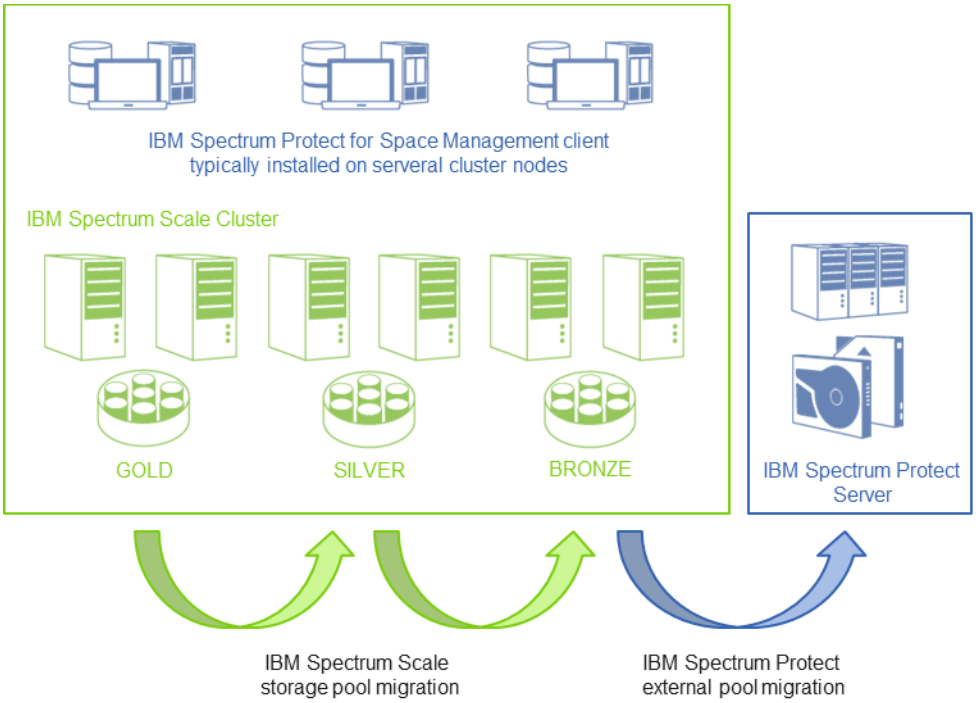

Figure 2. IBM Spectrum Protect for Space Management architecture

As shown in Figure 2, the IBM Spectrum Protect for Space Management client (HSM client) is installed on one or more cluster nodes that have access to the space-managed file systems. IBM Spectrum Protect for Space Management completes space management operations such as migration, premigration, and recall.

During *migration,* IBM Spectrum Protect for Space Management moves the content of a file to the IBM Spectrum Protect server and leaves a stub file in the file system. The stub file is essentially represented by the inode of the file. The inode of a stub file includes special metadata for IBM Spectrum Protect for Space Management such as an object identifier that uniquely identifies each migrated file in the IBM Spectrum Protect server. It also contains the managed regions that indicate which parts of the file are migrated to IBM Spectrum Protect. Migrated files are stored within a storage pool of the IBM Spectrum Protect server. The storage pool can be located on any type of storage system, including SSD, disk, tape, or cloud. For more details about running migration, see Migration.

After a file is migrated, it is still visible in the file system. When the file is accessed, it is transparently *recalled* by IBM Spectrum Protect for Space Management. It thereby intercepts the file-open request, reads the metadata information from the inode, retrieves the file from the IBM Spectrum Protect server, stores it in the internal file system pool, and releases the file-open request for user access. The file is essentially copied from the IBM

Spectrum Protect server to the file system pool from which it was migrated. For more details about running migration, see Recall.

Files can also be *pre-migrated*, which makes the file dual resident: in the internal pool and the external pool represented by IBM Spectrum Protect. For more details about running migration, see Premigration.

In accordance with the space management operations, files in a space-managed file system can have different states: resident, migrated, or premigrated. For details, see File HSM states.

2 Installation and configuration

Guidance is provided for installing and configuring IBM Spectrum Protect for Space Management on a Red Hat Enterprise Linux (RHEL) based IBM Spectrum Scale cluster.

2.1 Prerequisites

We provide guidance to help you prepare for the installation and configuration of the IBM Spectrum Protect for Space Management client on IBM Spectrum Scale cluster nodes. We assume that the IBM Spectrum Scale cluster already exists and that file systems were created on internal pools. We also provide guidance for solution design.

For detailed information about the IBM Spectrum Protect for Space Management client, see the IBM Spectrum Protect for Space Management product documentation in IBM Knowledge Center.

2.1.1 Solution design guidance

The IBM Spectrum Protect for Space Management client can be installed and configured on multiple IBM Spectrum Scale clusters, whereby all enabled nodes actively participate in space management operations. In fact, depending on the number of nodes that are enabled for space management, the speed of these operations can be improved because the nodes work in parallel.

The IBM Spectrum Scale node where the IBM Spectrum Protect for Space Management client is installed can be an NSD server node or an NSD client. The node can also run another IBM Spectrum Scale service such as Cluster Export Services (CES), providing access via NFS, SMB, and object storage APIs (Swift and S3). The preferred method is to install the client on dedicated IBM Spectrum Scale client nodes with a dedicated network connection to the IBM Spectrum Protect server (in other words, do not use the cluster network for communication with the IBM Spectrum Protect server). To optimize the performance of IBM Spectrum Protect for Space Management, provide a reasonable amount of random access memory (RAM) on the IBM Spectrum Scale node that is used for data buffering (read ahead and write behind by using the IBM Spectrum Scale pagepool).

The network connection between the IBM Spectrum Protect for Space Management client and the IBM Spectrum Protect server should support migration and recall of the daily data volume within the required time frame. If the network between the space management enabled nodes and the IBM Spectrum Protect server becomes a bottleneck, separate networks can be configured between the space management nodes and the IBM Spectrum Protect server. If the IBM Spectrum Protect server becomes the bottleneck, multiple IBM Spectrum Protect servers can be configured to receive the space management data.

The IBM Spectrum Protect server should be sized according to the Blueprint guidelines (see Item 1 in References), taking into consideration the daily workload and the total managed capacity.

An important factor for sizing the solution is the daily volume of data to be migrated, premigrated, and recalled. These operations cause additional workloads in the IBM Spectrum Scale cluster, on the network connection to the IBM Spectrum Protect server, and within the IBM Spectrum Protect server. For example, when a file is migrated, it must be read from the file system internal pool, transferred over the network to the IBM Spectrum Protect server, and be processed by the IBM Spectrum Protect server, whereby metadata is extracted and the file is stored.

2.1.2 Preparation

The IBM Spectrum Scale file systems that must be space managed should exist already, but the Data Management Application Programming Interface (DMAPI) should *not yet* be enabled (file system parameter `-z no`). The file system *should not* be configured for automatic mount (file system parameter `-A automount`) because this is not supported by the IBM Spectrum Protect for Space Management client.

The IBM Spectrum Scale nodes that are running the IBM Spectrum Protect for Space Management client software should be members of the cluster and have access to the space-managed file system and a network connection to the IBM Spectrum Protect server. For the installation of the client root, access is required on the IBM Spectrum Scale nodes.

The IBM Spectrum Protect server should already be installed and configured. When multiple IBM Spectrum Scale nodes have the IBM Spectrum Protect for Space Management software installed, configure a proxy node setup with one proxy node per space management client (see IBM Spectrum Protect server configuration).

Ensure that the IBM Spectrum Protect for Space Management version supports the IBM Spectrum Scale version in the cluster (see Item 2 in References). Also ensure that the IBM Spectrum Protect for Space Management client version support the IBM Spectrum Protect server version (see Item 3 in References).

Ensure that space management operations (migration to IBM Spectrum Protect) are not configured for Cluster Export Services (CES) shared root directories.

For more information about preparation and prerequisites, see Item 7 in References.

2.2 Installation

The steps for installing the IBM Spectrum Protect for Space Management client on RHEL are summarized. For detailed installation steps, see Item 7 in References.

1. Download the IBM Spectrum Protect for Space Management client (see Item 4 in References). Copy the package to all IBM Spectrum Scale nodes where the client must be installed. Unpack the packages on all of these nodes.

2. On the cluster nodes where the client is to be installed, set the HSM install mode to `SCOUTFREE`. The scout daemon checks the occupation in the file system and, if a certain threshold is reached, starts migration. The scout daemon is not required with IBM Spectrum Scale because IBM Spectrum Scale implements a different method to start the migration when the file system pool occupation reaches a specified threshold.

   ```
   # export HSMINSTALLMODE=SCOUTFREE
   ```

3. To ensure that the installation mode is set correctly for subsequent upgrades, add the following statement to the root user profile, for example, to `~/.bash_profile`:

   ```
   # export HSMINSTALLMODE=SCOUTFREE
   ```

4. Install the IBM Spectrum Protect for Space Management client packages.

   ```
   # yum install TIVsm-API64.x86_64.rpm TIVsm-BA.x86_64.rpm
   TIVsm-HSM.x86_64.rpm gskcrypt64-8.0.50.52.linux.x86_64.rpm
   gskssl64-8.0.50.52.linux.x86_64.rpm
   ```

5. Check the packages that are being installed by using the following command:

   ```
   # rpm -qa | grep -E "TIV|gskcrypt|gskssl"
   ```

2.3 Configuration

The configuration steps include the configuration of the IBM Spectrum Scale file system, the IBM Spectrum Protect server, and the IBM Spectrum Protect for Space Management client. For detailed instructions, see Item 8 in References.

2.3.1 Operating system configuration

To configure the operating system, complete the following steps:

1. Verify whether the `rpcbind` process is running by issuing the following command:
   ```
   # systemctl status rpcbind
   ```

2. If the `rpcbind` process is not running, verifying whether it is installed by issuing the following command:
   ```
   # rpm -qa | grep rpcbind
   ```

3. If the `rpcbind` process is not installed, install it by issuing the following command:
```
# yum install rpcbind
```

4. Start and enable automatic start for the `rpcbind` process:
```
# systemctl start rpcbind
# systemctl enable rpcbind
```

2.3.2 IBM Spectrum Scale configuration

Enable the DMAPI for all file systems that are space managed. Complete the following steps:

1. Unmount the file systems on all nodes:
```
# mmumount fsname
# mmchfs fsname -z yes
# mmmount fsname
```

2. Verify and configure the following IBM Spectrum Scale parameters:
   - `workerThreads:` For guidelines about setting this parameter, see the IBM Spectrum Scale online product documentation.
   - `pagepool:` Set to 16 GB or more, depending on the file size.
   - `maxFilesToCache:` For small files, use a larger value. For guidelines about setting this parameter, see the IBM Spectrum Scale online product documentation.
   - `dmapiWorkerThreads:` Set to a maximum value of 64.
   - `dmapiEventTimeout:` If Network File System (NFS) export is configured, set the value to 1000. If NFS export is not configured, keep the default value.

2.3.3 IBM Spectrum Protect server configuration

On the IBM Spectrum Protect server, a policy domain must be configured with a space management class. Consider the following parameters:

- `SPACEMGTECHnique` should be set to `AUTO`.
- `AUTOMIGNOnuse` should be set to `0`.
- `MIGREQUIRESBkup` should be set to `yes` if the `mmbackup` command is used. In this way, you can ensure that files are migrated only when they have been backed up, to avoid recalls upon backup. If the `mmbackup` command is not used, this parameter should be set to `no`.
- `MIGDESTination` should specify the name of the storage pool that is used to store migrated files.

For each IBM Spectrum Scale node that is running the IBM Spectrum Protect for Space Management client, a proxy node name must be registered that is bound to the space

management domain (by using the `REGISTER NODE` command). In addition, each target node name must be bound to the same management class. And a proxy node relation must be defined that binds all proxy nodes to the target node (by using the `GRANT PROXYNODE` command).

Also obtain the IP address and the port number that are configured for the server by using the following command:

```
query status
```

2.3.4 IBM Spectrum Protect for Space Management client configuration

The installation of the IBM Spectrum Protect for Space Management client installs default `dsm.sys` and `dsm.opt` files, which must be edited. These files are in the `/opt/tivoli/tsm/client/ba/bin` directory.

In the `/opt/tivoli/tsm/client/ba/bin/dsm.opt` file, add the following parameters:

```
  HSMDISABLEAUTOMIGDAEMONS YES
  HSMGROUPEDMIGRATE YES
  HSMEXTOBJIDATTR YES
```

If the IBM Spectrum Protect for Space Management client manages a file system that includes a file set that is configured for IBM Spectrum Scale Active File Management (AFM), add the following option:

```
  AFMSKIPUNCACHEDFILES YES
```

If you use IBM Spectrum Protect for Space Management with the IBM Spectrum Scale backup function in combination with the IBM Spectrum Protect backup-archive client, see Client options for backup for more detailed guidance.

In the `/opt/tivoli/tsm/client/ba/bin/dsm.sys` file, create a server stanza with the following parameters:

```
  SErvername  hsmserver
    COMMMethod         TCPip
    TCPPort            server_port
    TCPServeraddress   server_address
    nodename           proxy_node
    asnodename         target_node
    passwordaccess     generate
    errorlogname
/opt/tivoli/tsm/client/ba/bin/dsmerror.log
    errorlogretention  5 d
    HSMLOGNAME         /tmp/hsm/hsm.log
    HSMLOGEVENTFLAGS   FILE
```

The values for *server_port*, *server_address*, *proxy_node*, and *target_node* are obtained from the IBM Spectrum Protect server configuration. Consider using the NODENAME and ASNODENAME options if multiple IBM Spectrum Scale cluster nodes are configured for IBM Spectrum Protect for Space Management. In this case, each node has an individual node name, and all use the ASNODENAME value to access the same data stock in the IBM Spectrum Protect server.

Optionally, you can use the ERRLOGNAME, HSMLOGNAME, and HSMLOGEVENTFLAG options to specify the files in which logs are stored.

Now add the file systems to be space managed to the IBM Spectrum Protect for Space Management configuration by using the command:

```
# dsmmigfs add [options] fspath
```

The parameter *fspath* is the file system path. You can find the value by reviewing the -T parameter of the file system (`mmlsfs fsname -T`).

The IBM Spectrum Protect for Space Management command dsmmigfs has multiple options that can be used to adapt the requirements of your system. Use the following command to get a list of all options:

```
# dsmmigfs help
```

The following subset of options is commonly used for configuration and should be considered:

- The quota option specifies the maximum number of megabytes of data that you can migrate and premigrate from the file system to the IBM Spectrum Protect server. The value defaults to the available file system capacity. However, the migration of files from the file system to IBM Spectrum Protect frees up space, so potentially much more data will be migrated to IBM Spectrum Protect. Set the -Quota value to the anticipated total capacity in the file system and in the IBM Spectrum Protect server.

- The inlinecopymode option must be considered only when running IBM Spectrum Protect backup-archive progressive incremental backup operations or mmbackup operations for the same file system. The option specifies how a backup operation handles a migrated file. If the inlinecopymode option is set to MIG and a migrated file is a candidate for a backup operation, the file is cloned inside the space management storage pool in the IBM Spectrum Protect server. The clone of the file will be stored in the backup storage pool in the IBM Spectrum Protect server. This operation is called inline copy.
  Restriction: IBM Spectrum Protect container pools (cloud and disk) do not support cloning. In this case, the inlinecopymode option must be set to NO.

- The stubsize option specifies the size of stub files that remain on the file system when files are migrated to storage. For the space management client on IBM Spectrum Scale file systems, you can specify 0 or a multiple of the file system

block size. The default value is 0. For all file system types, the maximum value for a stub file size is 1 GB. This option is important if you have applications that use the start sequence of a file to generate preview icons of the file to prevent unwanted recall activity.

Certain space management-related parameters of the file system can be changed later by using the `dsmmigfs update` command. For more information, see Item 9 in the References.

3 Space management operations

Internals regarding IBM Spectrum Protect for Space Management are discussed. You can learn how certain operations work and how these operations are executed and monitored.

Details of IBM Spectrum Scale ILM policies are not provided; for information about ILM policies, see Item 17 in References.

3.1 Internals

Learn about fundamental behaviors of the IBM Spectrum Protect for Space Management client, including how it uses an interface script, how file states change, and how file attributes work.

3.1.1 Interface script

As explained in IBM Spectrum Protect for Space Management, the space management client integrates with IBM Spectrum Scale via an interface script. The interface script is defined in an external pool rule and is invoked by the policy engine when this external pool is used as a destination for migration in a `MIGRATE` rule. The interface script performs the migration operation.

A default interface script is stored in the following file:
`/usr/lpp/mmfs/samples/ilm/mmpolicyExec-hsm.sample`

The script is essentially a shell script. If you use the default interface script in an external pool rule, the preferred method is to rename it and store it in the following directory:
`/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.migrate`

Ensure that the script is executable.

An external pool rule defines the external pool and the interface script that is invoked with it, as shown in the following example:

```
RULE EXTERNAL POOL 'hsm' EXEC
'/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.migrate'
OPTS '-v'
```

A migration rule defines the selection criteria for files to be migrated and the destination pool for migration, which might be an external pool. In the following example, the pool is named HSM:

```
RULE 'migration' MIGRATE FROM POOL 'system' TO POOL 'hsm'
WHERE
NOT (MISC_ATTRIBUTES LIKE '%M%')
```

The migration rule essentially migrates all files residing in the SYSTEM pool to the HSM pool, where the MISC_ATTRIBUTES does not contain the character M. The MISC_ATTRIBUTES is an extended attribute that encodes the HSM state of the file (see File attributes). In the previous example, this attribute means that all files that are not migrated are selected by this rule. The HSM pool is the external pool that is managed by the following interface script:
`/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.migrate`

The default interface script can also be adjusted, at your own risk, to facilitate specific operations such as premigration and recall (see Premigration and Recall).

3.1.2 File HSM states

Depending on the space management operation (migration, recall, or premigration), different file states are possible, as shown in Figure 3.

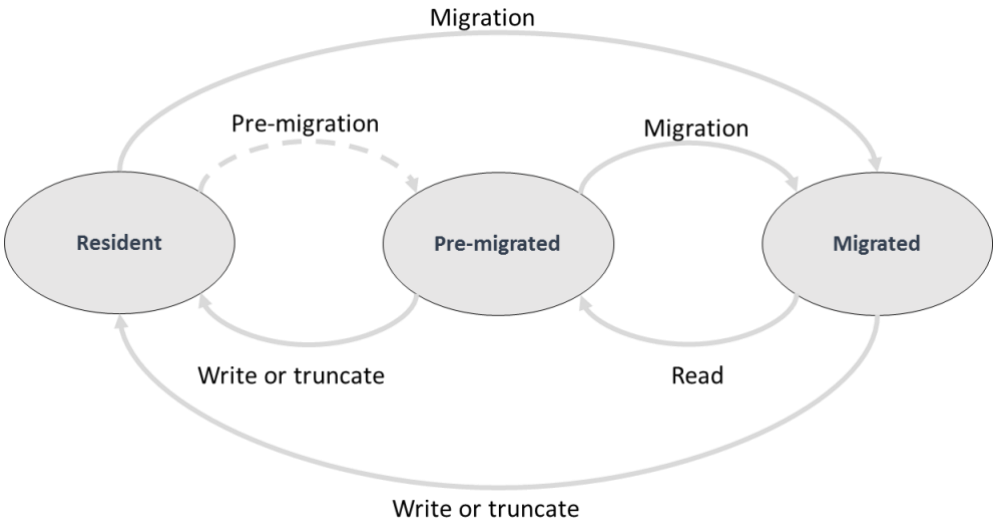

Figure 3: File states in a space managed file system

Figure 3 shows that a file can have three file states in a space-managed file system, whereby a file has only one state at a time. When a file is created in the internal file system pool, its state is *resident*. Resident means that the file is not space managed. If the file is moved to IBM Spectrum Protect, the file state is *migrated*. Migrated means that the file content is in the IBM Spectrum Protect server and a stub is left in the IBM Spectrum Scale file system. If the file is copied to IBM Spectrum Protect, the file state is *premigrated*. Premigrated means that the file is dual resident: in the file system and the IBM Spectrum Protect server. If a migrated file is read, it is recalled and changes to the premigrated state. If a migrated or premigrated file is modified, it might be recalled and its state changes to resident.

The file states are reflected in the extended attributes of files. For more details about file attributes associated with IBM Spectrum Protect for Space Management, File attributes.

3.1.3 File attributes

The IBM Spectrum Protect for Space Management client encodes information about the HSM state of the file in extended attributes. Each file or directory has extended attributes that are stored in the file inode. This means that, even if a file has been migrated, the metadata is still resident in the file system including the extended attributes.

To store HSM state information for a file, the MISC_ATTRIBUTES data structure is used (for details, see Item 10 in References). The following SQL expressions can be used to determine the HSM states:

Resident	MISC_ATTRIBUTES NOT LIKE '%M%'
Premigrated	MISC_ATTRIBUTES LIKE '%M%' AND MISC_ATTRIBUTES NOT LIKE '%V%'
Migrated	MISC_ATTRIBUTES LIKE '%V%'

The file status information (command: `stat`) encoded in the file system can also be used to determine the file state. In particular, the file size and allocated blocks determine the file state:

Resident	File size and allocated blocks match.
Premigrated	File size and allocated blocks match.
Migrated	File size is identical and the allocated blocks are 0.*

* The allocated blocks are 0 if the stub size is 0. Otherwise, the allocated blocks may match the stub size.

In addition, the IBM Spectrum Scale `mmlsattr` command can be used to list IBM Spectrum Protect for Space Management file attributes, as shown in the following example:

```
#mmlsattr -d -L --hex-attr mig_file | grep dmapi.IBM
dmapi.IBMObj:      0x010...000
dmapi.IBMexID:          0x303...000

#mmlsattr -d -L --hex-attr premig_file | grep dmapi.IBM
dmapi.IBMPMig:          0x010…E68
dmapi.IBMexID:          0x303…800
```

The attribute *IBMObj* indicates that the file is migrated or in transition. The attribute *IBMPMig* indicates that the file is in premigrated state. The attribute *IBMexID* reflects the unique identifier of the file within the IBM Spectrum Protect server. This attribute is required in combination with the file name when recreating stub files (see Recreating stub files).

3.2 Migration

During migration operations, files are migrated from an internal file system pool to an external pool represented by IBM Spectrum Protect. At the end of the migration process,

the file content resides in the IBM Spectrum Protect server while the file metadata (inode) is still present in the file system, and the HSM processing does not change the inode information. So, from a user perspective the file appears to be present in the file system with the same size as before the migration.

The IBM Spectrum Protect for Space Management configuration option `stubsize` can be used to control how much data of the file remains in the stub stored in the internal pool after migration. For example, if the `stubsize` option is set to 1 MB, for each migrated file, the first 1 MB of data is kept in the stub. Of course, this consumes additional capacity in the file system that must be accounted for. The `stubsize` value can be changed by using the following command (see also Item 9 in References):

```
# dsmmigfs update fspath -stubsize=size
```

In general, file migration can be driven by commands or the policy engine.
If the value of the `stubsize` option is changed by using the `dsmmigfs` command, the change has a global effect on all subsequent migration in the given file system. The stub size of a file that is already migrated with a larger stub size can be reduced to a smaller stub size by remigrating it after changing the stub size settings of the file system.

3.2.1 Command line

You can use the following basic command syntax for migration:

```
# dsmmigrate  [options] filespec | filelist=list
```

The options control the migration. The following list includes some of the available options (for more details, see Item 11 in References):

- `Server` IBM Spectrum Protect target server
- `Logname` Name of log file
- `Premigrate` Premigration state
- `Recursive` Recursively migrate files from the given path (*filespec*)

The files to be migrated are specified by a path name pattern (*filespec*) or a list of files (`filelist=`*`list`*) containing one fully qualified path and file name per line.

3.2.2 Policy driven

More commonly, the IBM Spectrum Scale policy engine is used for migration. As further explained in Item 17 in the References, migration policies can be active or scheduled. Active migration policies are activated in a file system and are triggered when a file system pool reaches a defined occupation threshold. Scheduled migration policies run periodically as initiated by a scheduler.

The general recommendation is to use scheduled migration policies to balance the file system pool occupation and to use active migration policies as last line of defense, in case the file system runs full. In other words, design and size the system in a way that active migration policies are never triggered. The reason for avoiding active migration policies is

that they are triggered when the pool occupation threshold is met, which can impact production workloads. With scheduled policies, the migration can be planned to occur at times when it does not impact production workloads.

Active migration is typically based on a `THRESHOLD` value, which is configured in a `MIGRATE` rule. The migration is invoked automatically when the occupation threshold for an internal file system pool is reached. The following example illustrates an active migration policy with IBM Spectrum Protect for Space Management where the SYSTEM pool is migrated to IBM Spectrum Protect when the pool reaches 90% occupation.

```
/* define exclude rule*/
RULE 'exclude' EXCLUDE WHERE (PATH_NAME LIKE '%/.SpaceMan/%'
OR
PATH_NAME LIKE '%/.snapshots/%')

/* here comes the migration rules from system to hsm*/
RULE EXTERNAL POOL 'hsm'
EXEC '/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-
hsm.migrate'
OPTS '-v'

RULE 'autoMig' MIGRATE FROM POOL 'system' THRESHOLD(90,70)
TO POOL 'hsm' WHERE (KB_ALLOCATED > 0)

/* here comes the placement rules for silver */
RULE 'default' SET POOL 'system'
```

The first rule is an `EXCLUDE` rule that excludes paths that shall not be migrated. All file names matching these patterns are excluded from processing with subsequent rules. In this example, it is important to exclude files stored in `.SpaceMan` and `.snapshots`.

The second rule defines the `EXTERNAL POOL` denoted by the IBM Spectrum Protect for Space Management interface script for migration.

The third rule is the `MIGRATE` rule that migrates all files that are allocated more than 0 KB. If the `stubsize` option is set to 0, files with a size of 0 KB are either empty files or files that have already been migrated.

The last rule is a `PLACEMENT` rule that must be included in an active policy.

Active migration policies as shown above must be activated in the file system by using the following command:

```
# mmchpolicy fsname policyfile -I yes | test
```

In addition, a callback must be configured to catch low disk space events indicating that a file system pool has reached its high occupation level (for more details, see Callback).

Scheduled migration is run manually or automated by a scheduler. It is typically independent of thresholds. With scheduled migration, ILM requirements independent of the storage pool occupation can be implemented, such as migration based on file age or other criteria. In the following example, the scheduled policy migrates files that have not been accessed for more than 30 days:

```
/* define macros */
define(access_age,(DAYS(CURRENT_TIMESTAMP) -
DAYS(ACCESS_TIME)))

/* define exclude rule*/
RULE 'exclude' EXCLUDE WHERE (PATH_NAME LIKE '%/.SpaceMan/%'
OR
PATH_NAME LIKE '%/.snapshots/%')

/* here comes the migration rules from system to hsm*/
RULE EXTERNAL POOL 'hsm'
EXEC '/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-
hsm.migrate' OPTS '-v'

RULE 'autoMig' MIGRATE FROM POOL 'system' TO POOL 'hsm'
WHERE (KB_ALLOCATED > 0) AND (access_age > 30)
```

This policy can now run periodically (based on the schedule) by using the following command:

```
# mmapplypolicy fsname -P policyfile -N hsmnodes [-B 1000 -m
3]
```

For more information about the optional parameters of the `mmapplypolicy` command, see Policy engine parameters. For more information about automating and scheduling policies, see Automation and Scheduling.

The policy engine must be started on a node that has the IBM Spectrum Protect for Space Management client installed. For more information about migration policies with IBM Spectrum Scale, see Item 17 in References.

A special form of migration is premigration. With a premigration policy, files are copied to an external pool. Thus, the file is dual resident, in the internal and external pool. Premigration works only when migrating files from an internal to an external pool (see Premigration).

3.2.3 Callback

A callback essentially catches one or more IBM Spectrum Scale events and executes a customizable script. A migration callback is required when an active policy is configured. The active policy defines an occupation threshold for an internal file system pool. If this

threshold is reached, the migration policy is started. To do this, IBM Spectrum Scale raises the event `lowDiskSpace` or `noDiskSpace`. The callback catches these events and runs a customizable script with a set of parameters. The customizable callback script starts the active migration policy for the file system. Such a script is available within IBM Spectrum Scale and is named `/usr/lpp/mmfs/bin/mmstartpolicy`.

Consequently, the command to create a callback defines the events to be caught (`--event lowDiskSpace,noDiskSpace`) and the customizable script to be invoked (`--comand /usr/lpp/mmfs/bin/mmstartpolicy`) and the parameter to be passed to the script (`--param "%eventName %fsName"`). Here is an example of the command creating a callback for active migration:

```
# mmaddcallback MIGRATION --command
/usr/lpp/mmfs/bin/mmstartpolicy
--event lowDiskSpace,noDiskSpace
--parms "%eventName %fsName"
```

3.3 Premigration

During premigration operations, files are copied from an internal file system pool to an external pool represented by IBM Spectrum Protect. At the end of the premigration process, the file content is dual resident: in the internal file system pool and in the IBM Spectrum Protect server.

Restriction: Do not use premigration for disaster recovery, unless you combine premigration with the "Scale out Backup and Restore" function of IBM Spectrum Scale (see Integration with Scale out Backup and Restore). Even though two copies of the premigrated file exist, if the copy in the internal pool disappears, there is no trivial way to recover the copy from the external pool.

Premigration can be used before migrating files. The migration of a premigrated file is fast because the file just needs to be stubbed. For example, if files are being migrated when they have not been accessed for 30 days, the files could be premigrated immediately, so the migration step after 30 days is quick. However, if the file changes several times within the 30 days, there is no point in premigrating files. So, this solution is appropriate for archives, where files do not change after creation.

3.3.1 Command line

Premigration can be driven with the following command, which is described in more detail in Item 11 in References:

`# dsmmigrate --premigrate` *`fspath`* `or` *`filelist`* `options`

The meaning of the parameters and options is identical to the `dsmmigrate` command shown in Migration.

3.3.2 Policy driven

You can implement premigration policies by using threshold-based policies or an adjusted interface script.

Using THRESHOLD

You can use a THRESHOLD command to define a premigration rule. The THRESHOLD clause allows for three values: `THRESHOLD(%high, %low, %premig)`

%high: The rule is applied only if the occupancy percentage of the source pool is greater than or equal to the `%high` value.

%low: If the occupancy percentage of the source pool is greater than the `%low` value, files are migrated until the `%value` is met.

%premig: Defines an occupancy percentage of a storage pool that is below the lower limit (`%low`). Files are premigrated if the storage pool occupancy is between the `%low` and the `%premig` limit. The value of `%premig` must be between 0 and the `%low` value. The amount of files being premigrated is equivalent to (`%low` - `%premig`) of the total pool capacity; however, the occupancy of the pool does not change because premigration copies the files to the external pool.

The premigration rule is triggered only if the file system pool occupancy exceeds the `%high` percentage. If this is the case and the pool occupancy percentage is between `%low` and `%premig`, files are premigrated. The amount of data that is premigrated corresponds to the occupation percentage minus `%premig`. Otherwise, if the storage pool occupancy exceeds `%low`, files are migrated until the `%low` value is met. Because the storage pool percentage does not change with premigration, the premigration processes a capacity that is equivalent to (`%occupied - %premig`).

For example, the following policy premigrates all files to IBM Spectrum Protect. The `%high` threshold is set to 0, so this rule is executed if the current occupation of the system pool is greater than 0%. The `%low` threshold is 100 and the `%premig` threshold is 0. This means as long as the system pool occupation is between 100 and 0, files are premigrated.

Restriction: Do not use this rule in an active policy because the rule will not free up capacity on the system pool.

```
/* define macros */
define( is_resident,(MISC_ATTRIBUTES NOT LIKE '%M%') )

/* external pool definition*/
RULE EXTERNAL POOL 'hsm'
EXEC '/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-
hsm.migrate' OPTS '-v'

RULE 'premig' MIGRATE FROM POOL 'system' THRESHOLD(0,100,0)
TO POOL 'hsm' where (is_resident)
```

This policy can now run periodically (based on a schedule) by using the following command:

```
# mmapplypolicy fsname -P policyfile -N hsmnodes [-B 1000 -m 3]
```

Using an adjusted interface script

Another way to run premigration is to create a separate interface script for this. The advantage of this approach is that no complex thresholds must be defined. To create an interface script for premigration, complete the following steps:

1. Copy the sample interface script to a `premigrate` interface script:
   ```
   # cp /usr/lpp/mmfs/samples/ilm/mmpolicyExec-hsm.sample
   /opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.premigrate
   ```

2. Adjust the `premigrate` interface script (`/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.premigrate`)
   Locate the token `$MigrateFormat` and add `-Premigrate`:
   ```
   $MigrateFormat = "%s %s -Premigrate -filelist=%s";
   ```

This special interface script for premigration can be used with policies to premigrate files. The following policy defines an external pool with the premigration interface script and premigrates all files that are resident:

```
/* define macros */
define( is_resident,(MISC_ATTRIBUTES NOT LIKE '%M%') )

/* external pool definition*/
RULE EXTERNAL POOL 'hsmPremig'
EXEC '/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.premigrate' OPTS '-v'

RULE 'premig' MIGRATE FROM POOL 'system' TO POOL 'hsm'
WHERE (is_resident)
```

This policy can now run periodically (based on a schedule) by using the following command:

```
# mmapplypolicy fsname -P policyfile -N hsmnodes [-B 1000 -m 3]
```

For more information about the optional parameters of the `mmapplypolicy` command, see Policy engine parameters. To run this policy periodically, see Automation and Scheduling.

The policy engine must be started on a node that has the IBM Spectrum Protect for Space Management client installed. For more information about migration policies with IBM Spectrum Scale, see Item 17 in References.

3.4 Recall

A recall operation copies migrated files from an external pool that is represented by IBM Spectrum Protect back to the internal file system pool. At the end of the recall process, the file content is typically in premigrated state. If the file is changed on the internal pool afterwards, the file state becomes resident. Files can also be recalled to resident state by using the `dsmrecall` command or policies.

When multiple migrated files must be recalled at the same time, certain recall operations can be tape optimized. Tape optimized recall provides much better throughput. With the tape optimized recall, the files to be recalled are sorted by the tape ID and the position on tape and are subsequently recalled in this order. Without this intelligent sorting, the recall of multiple files can require many tape mounts and the throughput is diminished.

The recall operation can be initiated transparently when a migrated file is opened in the file system (see Transparent recall). Alternatively, the recall operation can by started via command or policy engine (see Command line and Policy driven).

3.4.1 Transparent recall

The transparent recall of a migrated file is initiated when a user application opens the file for READ, WRITE, APPEND or TRUNCATE operations. If this happens, the DMAPI system that is part of IBM Spectrum Scale locks the user application and generates a DMAPI_READ, DMAPI_WRITE, or DMAPI_TRUNCATE event and queues this for HSM processing. The HSM recall service receives the event and initiates the appropriate data processing. Initially, it will read the file attribute IBMObj and extract the extObjId from the attribute that reflects the unique identifier of the file to locate it at the IBM Spectrum Protect server. Then the file data recall is initiated by connecting to the IBM Spectrum Protect server and requesting the data transfer to the file system. After the file data arrives at the file system, the HSM recall service responds to the DMAPI event. The DMAPI system unlocks the user application and allows access to the file data.

Restriction: If the event is a DMAPI_TRUNCATE event to a file size of 0 bytes, no recall of data occurs.

If the application accesses the file for a READ operation, the file is recalled to premigrated state. In all other cases, the file is recalled to resident state. Additionally, migrated and premigrated files can be brought to resident state by using the following command:
`#dsmrecall -resident premigrated_file`

If a premigrated file is opened for a WRITE, APPEND, or TRUNCATE operation, the file state is changed to resident without data transfer operations between the space management client and server.

Transparent recalls are handled from the HSM recall service and are distributed to other space management-enabled cluster nodes if any. The recall service can be described as a group of three or more `dsmrecalld` processes. Use the command similar to the following example to list the running recall daemon on a node:

```
#ps -ef --forest | grep dsmrecalld | grep -v grep

root      1453     1  0 Oct26 ?        00:00:00 dsmrecalld
root      1488  1453  0 Oct26 ?        00:01:27  \_
dsmrecalld
root      1489  1453  0 Oct26 ?        00:00:00  \_
dsmrecalld
```

The daemon with the parent pid 1 is the master recall daemon. On the space management node that owns the file system, the recall master receives the DMAPI data events. Find the owner node by using the following command:

```
#dsmmigfs query -detail file_system | grep Owner
```

The first daemon forked from the master is the recall distributor. The second fork is the recall receiver. The distributor distributes the recall jobs to recall receivers on the local node or talks to the recall receiver on other nodes to distribute the recall jobs. The distribution happens in a round robin fashion among all available cluster nodes. The recall receiver connects to the IBM Spectrum Protect server and executes the actual recall job.

Tip: The distribution mechanism can lead to a situation in which multiple files stored on the same tape cartridge are requested by different recall receivers. This concurrent access to tape cartridges can lead to multiple mounts and unmounts of the same cartridge. The recall distribution can be disabled by setting the following option in the `dsm.sys` file (see IBM Spectrum Protect for Space Management configuration):

```
HSMDISTRibutedrecall NO
```

Transparent recalls cannot be tape optimized. This means that, if multiple migrated files are opened at the same time, a tape mount might be required for each file because there is no sorting of the file by tape ID and position. This might slow down the recall operation and sometimes impact other file system operations. However, tape-optimized recall can be facilitated by using the `dsmrecall` command or policies (see Command line and Policy driven).

For transparent recalls, different recall modes can be configured, as explained in the next section (Recall modes).

Recall modes

Different recall modes can be configured on a per-file basis. The default recall mode (Normal recall) does not have to be configured because it always applies to migrated files.

Normal recall mode is the default for all files. Files are recalled completely from IBM Spectrum Protect server storage and can be accessed after the recall process is complete.

Restriction: The following options are honored only if a stub size > 0 is configured:

- The `readstartsrecall` option specifies whether a recall operation starts immediately when an application reads a stub file. To set this option, use the command:
  `# dsmmigfs update readstartsrecall` *`fspath`*
- In addition to the `readstartsrecall` option, the `previewsize` option takes effect when a recall is started. The `previewsize` option defines the leading bytes of a stub file that can be read without initiating a recall. Use the following command to set the option:
  `# dsmmigfs update previewsize` *`fspath`*

Streaming recall mode allows for an asynchronous recall of migrated files. The recalled portion of the file can be accessed while the rest of the file content is recalled. Streaming recall requires the file to be opened in read-only mode. To enable streaming recall for a migrated file use, the following command to change the file properties. The subsequent file open operation (read-only mode) initiates the streaming recall:

```
# dsmattr -recallmode=Streaming filename
```

For more information about the command, see Item 12 in References.

Certain space management options influence the behavior of streaming recalls, for example:

- The `minstreamfilesize` option is used to specify the minimum file size (in megabytes) for streaming recall mode. If the file size is small, a normal recall is performed. Small files should not be recalled in streaming mode because this adds more overhead.
- The `streamseq` option is used to specify the number of megabytes that are buffered before the recall daemon flushes the data to disk.

These options can be adjusted by using the following command:

```
# dsmmigfs update [-minstreamfilesize=100 | -streamseq=10]
fspath
```

Partial recall mode can be used to recall a portion of a migrated file. When an application makes a read request for a migrated file that is qualified for partial file recall, the space management client calculates which portion of the file to recall based on the offsets contained in the read request. This approach results in time and disk space savings because only a portion of the file is recalled. Partial recall requires the file to be opened in read-only mode. To enable partial recall for a migrated file, use the following command to change the file properties. The subsequent file open operation (read-only mode) will initiate the partial recall:

```
# dsmattr -recallmode=Partial filename
```

For more information about the command, see Item 12 in References.

Certain space management options influence the behavior of partial recalls, for example:

- The `minpartialrecallsize` option specifies the minimum size (in megabytes) for a file to qualify for partial file recall. If the file size is small, a normal recall is performed. Small files should not be partially recalled. This option can be updated by using the following command:
  `# dsmmigfs update [-MINPartialrecallsize=100] fspath`

Partial file recall mode takes precedence over streaming recall mode. If a file is smaller than the value of the `minpartialrecallsize` option, or if this option is set to 0, normal or streaming recall mode takes precedence.

3.4.2 Command line

You can run a recall operation by using the following command:
`# dsmrecall [options] [filespec or filelist=list] file_system_name`

For more information about the `dsmrecall` command, see Item 13 in References.

You can specify various options for the recall operation, for example:

- `Logname` Specifies the name of the log file
- `Resident` Specifies a recall to the resident state
- `Recursive` Specifies a recursive recall of all files relative to path name (*`filespec`*)
- `Offset` Specifies an offset from the beginning of the file for starting a partial recall
- `Size` Specifies the amount of data to be partially recalled relative to the offset

The files to be recalled are specified by a path name pattern (*`filespec`*) or a list of files containing one fully qualified path and file name per line.

Tape optimized recall: You can recall files in a *tape optimized manner*. Tape optimized recall operations sort the files by path name pattern (*`filespec`*) or file list (`filelist=`*`list`*) according to the tape ID and the location on tape. Files are recalled in this sequence. This method requires fewer tape mounts and provides better throughput when recalling many files. For example, to initiate a tape optimized recall of all files in directory `/ibm/gpfs0/archive` of file system `/ibm/gpfs0,` issue the following command. The trailing file system path is important:
`# dsmrecall -recursive /ibm/gpfs0/archive /ibm/gpfs`

Recall to resident state: Normally, a recall operation recalls the file to premigrated state, unless the `-resident` option is used to recall the file to resident state. Referring to the previous example, to recall all files to resident state, use the following command:

```
# dsmrecall -recursive -resident /ibm/gpfs0/archive
/ibm/gpfs
```

3.4.3 Policy driven

IBM Spectrum Scale MIGRATE policies can be used to recall files. The default interface script supports recall operations as well. You can run the following types of policy-driven recall operations: Standard recall to premigrated state, Recall to resident state, and Tape-optimized recall.

Standard recall

The following policy can be used to recall migrated files stored in directory `/ibm/gpfs0/archive`:

```
/* define macros */
define(is_migrated, (MISC_ATTRIBUTES LIKE '%V%'))
/* define directory to be recalled */
define(recall_dir, (PATH_NAME LIKE '/ibm/gpfs0/archive/%'))

/* exclude rule */
RULE 'exclude' EXCLUDE WHERE ( PATH_NAME LIKE
'%/.SpaceMan/%' OR PATH_NAME LIKE '.snapshots%' )

/* external pool definition */
RULE 'hsmpool' EXTERNAL POOL 'hsm' EXEC
'/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.migrate'
OPTS '-v'

/* recall rule */
RULE recall' MIGRATE FROM POOL 'hsm' TO POOL 'hsm'
WHERE (is_migrated)) AND (recall_dir)
```

This policy can be invoked by using the command:

```
# mmapplypolicy /ibm/gfps0 -P policyfile -N hsmnodes [-B
1000 -m 3]
```

For more information about the optional parameters of the `mmapplypolicy` command, see Policy engine parameters.

Recall to resident state

To recall files to resident state by using policies, you must edit the interface script. Complete the following steps:

1. Copy the sample interface script to a recall interface script:
   ```
   # cp /usr/lpp/mmfs/samples/ilm/mmpolicyExec-hsm.sample
   /opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.recall
   ```

2. Edit the recall interface script:
   (`/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.recall`)
   Locate the `$RecallFormat` token and add the `resident` option:
   `$ $RecallFormat = "%s %s -resident -filelist=%s";`

3. Use the new interface script in the previously described policy. The external pool definition looks like this:

```
/* external pool definition */
RULE 'hsmpool' EXTERNAL POOL 'hsm' EXEC
'/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.recall'
OPTS '-v'
```

Tape-optimized recall

When you use the policy engine for tape-optimized recall operations, you must change the interface script in order to accept the file system name that is associated with the external pool rule. Complete the following steps:

1. Copy the sample interface script to a `tor-recall` interface script:
   ```
   # cp /usr/lpp/mmfs/samples/ilm/mmpolicyExec-hsm.sample
   /opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.tor-recall
   ```

2. Edit the recall interface script:
   (`/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.tor-recall`)

   Locate the `$RecallFormat` token and add `%s` to the end. The `resident` option is optional if you want to recall to resident state:
   `$ $RecallFormat = "%s %s [-resident] -filelist=%s %s";`

   In the interface script where the processing options received from the policy engine are parsed, obtain the file system name (changes are marked in bold and italic):

```
# Process any options
# print "DEBUG1: ARGV = @ARGV\n";
# CHANGE: initialize fsName variable and get it from the
arguments
$fsName = " ";
foreach $opt (@ARGV) {
    if ($opt eq "-v") {
        $Verbose = 1;
        $VerboseOption = $VerboseCommandOption;
    }
```

```
#CHANGE: assume that the extra parameters given are -v and
fsName
    else {
      $fsName = $opt;
    }
}
```

In the script location where the recall command is composed, add the file system name:

```
  elsif ($command eq "RECALL") {
    # CHANGE: add $fsName at the end
        $syscmd = sprintf($RecallFormat,
                          $RecallCommand, $VerboseOption,
$hsmfilelist, $fsName);
  }
```

3. Use the new interface script in the external pool definition and add the file system name within the `OPTS` clause:

```
/* external pool definition */
RULE 'hsmpool' EXTERNAL POOL 'hsm' EXEC
'/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-hsm.tor-
recall'
OPTS '-v /ibm/gpfs0'
```

3.4.4 Recall storms

Recall storms occur when many migrated files are accessed simultaneously causing massive amount of recalls. This can cause a never-ending sequence of recalls which are not tape optimized. Recall storms in general impact the space management operations because they require many tape drives which are used inefficiently.

Let us look at the interaction between IBM Spectrum Scale and IBM Spectrum Protect for Space Management. Access to migrated files can be done from any node in the IBM Spectrum Scale cluster that has the file system mounted. Accessing a migrated file triggers a complex set of operations between IBM Spectrum Scale and IBM Spectrum Protect for Space Management:

- The file access request is received by the GPFS kernel extension and the kernel extension checks if the file is migrated.
- If this the file is migrated then the GPFS kernel extension registers a read event with the GPFS daemon, if the maximum of read events configured in parameter `dmapiWorkerThreads` is not exceeded. Otherwise the kernel extension blocks the read access until another registered read event completes. The maximum value of the parameter is `dmapiWorkerThreads` 64.

- The GPFS daemon issues a read event to the DMAPI application which is represented by the IBM Spectrum Protect for Space Management client recall daemons.
- The master recall daemon distributes each individual recall request if the maximum number of recalls (IBM Spectrum Protect for Space Management client option `MAXRECALLDAEMONS`) is not exceeded. If the maximum number of recalls is reached then the GPFS daemon requests are not executed until a recall daemon completed a recall job. The maximum value of the client option `MAXRECALLDAEMONS` is 99, whereby four of the recall daemons are used for internal purposes.
- If a recall request is completed then the GPFS daemon is notified to unblock the file access. The GPFS daemon informs the kernel extension and grants access to the file

Imagine there are many nodes in the cluster causing simultaneous access to migrated files. Each node can register several parallel read events whereby the number of read events is limited with the IBM Spectrum Scale configuration parameter `dmapiWorkerThreads`. The number of recalls in the cluster is limited by the IBM Spectrum Protect for Space Management client parameter `MAXRECALLDAEMONS`. Aligning the parameter `dmapiWorkerThreads` with the client option `MAXRECALLDAEMONS` is a challenging, because `dmapiWorkerThreads` is node specific and `MAXRECALLDAEMONS` is cluster-wide. Since each file to be recalled may be on a different tape or location on tape, it takes a long period of time to satisfy a massive number of transparent recalls. We talk about many hours or days.

If you experience massive transparent recalls there is a temporary cure:

- Enable the HSM client option: `hsmoptimizedrecallonly` (only available with IBM Spectrum Protect for Space Management client version 8.1.10 and above)
- Recycle the HSM daemon:

```
# dsmkilld
# sleep 3
# dsmrecalld
```

As a result, all transparent recall requests will fail, and no recall will be executed.

If you experience recall storms then think about preventing transparent recalls in general. The HSM client option `hsmoptimizedrecallonly` is a good starting point as shown in section Prevent recalls) . In addition, think about operational processes to allow users requesting files to be recalled and recall files requested by many users using the tape optimized recall (see section Tape-optimized recall). For more information see 23 in section References providing best practices for managing tiered storage file systems with tapes.

3.5 Reconciliation

To keep IBM Spectrum Scale file systems synchronized with the IBM Spectrum Protect server for space management services, the space management client automatically reconciles the file systems at preset intervals (see Automated reconciliation). You can also run reconciliation manually or control it by using a scheduler (see Manual reconciliation). And you can configure immediate reconciliation (for details, see Item 21 in References). For example, when you modify or delete a migrated or premigrated file from the local file system, an obsolete copy of the file remains in IBM Spectrum Protect server storage. During reconciliation any obsolete copies of migrated or premigrated files are marked for expiration.

You can specify how many days a migrated or premigrated file remains in IBM Spectrum Protect server storage after reconciliation by setting the `migfileexpiration` option in the `dsm.sys` options file. The default is 7 days. When the copies expire, they are removed from the server.

Reconciliation should be run only when absolutely necessary. Depending on the number of space-managed files in the file system, reconciliation can take a long time. For example, if the space-managed file system stores unchanged files that are kept for long periods of time, do not run reconciliation frequently. For this use case, you could run reconciliation manually once a month, once a quarter, or once a year.

3.5.1 Automatic reconciliation

Normally, the space management client automatically reconciles each file system for which space management is active. To specify how often reconciliation runs, set the `reconcileinterval` option in the `dsm.sys` options file (see IBM Spectrum Protect for Space Management configuration). The default is a reconciliation interval of 24 hours. To disable automatic reconciliation, set the `reconcileinterval` option to 0.

With IBM Spectrum Scale, however, the automatic reconciliation is controlled and started from the IBM Spectrum Protect for Space Management `dsmmonitord` daemon. If the HSM client was configured for IBM Spectrum Scale by exporting the environment variable `HSMINSTALLMODE=SCOUTFREE,` this daemon will not run. Hence, automatic reconciliation in an IBM Spectrum Scale environment typically does not work, unless the `dsmmonitord` daemon is started. Therefore, reconciliation should be initiated manually.

Automatic reconciliation can be run only for the entire space-managed file system.

3.5.2 Manual reconciliation

In large IBM Spectrum Scale environments, the preferred method is to use the `dsmreconcileGPFS.pl` script to initiate the reconciliation.

To manually reconcile a file system, use the following command:

```
# dsmreconcileGPFS.pl options fspath
```

The file system path (*fspath*) is the path name of the file system where stub files are recreated. To determine the file path, use the command:

```
# mmlsfs fsname  -T
```

This script invokes the IBM Spectrum Scale policy engine to generate a list of all migrated files. In parallel, the script queries the IBM Spectrum Protect server to get the list of migrated files from it. Then both lists are compared with the following results:

- Migrated files stored in the server, but not existing in the file system will be expired on the server.
- Migrated files with different attributes in the file system and the server will be updated on the server.
- Migrated files found in the file system, but not existing in the server will be reported as orphan stub files.

To run this command periodically, you can use a schedule, as described in Automation and Scheduling.

3.6 Recreating stub files

After file migration, a stub file is left in the file system. The stub file includes a pointer to the file content that is stored in the IBM Spectrum Protect server (see File attributes and Migration). If this stub file is lost (for example, if someone deleted it), the file content can no longer be accessed.

If a backup of the file exists (for example, as a result of the `mmbackup` command), the file can be restored by using the following command:

```
dsmc restore /path/filename
```

This command restores the entire file, including the access control list (ACL) and the Extended Accessibility (EA) attribute. If no file backup exists, you can use the `dsmmigundelete` command (for details, see Item 20 in References).

The `dsmmigundelete` command can be used to recreate the stub file of a file that is no longer in the file system. This command can also be used to recreate the stub file of a premigrated file that was lost. Consider the following:

- You can use the `dsmmigundelete` command to re-create stub files for the entire file system or for a list of files[1].
- The `dsmmigundelete` command does not recreate the ACL and EA metadata of the file. ACLs are applied according to the parent directory setting.
- The `dsmmigundelete` command creates a stub file that contains the necessary information to recall the corresponding file from storage. The stub file does not

[1] You can use a file list with the `dsmmigundelete` command beginning with the IBM Spectrum Protect for Space Management V8.1.4 client.

contain any leading bytes of data from the file, for example, if the stub size was set to greater than 0.

- The recall mode that you previously set for a migrated file is not stored in a re-created stub file. The recall mode for the file is set to normal.
- The space management client does not create a stub file if a directory path does not exist in the local file system for a migrated file. For this reason, you must create the directory path first.
- The `dsmmigundelete` command does not support hard linked files. To re-create a stub file for a hard linked file, all files that are hard linked together must be deleted from the local file system. When one file in a set of hard linked files is migrated, all of the hard linked files in the set become stub files. When the `dsmmigundelete` command re-creates a stub file for a hard linked file, the stub file has the same name as the file that was originally migrated. Stub files are not re-created for any other files that were previously in the hard linked set of files.

Stub files for deleted migrated and premigrated files can be recreated with the command:
`# dsmmigundelete -filelist=`*`list`* `[options]` *`fspath`*

You can specify options for the `dsmmigundelete` command, for example:

- `Server`: Specifies the name of the IBM Spectrum Protect server.
- `Logname`: Specifies the name of the log file.
- `Expiring`: Restores the stub file regardless of whether the file expired in the IBM Spectrum Protect server. Files in the server are expired upon reconciliation. Files that were expired and subsequently deleted from the IBM Spectrum Protect server cannot be recreated.
- `Recover`: Re-creates *all* existing stub files, overwriting previous versions. All premigrated and resident data that was previously stored in the file system is deleted. This option should only be used if the impact is understood. The use case for this option is a fast re-creation of all stub files in the file system without file existence verification.

The file system path (*`fspath`*) is the path name of the file system where stub files are recreated. To determine the file path, use the following command:
`# mmlsfs` *`fsname`* `-T`

Alternatively, specify a list of files with the command (option `--filelist=list`)[2]. The file list must have the following format:
`extObjID full_qualifying_path_of_the_file`

When you use the `filelist` option, only the files that are included in the file list are recreated. There are essentially two ways to create a list of files as an input file list for the command: By using an EXTERNAL LIST policy (see Creating a file list by using a LIST

[2] You can use a file list beginning with the IBM Spectrum Protect for Space Management V8.1.4 client.

policy) or by using a special space management client script (see Creating a file list by using the undelete script).

3.6.1 Creating a file list by using a LIST policy

To run the `dsmmigundelete` command with the `filelist` option, you must provide a list of files to be re-created. To create a file list that includes all migrated and premigrated files in the file system, you can use the following LIST policy:

```
RULE EXTERNAL LIST 'result' EXEC ''
RULE 'Undelete' LIST 'result'
   SHOW (XATTR('dmapi.IBMexID'))
   WHERE XATTR('dmapi.IBMexID') IS NOT NULL
```

To run this LIST policy, use the following `mmapplypolicy` command:

```
# mmapplypolicy fsname -P undelete.policy.rule -f ./undelete
-I defer
```

This command produces an output file that is named `undelete.list.result`. The name `undelete` comes from the `-f` option of the `mmapplypolicy` command; the name `list` denotes a list policy; and the name `result` comes from the EXTERNAL LIST name. Each file that is identified by the LIST policy is written on one line together with attributes. To filter out the attributes (extObjID and the path and file name) that are required by the `dsmmigundelete` command, you can use the following command:

```
# awk '{ $1=$2=$3=$5="";$0=$0;} NF=NF' <
scan.out.list.result > undelete.input.list
```

The file `undelete.input.list` is the input list for the `dsmmigundelete` command. To recreate the stub files for the files in a file list, run the following command:

```
# dsmmigundelete -filelist=undelete.inputlist -expiring
```

Use the `expiring` option if the file was expired through reconciliation.

Tip: To provision the case where individual stub files get lost and must be recreated, run the previous LIST policy periodically and store the output file. In case a stub (or premigrated) file is lost, the entry for this file can be extracted from the output file and fed into the `dsmmigundelete -filelist` command.

3.6.2 Creating a file list by using the undelete script

Alternatively, the IBM Spectrum Protect for Space Management V8.1.4 or later client provides a script that can be used to query the IBM Spectrum Protect server database to create a list of premigrated and migrated files. This script provides the most recent version of a migrated file available in the IBM Spectrum Protect server. To run this script, use the following command:

```
#
/opt/tivoli/tsm/client/hsm/multiserver/bin/dsmmigundelete.pl
```

```
--user=ADMIN_USER --passwd=ADMIN_PW --filespace=FILESPACE
--nodename=NODENAME --filename=SEARCHPATTERN
--filelist=/tmp/undelete.in
```

The parameters have the following semantics:

- The `user` and `passwd` parameters specify the credentials of an IBM Spectrum Protect server administrative user.
- The `filespace` parameter specifies the IBM Spectrum Protect server file space to be queried. This is typically the file system path.
- The `nodename` parameter specifies the name of the node that is used by the space management client for migration and recall operations. The node is configured in the `dsm.sys` file (NODENAME or ASNODENAME).
- The `filename` parameter specifies a search pattern. This can be a directory, a file name, or a partial file name, as shown in the following examples:
  - `/dir1` returns all pre- and migrated file information in the `dir1` directory and its subdirectories.
  - `/dir1/file` returns all pre- and migrated file information in `dir1` subdirectories that start with the word `file`.
  - `/dir1/fileA.txt` returns information from the specified file.
- The `filelist` option specifies the name of the output file with all files matching the query.

The output file (in this example: `=/tmp/undelete.in`) shows one line for each migrated or premigrated file that matches the query consisting of the following fields:
`NODE_NAME, FILESPACE_NAME, EXTOBJID, FILENAME, INSERT_DATE`

Use the output file as input for the `dsmmigundelete` command:

```
# dsmmigundelelte -filelist=/tmp/undelete.in -expiring
```

In the previous command, the `expiring` option is optional.
Yet another way to obtain the `extObjID` and the path and file name as input for the `dsmmigundelete -filelist` command is by issuing the `dsmmigquery` command. For more details, see Query migrated files on a server.

4 Integrating IBM Spectrum Scale functions

In an IBM Spectrum Protect for Space Management solution, you can integrate additional IBM Spectrum Scale functions, such as backup operations, active file management, and the Scale Out Backup and Restore (SOBAR) mechanism. We also provide guidance about automating migration operations within an IBM Spectrum Scale cluster.

4.1 Integration with backup operations

IBM Spectrum Scale includes a backup function that is run in conjunction with IBM Spectrum Protect. The backup function, based on the `mmbackup` command, is described in more detail in Item 14 in References. This backup function uses the IBM Spectrum Scale policy engine to identify files that changed since the last backup operation and feeds these files into the IBM Spectrum Protect backup-archive client, which backs up the files. This backup function is extremely scalable because multiple backup-archive clients running on different IBM Spectrum Scale nodes can execute the backup operation in parallel. In addition, different IBM Spectrum Protect servers can be used (see Item 19 in References).

Space management and backup can be performed on the same IBM Spectrum Scale file system. To benefit from the integration, the same IBM Spectrum Protect server must be used for backup and migration operations. IBM Spectrum Protect for Space Management integrates with the IBM Spectrum Scale backup function in several ways as explained below.

4.1.1 Client options for backup

The options that control the backup-archive client are stored in the same file as the option for the IBM Spectrum Protect for Space Management client:
`/opt/tivoli/tsm/client/ba/bin/dsm.opt`

If you use the IBM Spectrum Scale backup function in combination with the IBM Spectrum Protect backup-archive client, consider adding the following options to the client option file:

- `UPDATECTIME YES`: If the change time (`ctime`) of a file changes, the file is updated during the next backup cycle. There is no requirement to send the entire file, only the `ctime`.

- `QUOTESARELITERAL YES`: If a file name contains quotation marks, the quotation marks are interpreted as strings.

- `WILDCARDSARELITERAL YES`: If a file name contains one or more wildcards (like `*` or `?`), the wildcards are interpreted as strings.

- `SKIPACL YES`: File ACLs are not backed up, and if an ACL changes for a file, the change does not cause the file to be backed up. However, upon restore the file

ACL cannot be restored. Before you set this option, consider the consequences carefully.

- `SKIPACLUPDATECHECK YES`: The system does not monitor for changes in file ACLs. If the ACL for a file changes, the ACL is not backed up. This option should be used in conjunction with the `SKIPACL YES` option. Take care when using this option because files cannot be restored with the latest ACL.

Tips about using other client options:

- If you use the client option `AFMSKIPUNCACHEDFILES YES` in an AFM environment to prevent IBM Spectrum Protect for Space Management file migration of uncached or evicted files, the option has the same effect on backup operations. Uncached or evicted files are omitted from the backup operation.

- The values for the client options `SCROLLPROMPT, SCROLLLINES`, and `QUIET` are ignored during `mmbackup` operations. During the backup operations, the values for the three options are reset to default values.

4.1.2 Backup requirement prior to migration

The space management class in the IBM Spectrum Protect server can be configured to prevent files from being migrated if the files have not been backed up. For this purpose, the management class parameter `MIGREQUIRESBkup` can be set to yes. This setting ensures that files are migrated only after they have been backed up, to avoid recalls upon backup.

4.1.3 Inline copy or clone function

The inline copy function backs up a migrated file within the IBM Spectrum Protect server. In this case, the migration is completed before the backup operation. If a migrated file is being backed up, it is not recalled again, but is cloned within IBM Spectrum Protect from the space management pool to the backup pool. The inline copy function does not work under the following circumstances:

- If the backup operation is done to the different IBM Spectrum Protect server than space management.
- If the file to be backed up has ACL or extended attribute metadata. ACL is configured in the IBM Spectrum Scale file system with the `k` parameter (`mmlsfs fsname -k`).
- If the space management pool or the backup pool in the IBM Spectrum Protect server are container pools.

The inline copy function can be configured with the following command:

```
# dsmmigfs add [-INlinecopymode=MIG|PREMIG|OFF] fspath
```


The parameter -INlinecopymode must be considered only when running the mmbackup command for the same file system. The parameter specifies how a backup operation handles a migrated file. If this option is set to MIG and a migrated file is a candidate for a backup operation, the file is cloned inside the space management storage pool in the IBM Spectrum Protect server. The clone of the file is stored in the backup storage pool in the IBM Spectrum Protect server. This operation is called inline copy.

Restriction: IBM Spectrum Protect container storage pools (cloud and disk) do not support cloning. In this case, the inlinecopymode option must be set to NO.

4.1.4 Changes to Access Control Lists (ACL)

The backup operation (mmbackup) stores Access Control List (ACL) and Extended Attributes (EA) of the files along with the file data in the backup storage. This can indirectly impact the space management function: If a file is migrated and its ACL or EA are changed, then the file will not be recalled, however, the subsequent backup operation will select the file for backup. This is because the file's change time (ctime) has changed. The backup operation requires the file to be recalled.

The default behavior of the IBM Spectrum Scale backup function is to not recall migrated files that require backup (mmbackup option --skip-migrated). Instead the backup operation will end with a warning message indicating that migrated files must be backed up (see Backup skips migrated files). However, this mmbackup warning may become annoying. There are some options to deal with this.

To prevent backing up files upon ACL changes, the client options: SKIPACL or SKIPACLUPDATECHECK and UPDATECTIME can be set (see Client options for backup). While SKIPACL does not backup any ACL the option SKIPACLUPDATECHECK backs up the initial file ACL but no subsequent ACL changes. The option UPDATECTIME updates the CTIME in the IBM Spectrum Protect database which does not require a file recall. In addition, the IBM Spectrum Protect server management class parameter MIGREQUIRESBKUP can be set to prevent migration of files that have not been backed up.

Attention: Skipping ACL for backup comes for a price: If ACL are not backed up, then ACL cannot be restored. This may cause files to be inaccessible after restoring. Manual intervention is required by an authorized user to set the appropriate ACL after restore operation.

If you plan for backup and space management for a new file system and you anticipate ACL changes and do not care about ACL upon restoring then set the client options SKIPACL and UPDATECTIME and set the option MIGRREQUIRESBKUP in the management class of the IBM Spectrum Protect server. You can run mmbackup with option --backup-migrated to avoid the skipped files warning messages.

If you plan for backup and space management for an existing file system and you experience skipped files during backup upon ACL changes then set the client options SKIPACLUPDATECHECK and UPDATECTIME and set the option MIGRREQUIRESBKUP in the management class of the IBM Spectrum Protect server. You can run `mmbackup` with option `--backup-migrated` to avoid the skipped files warning messages. Be aware that the restore operation will restore the initial ACL of the files and perhaps not the latest ACL.

Note, if the IBM Spectrum Scale file system does not have ACL enabled (file system option -k posix), then ACL are not present. In this case the backup client stores the POSIX permission in the IBM Spectrum Protect server database and an update to file permissions does not cause a recall (or skip files respectively). Only if ACL are enabled in the file system (option -k nfs4) then the backup client will back up file ACL and the above applies.

For more information about the integration of IBM Spectrum Scale backup functionality with IBM Spectrum Protect for Space Management, see Item 19 in References.

4.1.5 Backup skips migrated files

The IBM Spectrum Scale backup function (`mmbackup`), when used with default options, does not recall migrated files if they are candidates for backup operations. Instead, the backup function creates a list of migrated files that are candidates for backup operations and informs the administrator about the files with a console message. The administrator can then recall these files and run the backup operation again.

This approach prevents massive recalls that would all be transparent. The backup function backs up all files that are candidates and not migrated and exits with a return code of 1 if any backup candidates are migrated.

The `mmbackup` option that control the recall behavior for migrated files are:

- `skip-migrated`: does not backup migrated files, even if these files must be backed up. Instead it creates a list of migrated files that require backup. This option is default. If migrated files are skipped then mmbackup end with a waring message.
- `backup-migrated`: backs up migrated files. This may cause recall storms if not used wisely.

In general, when configuring the IBM Spectrum Protect server management class with the option MIGREQUIRESBKUP yes then files are not migrated if the current version has not been backed up. An exception are ACL and EA changes, see section Changes to Access Control Lists (ACL) for guidance for managing ACL changes.

4.2 Integration with Active File Management (AFM)

IBM Spectrum Scale AFM is a scalable, high-performance file-system caching layer that is integrated with the IBM Spectrum Scale cluster file system. AFM is based on a home-cache model. A single home provides the primary file storage that is exported. One or

more caches provide a view into the exported home file system, through distinct file sets, without storing the file data locally. Upon file access in the cache, the data is fetched from home and stored in the cache. Another way to get files transferred from home to the cache is through prefetching. Prefetching can use the IBM Spectrum Scale policy engine to quickly identify files that match certain criteria.

IBM Spectrum Protect for Space Management and IBM Spectrum Protect backup and restore operations can be used in combination with AFM. If you use AFM, set the following IBM Spectrum Protect for Space Management option in the client option file:

```
AFMSKIPUNCACHEDFILES YES
```

For more information, see IBM Spectrum Protect for Space Management client configuration. By setting this option, you ensure that IBM Spectrum Protect for Space Management does not try to migrate files that are not cached or that are not yet replicated. For more detailed guidance, see Item 15 in References.

4.3 Integration with IBM Spectrum Scale snapshots

Integrating the space management function with IBM Spectrum Scale snapshots in the same file system or file set bear some risk. The usage of snapshots in a space managed file system can lead to recall storm under some circumstances.

If a migrated file is part of a snapshot and the file is deleted from the file system, then the file is transparently recalled into the snapshot. This is required to make the snapshot consistent. If many files are deleted at the same time then this may cause a recall storm, resulting in high recall activity and increased file system space usage. It makes no difference if the file was migrated before the snapshot was generated or afterward.

To prevent this situation, it is recommended to not create snapshots in a file system or file set that is space managed by IBM Spectrum Protect for Space Management.

4.4 Integration with Scale Out Backup and Restore (SOBAR)

IBM Spectrum Scale SOBAR is a method designed for fast disaster recovery of an entire file system that is space managed by IBM Spectrum Protect for Space Management. (For more information about SOBAR, see Item 18 in References.) SOBAR includes a backup and a restore methodology.

The basic idea of SOBAR is to premigrate or migrate, or to premigrate and migrate, all files in a file system to the IBM Spectrum Protect server by using the space management client and by periodically capturing and backing up the cluster and file system configuration as well as the file system metadata (image dump). The point in time when the last file system configuration and image dump were captured is the recovery point. If the entire file system must be recovered, the cluster and file system configuration is re-established and the file system metadata (inodes) is restored. After restoring the file system

metadata, the file system structure is rebuilt, and files can be gradually recalled from IBM Spectrum Protect. The recovery process for a very large file system is fast; however, at the end of the recovery process, the files are in migrated state.

The restore operation can be run for an entire file system, but not for individual files. If individual stub files that were lost in the file system must be restored, consider the procedure for Recreating stub files.

4.4.1 SOBAR backup

The SOBAR backup comprises the following high-level steps. For details, see Item 18 in References:

1. Run a migration or premigration operation, or both, for all files in the file system that must be protected. For more information about running migration and premigration operations, see Migration and Premigration.

   Optionally, run periodic checks for files that were not migrated or premigrated. For this purpose, you can use LIST policies (see Gathering space management statistics).

2. Back up the file system configuration by using the `mmbackupconfig` command:
   ```
   # mmbackupconfig fsname -o config.backup
   ```

   The file system configuration is backed up to the file denoted by the `-o config.backup` parameter, which specifies a fully qualified path and file name.

   The file system configuration can also be translated into a readable format by using the following command:
   ```
   # mmrestoreconfig fsname -i config.backup -F result.file
   ```

   The file that is specified by `result.file` includes a translated file system configuration. This includes the file system parameters as well as the NSD definitions (stanza).

3. Back up the file system metadata (such as inodes) from a snapshot:

   a. To have a consistency point, create a global snapshot (on the file system level) before taking the image backup of the file system:
   ```
   # mmcrsnapshot fsname snapshotname
   ```

   b. List the snapshot that was created:
   ```
   # mmlssnapshot fsname
   ```

   c. Create the image backup of the source file system. Ensure that the `sobar-filepath` provides sufficient capacity to store the image in dump format.

   Tip: When the image dump is created from a snapshot, it might be possible to store the

image in the file system that is being backed up. With the `notsm` option, the image dump is not automatically backed up to the IBM Spectrum Protect server.

```
# mmimgbackup fsname -S snapshotname -g sobar-filepath
-N acting_nodes
```

Delete the snapshot:

```
# mmdelsnapshot fsname snapshotname
```

At the end of this procedure, the following files should exist in the directory that was created for the image dump and configuration files (*sobar-filepath*) on the source. These files must be backed up, either to IBM Spectrum Protect or to another backup location:

- Config.backup: File containing the file system configuration data.
- Image dump: Typically, the `.idx` and `.sbr` file in a subdirectory structure similar to the following example:
  `sobar-filepath/imgbackup_xxxxx/mmPolicy.yyyyyyyy.zzzzzzzz`

  where `x`, `y`, and `z` are numbers or letters. Sometimes the directory containing the image dump is designated by a number (in IBM Spectrum Scale V4.1).

In addition to the file system configuration and the image dump, the IBM Spectrum Protect client configuration files (`dsm.opt` and `dsm.sys`, as described in IBM Spectrum Protect for Space Management client configuration) as well the cluster configuration should be backed up periodically.

This backup process can be automated in a script and run periodically by using schedules. For automation of storage services, see Automation and Scheduling and Item 16 in References.

4.4.2 SOBAR restore

In case of a disaster that makes the file system no longer usable, the SOBAR restore procedure can be used to re-create the file system from the stub files and the saved file system configuration. Prior to executing the SOBAR restore operation, ensure that the cluster is online and that the previous file system was removed from the configuration. In addition, restore the required files such as the file system configuration (the `backup.config` file) and ensure that the image dump was backed up. Also make available the IBM Spectrum Protect client configuration files (`dsm.opt` and `dsm.sys`).

SOBAR restore comprises the following high-level steps, which are describe in more detail in Item 18 in References:

1. Create a file system based on the configuration of the previous file system by using the `mmcrfs` command.

   The configuration parameters of the previous file system, including the NSD definition, can be obtained from the `backup.config` file that was created during the SOBAR

backup operation. To view the previous file system configuration, restore the `backup.config` file and translate it:

```
# mmrestoreconfig fsname -i config.backup -F result.file
```

The *`result.file`* includes the command to create the file system with the proper parameters (command: `mmcrfs`). Some file system parameters should not be changed, for example:

`-Q yes`. ***Remove this option*** from the command; the option will be set in a separate step later.

`-F nsd_stanza_file_of_IBM_Spectrum_Scale`. This is the NSD stanza file according to the NSD that existed in the previous file system. The NSD definition is included in the *`result.file`*.

`-k nfs4`. This parameter is required for Cluster Export Services (CES) on IBM Spectrum Scale.

`-B blocksize`. This parameter value should match the file system block size of the source system, but the value can also be larger (not smaller). To obtain the file system block size in the source system, use the command: `mmlsfs <fsname> -B`. If you change the block size, you must also adjust the size of the pools and associated disk.

`-i inodesize`. This parameter value should match the file system inode size of the previous file system, but the value can also be larger (not smaller) *if* there are no independent file sets in the source file system.

Tip: For IBM Spectrum Scale, consider using an inode size of 4K because this size aligns well with the disk I/O.

`-z yes`. This parameter must be set.

`--version version`. This parameter should match the file system version of the source system. To obtain the file system version in the source system, use the command: `mmlsfs fsname -V`

`--inode-limit total:preallocated`. Set the total and pre-allocated number of inodes equal to or higher than the values that are set in the source system. To obtain these numbers from the source system, use the `mmdf` command.

`-M maxMetadataReplication`. Adjust this parameter if you want to change the maximum metadata replication factor. The actual metadata replication factor (parameter `-m`) must be identical to the factor that was used in the previous file system. The factor can be changed on the new file system after the restore operation is finished and the new file system, created by using the `mmcrfs` command, is running.

`-R maxDataReplication`. Adjust this parameter if you want to change the maximum data replication factor. The actual data replication factor (parameter `-m`) must be identical to the factor that was used in the previous file system. The factor can be changed on the new file system after the restore operation is finished and the new file system, created with the `mmcrfs` command, is running.

> **Attention:** Do not change the data and metadata replication setting and the disk I/O latency (parameters `-m`, `-r`, and `-w`). These changes can cause errors during `mmrestoreconfig` and `mmimgrestore` operations.
>
> If the previous file system included independent file sets, the inode size, the metadata block size, or both must be identical on the new file system.
>
> In addition, if you change the block size on the new file system, the capacity of the file system pools also must be changed. For example, if the block size doubles, the capacity of the pools also must be doubled.

2. Verify the file system parameters on the new file system by using the command:
   ```
   # mmlsfs fsname
   ```

3. Verify that the size of the file system pools is identical to or larger than the previous file system:
   ```
   # mmdf fsname
   ```

4. Restore the file system configuration from the *backup.config* file. In this way, you can configure the essential file system parameters and link to file sets:
   ```
   # mmrestoreconfig fsname -i config.backup --image-restore
   ```

   This step restores the active policy that was configured on the previous file system.

5. Review and adjust the active policy. To review the policy, run the following command:
   ```
   # mmlspolicy fsname -L
   ```

6. Adjust the policy rules as required. For this purpose, copy the active policy obtained with the previous command to a file (*policyfile*) and edit this file. Be aware that the EXTERNAL POOL rule might have to be changed to reflect the correct name of the external pool script.

7. After you adjust the policy for this file system, activate and verify the policy by using the following commands:
   ```
   # mmchpolicy fsname policyfile -I test
   # mmchpolicy fsname policyfile -I yes
   # mmlspolicy fsname -L
   ```

8. Mount the new file system is mounted for the first time, in read-only mode:
   ```
   # mmmount fsname -o ro
   ```

9. Start the image restore operation. This operation requires the image dump files that were captured during SOBAR backup (command: `mmimgbackup`). Within the image dump, there is a subdirectory named imgbackup_*xxxxx* (*xxxxx* is a number) generated on the previous file system. Within this subdirectory, there is another subdirectory

named `mmPolicy*`. Within this directory is the image dump. To restore the image dump, use the following command:

```
# mmimgrestore fsname image-dump-dir/xxxxxxx/mmPolicy*
```

When the image dump restore is completed, inodes are created in the new file system.

10. To check for files, use one of the following commands (*fspath* is a path in the new file system):

    ```
    # ls -l fspath
    # find fspath | more
    ```

    Important: Do not access any files in the file system because HSM is not yet enabled.

11. Unmount the file system again:

    ```
    # mmumount fsname -a
    ```

12. Apply only the Quota settings as stored in the *config.backup file*:

    ```
    # mmrestoreconfig fsname -i sobar-filepath-
    target/config.backup -Q only
    ```

13. Mount the file system in read-write mode to prepare the HSM configuration:

    ```
    # mmmount fsname -a
    ```

14. Clean up the HSM configuration for the new file system and add this file system to HSM.

    Ensure that the IBM Spectrum Protect client is configured properly in regard to the client configuration files (`dsm.opt` and `dsm.sys`).

    Add the file system to the space management configuration. Consider additional options such as the HSM quota.

    ```
    # rm -rf fspath/.SpaceMan/
    # dsmmigfs add fspath [-Quota=megabytes]
    ```

    Verify that the file system has been added to HSM:

    ```
    # dsmmigfs query fspath -detail
    ```

At this point, the file system configuration and metadata are restored. Files can be selectively recalled by using transparent, command, or policy-driven recalls (see Recall).

4.5 Integration with Transparent Cloud Tiering

The IBM Spectrum Scale Transparent Cloud Tiering function (TCT) allows to migrate, premigrate and recall files from the IBM Spectrum Scale file system to on object storage. This function is not compatible with IBM Spectrum Protect for Space Management when used in the same file system. However, IBM Spectrum Protect for Space Management can

be configured on one file system and TCT can be configured on another file system within the same IBM Spectrum Scale cluster.

Attention: do not configure IBM Spectrum Protect for Space Management and TCT on the same file system.

4.6 Automation and Scheduling

You can automate IBM Spectrum Protect for Space Management operations. As shown in the previous examples, the policy engine that is represented by the `mmapplypolicy` command is used to start migration operations (see Migration). To automate policy runs, this command must be scheduled on at least one node that is configured for space management.

At first glance, it seems that you could use a system scheduler such as `cron` or `timerd`. However, it is not that trivial. For example, if you want to run a migration in a cluster, on which node do you schedule it? And what happens if this node happens to be down? Will migration run? And what happens if the file system is not available when the migration job is scheduled? How do you manage log files for processes that are running in the background?

All these questions and many more are addressed in the IBM Spectrum Scale storage services automation framework. See Item 16 in References.

5 Hints and tips

We provide further guidance with hints and tips.

5.1 Verifying HSM service availability

The IBM Spectrum Protect for Space Management CLI command `dsmmigfs` provides the option to list the current state of the HSM services. Use the command:

```
#dsmmigfs query -node=all
```

This command displays the status of the space management client processes and nodes. For example:

```
#dsmmigfs query -detail -node=all
…
GPFS Node Name:                 blackpearl
GPFS Node ID:                   1
GPFS Status:                    active
HSM Status:                     active
Recall Daemon Session ID:       59F1F00C
Mount Disposition:              YES
Ping Recall Daemon:             YES
Watch Daemon Session ID:        0
```

If the space management client is able to handle requests, the output shows:

```
HSM Status: active
```

If you add the `-detail` parameter, the command also collects information about the state of the IBM Spectrum Scale cluster. The `-detail` argument increases the runtime of the command. The `GPFS Status` information that is provided is identical to the information of the IBM Spectrum Scale `mmgetstate` command.

The `pars` option shows the output in a machine readable manner, for example:

```
#dsmmigfs query -detail -pars -node=all
dsmmigfs:queryNode:blackpearl:1:active:active:59F1F00C:YES:Y
ES:0:
```

To check the failover status of the HSM component on a given node, use the command:

```
# dsmmigfs q -f
```

In addition to this command, you can verify the space management client services by using the following command:

```
# ps -ef | grep dsm
```

Normally, three or more recall daemons run simultaneously. An unstable recall process could indicate a problem.

5.2 Stopping and starting HSM services

In some cases, you might have to restart the HSM services on HSM client nodes. For example, if you changed settings in the management class of the IBM Spectrum Protect server that is used by the HSM client, you might have to restart the space management services.

To stop HSM services, complete the following steps:

1. Ensure that no recall or migration processes are running. To check for pending or running recalls, use the following command:
   ```
   # dsmq
   ```

   If recalls are running, remove recall processes by using the following command. The recall-ID is a result of the previous command:
   ```
   # dsmrm recall-id
   ```

2. To check for migration processes, use the following command:
   ```
   # ps –ef | grep dsmmigrate
   ```

   If migration processes are running, let them finish.

3. If no recall or migration processes are running, stop the recall daemons by using the following command:
   ```
   # dsmkilld
   ```

4. Stop all HSM daemons by using the following command:
   ```
   # dsmmigfs stop file_system_path
   ```

5. Check that all daemons are stopped by using the following command:
   ```
   # ps –ef | grep dsm
   ```

6. If the watch daemon is still running, stop it by using the following command:
   ```
   # systemctl stop hsm
   ```

7. To start the HSM services on an HSM client node, run the following commands:
   ```
   # systemctl start hsm
   # dsmmigfs start file_system_path
   ```

8. Verify that the HSM services are running by using the following command:
   ```
   #dsmmigfs query -detail -node=all
   ```

5.3 Unmounting a space-managed file system

An operation to unmount a space-managed file system in an IBM Spectrum Scale cluster can fail with an error message indicating that the resource is busy. In this case, you might have to stop the space management services. To stop the space management services, complete the following steps on all space management nodes:

1. To ensure that no processes are accessing space-managed files, disable access to the file system. (If processes are accessing files while the space management services are stopped, errors can result.)

2. Ensure that no automatic or scheduled migration policies are running or are scheduled to run. To list the automatic policies, use the command:
   `# mmchpolicy fsname -L`
   If there are threshold-based migration rules, create a policy file without these rules and apply the policy by using the following command:
   `# mmchpolicy fsname -P policyfile`

3. On each space management node, complete the following steps:

   a. Deactivate space management for the file system:
      `# dsmmigfs deactivate fspath`

   b. Verify that all migration and recall processes are stopped:
      `# ps -ef | grep dsmmigrate`

      Tip: No processes should be listed.

      `# dsmq`

      Tip: No recalls should be listed.

   c. Verify that no migration or recall processes stopped accidentally, leaving files in an incorrect file state.

      Determine the node ID for each HSM client node:
      `# dsmmigs query -failover`

      Check whether files exist in the following directory, whereby the token *node_id* stands for each node ID determined in the previous step:
      `# ls /file_system_mount/.SpaceMan/logdir/translog/node_id/`

      Tip: The directory should be empty. If the directory is not empty, restart the recall daemon processes on the local GPFS node:
      `# dmkilld; sleep 10; dsmrecalld`

Tip: This command sequence will correct all transaction failures and remove the entries from the `translog` directory.

4. Stop the space management services:
   ```
   # systemctl stop hsm
   ```
   By running this command, you disable the failover and stop all HSM services and daemons.

5. Unmount the file system:
   ```
   # mmumount fsname
   ```

6. Before you mount the file system, start the HSM services on each space management node:
   ```
   # systemctl start hsm
   ```
   By running this command, you automatically enable the previous failover mode and start all required HSM daemons.

7. Mount the file system:
   ```
   # mmmount fsname [-a]
   ```

8. Reactivate space management for the file system on each space management node:
   ```
   # dsmmigfs reactivate fspath
   ```

9. Verify the availability of the HSM service. For details, see Verifying HSM service availability.

10. Apply the original automatic migration policy by using the command:
    ```
    # mmchpolicy fsname -P policyfile.original
    ```

5.4 Logs and commands

Learn about log files and commands that can be used for logging.

5.4.1 Space management client logging

The IBM Spectrum Protect for Space Management client implements two logging engines. One of the engines is for error logging. This engine is identical to that of other IBM Spectrum Protect client products. The error logging can be adjusted by adding the following options to the server stanza in the `dsm.sys` file:

```
ERRORLOGNAME        /path/dsmerror.log
```

The same procedure can be used to enable space management client logging. Add the following lines to the server stanza:

```
HSMLOGNAME          /path/hsm.log
HSMLOGEVENTFLAGS    FILE
```

Tip: If you use the IBM Spectrum Protect backup-archive client, a third logging engine is available. The content logged is similar to the space management client logging. Enable backup-archive client logging by adding the following lines to the server stanza:

```
AUDITLOGGing        FULL
AUDITLOGNAME        /path/dsmaudit.log
```

For more details, see Item 22 in References.

The space management client also writes to the system log (for example, to `/var/log/messages` or `journalctl`). For more information, see System log. In addition to the log file entries in the syslog file, the HSM client creates dump files of the HSM internal state and the DMAPI system state each time the HSM service is restarted manually or must be recycled automatically. These dump files are written to the `/tmp/hsm/` directory. The files help IBM Support to identify the root cause of HSM service issues.

5.4.2 System log

The space management client creates entries in the syslog file of the server where the client runs. The HSM entries in the syslog file can be used to identify HSM problems and potential required configuration changes. The following HSM events are logged to the syslog file:

- Start and stop of recall daemon processes.

```
Nov 20 08:48:38 nodeA dsmrecalld: HSM(pid:13756): start master
Nov 20 08:48:38 nodeA dsmrecalld: HSM(pid:13797): start distributor
Nov 20 08:48:38 nodeA dsmrecalld: HSM(pid:13798): start receiver
Nov 20 08:48:38 nodeA dsmrecalld: HSM(pid:24026): start PERMANENT recall worker (ID:3;MIN:3;MAX:5)
Nov 20 08:48:38 nodeA dsmrecalld: HSM(pid:24027): start TEMPORARY recall worker (ID:4;MIN:3;MAX:5)
Nov 20 08:48:38 nodeA dsmrecalld: HSM(pid:24027): stop TEMPORARY recall worker (ID:4)
Nov 20 08:48:38 nodeA dsmrecalld: HSM(pid:20964): stop master
```

- MASTER (core component for file system management, receives DMAPI events) – PPID 1.
- DISTRIBUTOR (distributes recall requests to local and remote recall worker) – PPID master.pid.
- RECEIVER (receives recall requests from remote nodes) – PPID master.pid.
- WORKER (performs the real data recall activity) – PPID distributor.pid.

  Hint: If many TEMPORARY workers will be started and stopped frequently, the value for the HSM option `minrecalldaemon` should be increased. If the number of TEMPORARY recall workers is often close to the limit (ID == MAX), the value for the HSM option `maxrecalldaemon` should be increased.
- Send and receive signals.
- Creation of dump files.
- System events that cause the restart of HSM daemons or the failover of file system space management to other cluster nodes.

```
Nov 20 08:40:41 nodeA dsmwatchd: HSM(pid:7823): Stop local recall service. Reason: GPFS down
Nov 20 08:48:32 nodeA dsmwatchd: HSM(pid:7823): Restart local recall service. Reasons: invalid process list
```

- GPFS down (the GPFS daemon was not able to handle requests or was down).
- Invalid process list (the process list of running HSM recall daemons is incorrect; for example, there is more than one master daemon).

Tip: The most common root cause for the restart of the recall service with the reason `invalid process list` is the RPCbind service that might not be active. Verify if this service is active on all space management nodes by using this command:

```
# systemctl status rpcbind
```

If the service status is not active, start and enable the `rpcbind` service by using the commands:

```
# systemctl status rpcbind
# systemctl enable rpcbind
```

If the service is not installed, install it.

5.5 Gathering space management statistics

For a space-managed file system, you might have to gather statistics about the number of files that are migrated, premigrated, or resident. For this purpose, LIST policies can be used.

The following LIST policy includes rules to identify files according to their states:

```
/* define exclude list */
define( exclude_list,(PATH_NAME LIKE '%/.SpaceMan/%' OR
PATH_NAME LIKE '%/.snapshots/%') )

/* Define is migrated */
define( is_migrated,(MISC_ATTRIBUTES LIKE '%V%') )
/* list rule to list all migrated files */
RULE EXTERNAL LIST 'mig' EXEC ''
RULE 'list_mig' LIST 'mig' WHERE (is_migrated) AND NOT
(exclude_list)

/* Define is premigrated */
define( is_premigrated,(MISC_ATTRIBUTES LIKE '%M%' AND
MISC_ATTRIBUTES NOT LIKE '%V%') )
/* list rule to list all premigrated files */
RULE EXTERNAL LIST 'pmig' EXEC ''
RULE 'list_pmig' LIST 'pmig' WHERE (is_premigrated)
AND NOT (exclude_list)

/* Define is resident */
define( is_resident,(MISC_ATTRIBUTES NOT LIKE '%M%') )
/* list rule to list all resident files */
RULE EXTERNAL LIST 'res' EXEC ''
RULE 'list_res' LIST 'res' WHERE (is_resident) AND NOT
(exclude_list)
```

This LIST policy can be executed with the following command:

```
# mmapplypolicy fsName -P listpolicy -f ./gpfs0 -I defer
```

The policy run generates three output files, which are based on `EXTERNAL LIST` names and the prefix set by the `f` parameter:

gpfs0.list.res	Includes all file names of resident files
gpfs0.list.mig	Includes all file names of migrated files
gpfs0.list.pmig	Includes all file names of premigrated files

Restriction: If any of the output files does not exist, no files are identified according to the LIST rule.

For more information about using LIST policies, see Item 17 in References.

5.6 Policy engine parameters

The IBM Spectrum Scale policy engine provides a range of parameters to control processing. The following table summarizes the most important parameters:

Parameter	Description	Consideration

`-m`	Number of threads dispatched for migration. Each thread processes a bucket.	Depends on the number of nodes, the bucket size, and the size of files.
`-B`	Bucket size: number of files per bucket (per file list).	1000 – 10,000, depending on the number and size of files.
`-n`	Number of directory scan threads.	1
`-N`	Node names to be involved in policy execution.	Nodes where the HSM client is installed and running.
`--single-instance`	Only one instance of `mmapplypolicy` can run for this file system.	Single instance should be used.
`-I`	Run mode of policy (yes, test, or defer).	Changing the run mode can be useful for testing.
`-s` *`directory`*	Directory for temporary files created by the policy engine.	Set this parameter for large file systems.
`-M`	Makes it possible to substitute strings in the policy file.	This parameter is for special use cases.

The `-m` parameter controls how many parallel migration processes are started on each participating node. The `-B` parameter defines the number of files per bucket that are processed by one migration process at a time. Both parameters must be set in accordance with the average file size and the number of files to be migrated (see Aligning configuration parameters).

A migration policy must be started on a cluster node that has the IBM Spectrum Protect for Space Management client installed. The `-N` parameter is used to specify other nodes that have the client installed and participate in policy-based migration. Migration is done in parallel on multiple nodes.

The `--single-instance` parameter specifies that only one `mmapplypolicy` command is allowed to run at a time. If one `mmapplypolicy` command is started with the `--single-instance` parameter and another one is started while the first one is running, the operation will abort. This parameter ensures that there are no overlapping policy runs that could severely impact system performance. This parameter is especially important when migrating files directly to an IBM Spectrum Protect server tape pool (which is not recommended).

The `-s` parameter specifies a directory for temporary files that are created by the policy engine. The default directory is `/tmp`. If there are many files in the file system to be processed by the policy engine, the directory might become full, and this situation causes the policy run to abort. Therefore, especially for large file systems, you must specify a directory in a file system that is large enough. You might want to specify a directory on a file system that uses Flash or SSD drives because the operations on these temporary file lists enjoy low latency and help to make policy runs faster.

It is possible to pass strings from the `mmapplypolicy` command to the policy rule before this is evaluated. You can use the `mmapplypolicy` command parameter `-M "STRING=VALUE"` for this. STRING is a string in a rule and VALUE is the substitute. By using this method, you can create generic policies that can be applied to any file system by using a wrapper script.

Imagine having a migration policy that should run for four different file sets at different times. Instead of creating four almost identical policy files, you can create one policy file and substitute the file set name in the `MIGRATE` policy. The following example shows such a rule; take note of the `FOR FILESET ('FSETNAME'` clause:

```
/* define macros */
define(access_age,(DAYS(CURRENT_TIMESTAMP) -
DAYS(ACCESS_TIME)))

/* define exclude rule*/
RULE 'exclude' EXCLUDE WHERE (PATH_NAME LIKE '%/.SpaceMan/%'
OR
PATH_NAME LIKE '%/.snapshots/%')

/* here comes the migration rules from system to hsm*/
RULE EXTERNAL POOL 'hsm'
EXEC '/opt/tivoli/tsm/client/hsm/bin/mmpolicyExec-
hsm.migrate'
OPTS '-v'

RULE 'autoMig' MIGRATE FROM POOL 'system' TO POOL 'hsm'
FOR FILESET ('FSETNAME') WHERE (KB_ALLOCATED > 0) AND
(access_age > 30)
```

Now the policy engine can be started for a given file set (for example, MYFILESET) by using the `-M FSETNAME=myfileset` parameter:

`# mmapplypolicy` *`fsname`* `-P` *`policyfile`* `-N` *`hsmnodes`* `[-B 1000 -m 3]-M "FSETNAME=myfileset"`

The policy engine substitutes the `FOR FILESET ('myfileset')` clause and applies the policy run to the given file set.

5.7 Aligning configuration parameters

When running policies for migration or recall operations, you must consider several components and configuration parameters. The following components are involved in these jobs:

- The IBM Spectrum Scale policy engine represented by the command: `mmapplypolicy`
- The IBM Spectrum Protect for Space Management client
- The IBM Spectrum Protect server

Figure 4 provides an overview of the components and parameters.

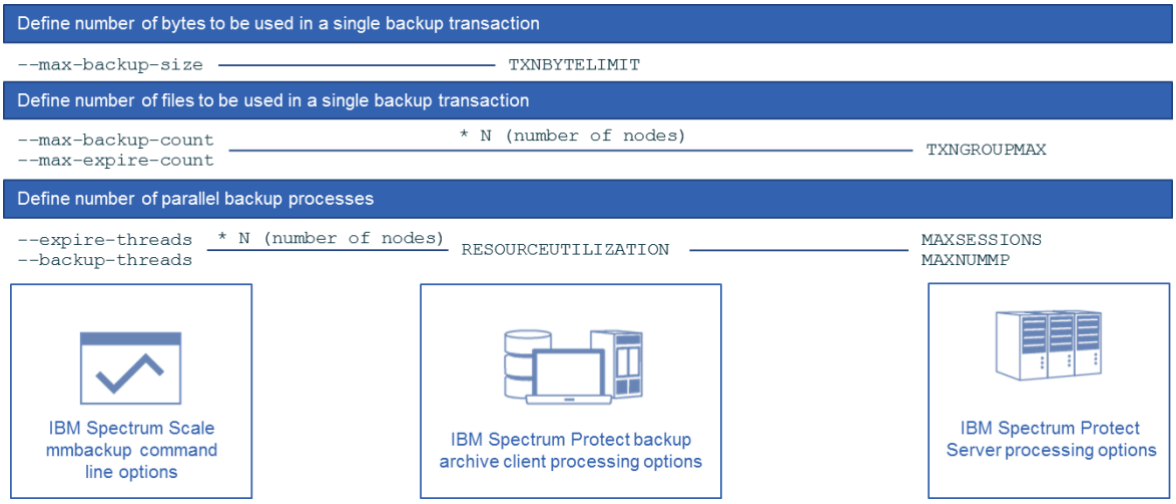

Figure 4: Components and respective parameters involved in migration and recall operations

As shown in Figure 4, the number of bytes transferred by a single client thread transaction is configured with the `TXNBYTELIMIT` client parameter. You can specify a value in the range 300 - 34359738368 (32 GB). The default is 25600 KB. For large files, this parameter should align to the file size or a multiple thereof. This parameter is configured in the client option file (`dsm.sys`).

The number of files transferred in a single transaction is configured with the server parameter `TXNGROUPMAX`. This parameter specifies a number in the range 4 - 65000 for the maximum number of files per transaction. The default is 4096. This parameter should be aligned with the `-B` parameter of the policy engine (see Policy engine parameters). This parameter is set in the server by using the following command:

```
setopt TXNGROUPMAX
```

This option is related to the `TXNBYTELIMIT` option in the client options file. After an object is transferred, the client commits the transaction if the number of bytes transferred during the transaction reaches or exceeds the value of `TXNBYTELIMIT`, regardless of the

number of objects transferred (option `TXNGROUPMAX`). So, these two options must be aligned according to the file size and the bucket size. The desired behavior is that one bucket of files is transferred within one transaction, if possible.

The number of parallel migration processes is defined with the `-m` parameter of the policy engine, which defines the number of parallel migration processes per participating node (see Policy engine parameters). In addition, the client option `RESOURCEUTILIZATION` must be considered, which regulates the level of resources the server and client can use during processing. These two options must align with the server parameters `MAXSESSIONS` and `MAXNUMMP`. The number of parallel processes per node (`-m` parameter) multiplied by the client option `RESOURCEUTILIZATION` represents the maximum numbers of sessions established between the IBM Spectrum Scale cluster and the IBM Spectrum Protect server during migration. Consequently, these server parameters must be set accordingly.

Do not configure file migration directly to an IBM Spectrum Protect server tape pool. Instead, configure a small disk buffer, large enough to absorb a daily migration volume. The tape pool can be configured as the next pool for this disk buffer.

5.8 Query migrated files on a server

In some cases, you might have to check the state of a file on the IBM Spectrum Protect server. The following command can be used from an IBM Spectrum Scale node that is enabled for space management to determine the file state on the server:

```
# dsmmigquery -serverinfo path
```

This command is not documented in the official product documentation and not officially supported. It can be used in support cases to get detailed information about a migrated file and how it is represented on the IBM Spectrum Protect server. The command output is similar to the following example[3]:

```
#dsmmigquery -serverinfo /gpfs/big
…
Alias      : /big
insert date: 2017-10-26 17:11:49
extobjid   :
0101020C000000009B09283459DA907502ED3D8AD9AFF174E6248E70
migr state : MIGRATED
times      : c 2017-10-26 16:21:24 m 2017-10-26 16:21:24 a
2017-10-26 16:21:28
mode OCT   :     100666 uid: 0 gid: 0 size:1073741836
inode      :      23051 ACL size: 0 checksum: 1745981440
```

[3] The format of the output is based on IBM Spectrum Protect for Space Management V8.1.2. The output might vary from version to version.

```
DMI handle : 099B3428593128B7-0000000000005A0B-
0000000000000000-0090116800000000
local:file path    : /gpfs/big
local:extobjid     :
0101020C000000009B092834DCAC9075002CBD1E4C3098F2C2C30E68
local:migr state   : MIGRATED
```

The first few lines describe the file state on the IBM Spectrum Protect server. Near the end of the output, the lines that start with `local` represent the state of the file from the IBM Spectrum Scale file system. Typically, the information from the file system is equal to the information that is returned from the server. In some situations, the information can be different. If the migrated file was deleted locally, the `local:migr_state` indicates `ENOENT`.

The query can include multiple results for a single file. Some of the results will pertain to the active copy that was migrated most recently. Other results can be for previous versions of the file that were expired, but not deleted from the server. The following example shows three versions of the file `/gpfs/file0:`

```
 # dsmmigquery -serverinfo /gpfs/file0 | grep -A 2 Alias

Alias      : /file0
insert date: 11/17/2017 09:56:03
extobjid   : 0101020C000000009B0970341911A075019537C02399B5CB219784DD
migr state : MIGRATED
--
Alias      : /file0
insert date: 11/03/2017 10:08:07
extobjid   : 0101020C000000009B0970344903A07500D4034B9F938159A40C2645
migr state : EXPIRED at 11/17/2017 09:05:40
--
Alias      : /file0
insert date: 11/08/2017 15:32:28
extobjid   : 0101020C000000009B097034A9A8A075031BF49D953647A0B303977F
migr state : EXPIRED at 11/17/2017 09:05:40
```

By using the `insert_date` and the `migr_state` tokens, the most current version of the file and its `extObjID` can be identified. In the previous example, the file `/gpfs/file0` was migrated three times. The first migration happened at 11/03/2017, and the second migration happened at 11/08/2017. Both versions where expired at 11/17/2017 09:05:40. A new version of the file was migrated at 11/17/2017 09:56:03.

The `dsmmigquery` command can also be used to create a file list for recreating stub files that where lost in a file system by using the command `dsmmigundelete -filelist` (see Recreating stub files). The file list for this command must include the `extObjID` and the path and file name. This information can be extracted from the output of the `dsmmigquery` command.

Restriction: To recreate a previous version of a file, the existing stub file in the live file system must be renamed and then the `dsmmigundelete -filelist` command must be used with a file list that contains the command and the file name of one of the previous versions of the file.

In addition to the query initiated from the space management client by using the `dsmmigquery` command, the IBM Spectrum Protect server database can be queried as well by using a SELECT statement. The file's metadata, inserted by the space management client into the server, is stored in the `SPACEMGFILES` table. To initiate a query, the file system path is required (`mmlsfs fsname -T`), along with the path and file name relative to the file system path. In the following example, a query is done for the file `/gpfs/path/file01`, whereby the file system path is `/gpfs` and the path and file name is `/path/file01`:

```
tsm: TSM01>select * from spacemgfiles where
filespace_name='/gpfs and file_name='/path/file01'
  NODE_NAME: HSMNODE
  FILESPACE_NAME: /gpfs
  STATE: ACTIVE_VERSION
  EXTOBJID:
0101020C00000000FEA902080228707401A942CB282BFD569E3F8EE7
  OBJECT_ID: 21744
  FILE_NAME: /path/file01
  INSERT_DATE: 2016-08-08 19:00:36.000000
  DELETE_DATE:
  CLASS_NAME: DEFAULT
```

If a migrated file is not on the server, this query does not return a result.

5.9 Prevent recalls

In environments with direct end-user interaction in the space management file system, you might want to prevent transparent recall operations because these could result in recall storms (see section Recall storms). For more information about best practices for space management file systems see item 23 in section References.

When disabling transparent recalls then operational procedures must be established to get files back from IBM Spectrum Protect. In the course of these procedures the end-user provides the file names to the administrator, the administrator recalls the files from may end-users leveraging the tape optimized recall function and informs the user that the requested files are accessible.

There are two ways to prevent recalls. One way to prevent transparent recalls is to use the client configuration parameter:
`hsmoptimizedrecallonly`

This option is available with IBM Spectrum Protect for Space Management client version 8.1.10 and above. With this option set, migration and read/write operations still work for premigrated files. Furthermore, truncation to the size of zero bytes can be performed on migrated files. In general, this option prevents any transparent recall that requires data transfer from the IBM Spectrum Protect server to the client system.

Alternatively, another option can be configured for environments where users access data in a space-managed file system via SMB. For the SMB share that is space managed, the following SMB option can be set (for more information, see Item 22 in References):

```
# mmsmb export change gpfs:recall=no
```

If the option `gpfs:recalls` is set to NO, files are not recalled on access and the SMB client receives an `ACCESS_DENIED` message. The default value of this option is YES, which initiates the recall upon access.

5.10 Partial file recall

In some cases, it may be required to recall only a portion of a file. One example can be that you restored the file system after a disaster using SOBAR and all files are in stub format now. Your business requirements may be that all files must have a resident portion (stub size) of the files. This can be achieved using the following procedure.

IBM Spectrum Protect for Space Management implements a function for partial recall that can be enabled on file level using the `dsmattr` command or globally for the entire file system using the `dsmmigfs update` command. The partial recall itself can be performed using the `dsmrecall` command. Due to limitations of the DMAPI standard a file can have 32 parts only. When you recall more than 32 portions of a file the recall mode will be set to normal automatically and the file will be recalled entirely.

When starting the procedure, the file is in recall mode normal and has 0 resident bytes:

Switch a file to partial recall mode:

```
# dsmls test02 | grep test02
   411506721            0            0   m      test02

# dsmattr -recall=partial test02 | grep test02
File attributes for /gpfs/partial_recall_test/test02 set
successfully.

# dsmls test02 | grep test02
   411506721            0            0   m (p)  test02
```

Now the file is in recall mode partial (indicated by the (p)) and the leading portion of the file can be recalled using the partial recall:

```
# dsmrecall -offset=0 -size=16M test02

# dsmls test02 | grep test02
   411506721     16777216        16384   m (p)  test02
```

Now the leading portiuon of the file are resident and the file can be switched to normal recall mode:

```
# dsmattr -recall=normal test02 | grep test02
File attributes for /gpfs/partial_recall_test/test02 set
successfully.

# dsmls test02 | grep test02
   411506721     16777216        16384   m      test02
```

By adjusting the setting for the stub size and the specific behavior of the IBM Spectrum Protect for Space Management client for stubs in this mode the wanted behavior can be set. Use the below options (See the product documentation for more details):

```
# dsmmigfs query -detail
…
Stub Size:              0
Read Starts Recall:     no
Preview Size:           0
```

APPENDIX

References

[1] IBM Spectrum Protect Blueprints:
https://www.ibm.com/support/pages/ibm-spectrum-protect-blueprints

[2] IBM Spectrum Protect for Space Management system requirements:
https://www.ibm.com/support/pages/node/86837

[3] Compatibility matrix for the IBM Spectrum Protect client and server:
https://www.ibm.com/support/pages/node/660949

[4] IBM Spectrum Protect client downloads:
https://www.ibm.com/support/pages/node/1108473

[5] IBM Spectrum Protect for Space Management product documentation:
http://www.ibm.com/support/knowledgecenter/SSERBH_8.1.9/hsmul/welcome.html

[6] IBM Spectrum Protect for Space Management installation prerequisites:
http://www.ibm.com/support/knowledgecenter/SSERBH_8.1.9/hsmul/r_gen_reqs.html

[7] Installing the IBM Spectrum Protect for Space Management client:
http://www.ibm.com/support/knowledgecenter/SSERBH_8.1.9/hsmul/c_inst_lnx_gpfs_ovw.html

[8] Configuring IBM Spectrum Protect for Space Management client:
http://www.ibm.com/support/knowledgecenter/SSERBH_8.1.9/hsmul/c_ovw_selopt.html

[9] Command `dsmmigfs add` for registering a file system for space management:
http://www.ibm.com/support/knowledgecenter/SSERBH_8.1.9/hsmul/r_cmd_dsmmigfs_add.html

[10] IBM Spectrum Scale `MISC_ATTRIBUTES` file attribute:
http://www.ibm.com/support/knowledgecenter/STXKQY_5.0.4/com.ibm.spectrum.scale.v5r04.doc/bl1adv_usngfileattrbts.htm

[11] The `dsmmigrate` command:
http://www.ibm.com/support/knowledgecenter/SSERBH_8.1.9/hsmul/r_cmd_dsmmigrate.html

[12] The `dsmattr` command:
http://www.ibm.com/support/knowledgecenter/SSERBH_8.1.9/hsmul/r_cmd_dsmattr.html

[13] The `dsmrecall` command:
http://www.ibm.com/support/knowledgecenter/SSERBH_8.1.9/hsmul/r_cmd_dsmrecall.html

[14] The `mmbackup` command:
http://www.ibm.com/support/knowledgecenter/STXKQY_5.0.4/com.ibm.spectrum.scale.v5r04.doc/bl1adm_mmbackup.htm

[15] Configuring IBM Spectrum Protect for Space Management with IBM Spectrum Scale AFM:
https://www.ibm.com/support/pages/node/5692358

[16] Automation of IBM Spectrum Scale storage services:
http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102676

[17] Configuring IBM Spectrum Scale ILM policies:
http://w3.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102642

[18] IBM Spectrum Scale Out Backup and Restore (SOBAR):
http://www.ibm.com/support/knowledgecenter/STXKQY_5.0.4/com.ibm.spectrum.scale.v5r04.doc/bl1adv_sobar.htm

[19] The `dsmmigundelete` command:
http://www.ibm.com/support/knowledgecenter/SSERBH_8.1.9/hsmul/c_mig_stub_recreat.html

[20] File system reconciliation with IBM Spectrum Protect for Space Management:
http://www.ibm.com/support/knowledgecenter/SSERBH_8.1.9/hsmul/c_recon_ovw.html

[21] Space management client logging options:
http://www.ibm.com/support/knowledgecenter/SSERBH_8.1.9/hsmul/r_setup_hsm_logging.html

[22] IBM Spectrum Scale `mmsmb` command:
http://www.ibm.com/support/knowledgecenter/en/STXKQY_5.0.4/com.ibm.spectrum.scale.v5r04.doc/bl1adm_mmsmb.htm

[23] Managing tiered storage file systems with tape
https://community.ibm.com/community/user/storage/blogs/nils-haustein1/2020/01/14/managing-files-in-tiered-storage

[24] HSM client option to prevent transparent recalls:
https://www.ibm.com/support/knowledgecenter/SSERBH_8.1.10/hsmul/r_opt_hsmoptimizedrecallonly.html

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBMproducts. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.