Dell EMC PowerScale: CloudPools and Microsoft Azure

Abstract
This white paper provides an overview of Dell EMC™ PowerScale™ CloudPools software in OneFS™ 9.0 and describes its policy-based capabilities that can reduce storage costs and optimize storage by automatically moving infrequently accessed data to Microsoft® Azure®.

June 2020
Revisions

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>April 2019</td>
<td>Initial release</td>
</tr>
<tr>
<td>October 2019</td>
<td>Updated snapshot efficiency</td>
</tr>
<tr>
<td>June 2020</td>
<td>Updated best practice</td>
</tr>
</tbody>
</table>

Acknowledgements

This paper was produced by the following:

Author: Jason He (Jason.He@dell.com)

Dell EMC and the authors of this document welcome your feedback on this white paper.

The information in this publication is provided “as is.” Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any software described in this publication requires an applicable software license.

Copyright © 2019 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, Dell EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners. [6/2/2020] [Technical White Paper] [H17746.2]
# Table of contents

Revisions ................................................................................................................. 2
Acknowledgements ..................................................................................................... 2
Table of contents ......................................................................................................... 3
Executive summary ...................................................................................................... 5
Audience ................................................................................................................... 5

1 CloudPools solution architectural overview .............................................................. 6
   1.1 PowerScale .......................................................................................................... 6
   1.1.1 SmartPools ....................................................................................................... 6
   1.1.2 SmartLink files ................................................................................................ 7
   1.1.3 File pool policies .............................................................................................. 7
   1.2 Microsoft Azure .................................................................................................. 9
   1.2.1 Cloud metadata object ..................................................................................... 9
   1.2.2 Cloud data object ........................................................................................... 9
   1.3 CloudPools operations ....................................................................................... 9
      1.3.1 Archive .......................................................................................................... 9
      1.3.2 Recall ............................................................................................................ 10
      1.3.3 Read ............................................................................................................. 11
      1.3.4 Update ......................................................................................................... 12

2 CloudPools 2.0 ........................................................................................................ 14
   2.1 NDMP and SyncIQ support ............................................................................... 14
   2.2 Non-disruptive upgrade support ........................................................................ 15
   2.3 Snapshot efficiency ......................................................................................... 15
      2.3.1 Scenario 1 .................................................................................................... 16
      2.3.2 Scenario 2 .................................................................................................... 16
      2.3.3 Scenario 3 .................................................................................................... 17
      2.3.4 Scenario 4 .................................................................................................... 17
      2.3.5 Scenario 5 .................................................................................................... 18
   2.4 Sparse files handling ......................................................................................... 19
   2.5 Quota management ......................................................................................... 19
   2.6 Anti-virus integration ...................................................................................... 20
   2.7 WORM integration ......................................................................................... 20

3 Best practices for PowerScale storage and Microsoft Azure ..................................... 21
   3.1 PowerScale configuration ................................................................................. 21
      3.1.1 CloudPools settings ..................................................................................... 21
Executive summary

This white paper describes about how Dell EMC™ PowerScale™ CloudPools in OneFS™ 9.0 integrates with Microsoft® Azure® and it covers the following topics:

- CloudPools solution architectural overview
- CloudPools 2.0 introduction with a focus on the following improvements:
  - Dell EMC PowerScale NDMP and Dell EMC PowerScale SyncIQ support
  - Non-disruptive upgrade (NDU) support
  - Snapshot efficiency
  - Sparse files handling
  - Quota management
  - Anti-virus integration
  - WORM integration
- General considerations and best practices for a CloudPools implementation
- CloudPools reporting, commands, and troubleshooting

Audience

This white paper is intended for experienced system administrators, storage administrators, and solution architects interested in learning how CloudPools works and understanding the CloudPools solution architecture, considerations, and best practices.

This guide assumes the reader has a working knowledge of the following:

- Network-attached storage (NAS) systems
- Dell EMC PowerScale scale-out storage architecture and Dell EMC PowerScale OneFS operating system
- Microsoft Azure

The reader should also be familiar with PowerScale and Azure documentation resources including the following:

- Dell EMC OneFS release notes, available on Dell EMC Support, containing important information about resolved and known issues
- Dell EMC PowerScale OneFS Best Practices
- Microsoft Azure
1 CloudPools solution architectural overview

The CloudPools feature of OneFS allows tiering cold or infrequently accessed data to lower-cost cloud storage. It is built on the Dell EMC PowerScale SmartPools file pool policy framework, which provides granular control of file placement on a PowerScale cluster.

CloudPools extends the PowerScale namespace to the public cloud, Microsoft Azure, as illustrated in Figure 1. It allows applications and users to seamlessly retain access to data through the same network path and protocols regardless of where the file data physically resides.

![CloudPools Solution Architectural Overview Diagram](image)

**Figure 1** CloudPools solution overview

**Note:** A SmartPools license and a CloudPools license are required on each node of the PowerScale cluster. A minimum of Dell EMC Isilon OneFS version 8.0.0 is required for CloudPools 1.0, and Dell EMC Isilon OneFS version 8.2.0 for CloudPools 2.0.

The tiering of data is driven by policies defined on the PowerScale cluster. The archived data can be accessed by clients through a variety of protocols including SMB, NFS, HDFS, and S3.

1.1 PowerScale

This section describes key CloudPools concepts including the following:

- SmartPools
- SmartLink files
- File pool policies

1.1.1 SmartPools

SmartPools is the OneFS data tiering framework of which CloudPools is an extension. SmartPools alone provide the ability to tier data between different node types within a PowerScale cluster. CloudPools also adds the ability to tier data outside of a PowerScale cluster.
1.1.2 SmartLink files

Although file data is moved to cloud storage, the files remain visible in OneFS. After file data has been archived to the cloud storage, the file is truncated to an 8 KB file. The 8 KB file is called a SmartLink file or stub file. Each SmartLink file contains a data cache and a map. The data cache is used to retain a portion of the file data locally, and the map points to all cloud objects.

Figure 2 shows the contents of a SmartLink file and the mapping to cloud objects.

![SmartLink file diagram](image)

Figure 2  SmartLink file

1.1.3 File pool policies

Both CloudPools and SmartPools use the file pool policy engine to define which data on a cluster should live on which tier or be archived to a cloud storage target. The SmartPools and CloudPools job has a customizable schedule that runs once a day by default. If files match the criteria specified in a file pool policy, the content of those files is moved to cloud storage during the job execution, and a SmartLink file is left behind on the PowerScale cluster that contains information about where to retrieve the data. In CloudPools 1.0, the SmartLink file is sometimes referred to as a stub, which is a unique construct that does not behave like a normal file. In CloudPools 2.0, the SmartLink file is an actual file that contains pointers to the CloudPool target where the data resides.

This section describes the key options when configuring a file pool policy, which include the following:

- Encryption
- Compression
- File matching criteria
- Local data cache
- Data retention

1.1.3.1 Encryption

CloudPools provides an option to encrypt data before it is sent to the cloud storage. It leverages the PowerScale key management module for data encryption and uses AES-256 as the encryption algorithm. The benefit of encryption is that only encrypted data is being sent over the network.

1.1.3.2 Compression

CloudPools provides an option to compress data before it is sent to the cloud storage. It implements block level compression using the zlib compression library. CloudPools does not compress data that is already compressed.
1.1.3.3 File matching criteria

When files match a file pool policy, CloudPools moves the file data to the cloud storage. File matching criteria enable defining a logical group of files as a file pool for CloudPools. It defines which data should be archived to cloud storage.

File matching criteria include the following:

- File name
- Path
- File type
- File attribute
- Modified
- Accessed
- Metadata changed
- Created
- Size

Any number of file matching criteria can be added to refine a file pool policy for CloudPools.

1.1.3.4 Local data cache

Caching is used to support local reading and writing of SmartLink files. It reduces bandwidth costs by eliminating repeated fetching of file data for repeated reads and writes to optimize performance.

**Note:** The data cache is used for temporarily caching file data from the cloud storage on PowerScale disk storage for files that have been moved off cluster by CloudPools.

The local data cache is always the authoritative source for data. CloudPools looks for data in the local data cache first. If the file being accessed is not in the local data cache, CloudPools fetches the data from the cloud. CloudPools writes the updated file data in the local cache first and periodically sends the updated file data to the cloud.

CloudPools provides the following configurable data cache settings:

- **Cache expiration**: Specifies the number of days until OneFS purges expired cache information in SmartLink files. The default value is one day.
- **Writeback frequency**: Specifies the interval at which OneFS writes the data stored in the cache of SmartLink files to the cloud. The default value is nine hours.
- **Cache read ahead**: Specifies the cache read ahead strategy for cloud objects (partial or full). The default value is partial.
- **Accessibility**: Specifies how data is cached in SmartLink files when a user or application accesses a SmartLink file on the PowerScale cluster. Values are cached (default) and no cache.

1.1.3.5 Data retention

Data retention is a concept used to determine how long to keep cloud objects on the cloud storage. There are three different retention periods:

- **Cloud data retention period**: Specifies the length of time cloud objects are retained after the files have been fully recalled or deleted. The default value is one week.
• **Incremental backup retention period for NDMP incremental backup and SyncIQ:** Specifies the length of time that CloudPools retains cloud objects referenced by a SmartLink file that has been replicated by SyncIQ or an incremental NDMP backup. The default value is five years.

• **Full backup retention period for NDMP only:** Specifies the length of time that OneFS retains cloud data referenced by a SmartLink file that has been backed up by a full NDMP backup. The default value is five years.

**Note:** If more than one period applies to a file, the longest period is applied.

1.2 **Microsoft Azure**

This section describes the following cloud objects in Microsoft Azure:

- Cloud metadata object
- Cloud data object

1.2.1 **Cloud metadata object**

A cloud metadata object (CMO) is a CloudPools object in Microsoft Azure that is used for supportability purposes.

1.2.2 **Cloud data object**

A cloud data object (CDO) is a CloudPools object that stores file data in Microsoft Azure. File data is split into 2MB chunks to optimize performance before sending it to Microsoft Azure. The chunk is called a CDO. If file data is less than the chunk size, the CDO size is equal to the size of the file data.

**Note:** The chunk size is 1 MB in CloudPools 1.0 and versions prior to OneFS 8.2.0.

1.3 **CloudPools operations**

This section describes the workflow of CloudPools operations:

- Archive
- Recall
- Read
- Update

1.3.1 **Archive**

The archive operation is the CloudPools process of moving file data from the local PowerScale cluster to cloud storage. Files are archived either using the SmartPools Job or from the command line. The CloudPools archive process can be paused or resumed. Refer to the section 5.1 for details.
Figure 3 shows the workflow of the CloudPools archive.

Additional workflow details include the following:

- The file pool policy in step 1 (see section 1.1.3) specifies a cloud target and cloud-specific parameters. Example policies include the following:
  - Encryption (section 1.1.3.1)
  - Compression (section 1.1.3.2)
  - Local data cache (section 1.1.3.4)
  - Data retention (section 1.1.3.5)

When chunks are sent from the PowerScale cluster to Azure in step 3, a checksum is applied for each chunk to ensure data integrity.

1.3.2 Recall

The recall operation is the CloudPools process of reversing the archive process. It replaces the SmartLink file by restoring the original file data on the PowerScale cluster and removing the cloud objects in Azure. The recall process can only be performed using the command line. The CloudPools recall process can be paused or resumed. Refer to the section 5.1 for detailed instructions on commands.

Figure 4 shows the workflow of CloudPools recall.
CloudPools solution architectural overview

Figure 4 Recall workflow

1.3.3 Read
The read operation is the CloudPools process of client data access, known as inline access. When a client opens a file for read, the blocks will be added to the cache in the associated SmartLink file by default. This can be disabled by setting the accessibility. For more detail, refer to the section local data cache.
Figure 5 shows the workflow of CloudPools read by default.

In step 1, OneFS looks for data in the local data cache first and moves to step 3 if the data is already in the local data cache.

### Update

The update operation is the CloudPools process that occurs when clients update data. When clients make changes to a SmartLink file, CloudPools first writes the changes in the data local cache and then periodically sends the updated file data to Azure. The space used by the cache is temporary and configurable. For more information, refer to the section [local data cache](#).
Figure 6 shows the workflow of the CloudPools update.

1. Client accesses the file through the SmartLink file.

2. OneFS retrieves CDOs from Azure, putting the file data in the local cache.

3. Client updates the file and those changes are stored in the local cache.

4. OneFS sends the updated file data from the local cache to Azure.

5. OneFS purges expired cache information for the SmartLink file.
CloudPools 2.0

CloudPools 2.0 is the next generation of CloudPools, released in OneFS 8.2.0. This chapter will describe the following improvements in CloudPools 2.0:

- NDMP and SyncIQ support
- Non-disruptive upgrade (NDU) support
- Snapshot efficiency
- Sparse files handling
- Quota management
- Anti-virus integration
- WORM integration

2.1 NDMP and SyncIQ support

When the CloudPools version differs between the source cluster and the target PowerScale cluster, the CloudPools cross-version compatibility is handled.

NDMP and SyncIQ provide two types of copy or backup: shallow copy and deep copy. For more information on NDMP and SyncIQ protection, refer to the white paper High Availability and Data Protection with Dell EMC PowerScale Scale-out NAS.

- **Shallow copy (SC)/backup:** Replicates or backs up SmartLink files to the target PowerScale cluster or tape as SmartLink files without file data.
- **Deep copy (DC)/backup:** Replicates or backs up SmartLink files to the target PowerScale cluster or tape as regular files or unarchived files.

Table 1 shows the CloudPools and OneFS mapping information. CloudPools 2.0 is released along with OneFS 8.2.0. CloudPools 1.0 is running in OneFS 8.0.x or 8.1.x.

<table>
<thead>
<tr>
<th>OneFS version</th>
<th>CloudPools version</th>
</tr>
</thead>
<tbody>
<tr>
<td>OneFS 8.0.x/OneFS 8.1.x</td>
<td>CloudPools 1.0</td>
</tr>
<tr>
<td>OneFS 8.2.0 or higher</td>
<td>CloudPools 2.0</td>
</tr>
</tbody>
</table>

Table 2 shows the NDMP and SyncIQ supported use cases when running a different version of CloudPools on the source and target clusters. As noted below, if CloudPools 2.0 is running on the source PowerScale cluster and CloudPools 1.0 is running on the target PowerScale cluster, shallow copies are not allowed.

<table>
<thead>
<tr>
<th>Source</th>
<th>Target</th>
<th>SC NDMP</th>
<th>DC NDMP</th>
<th>SC SyncIQ replication</th>
<th>DC SyncIQ replication</th>
</tr>
</thead>
<tbody>
<tr>
<td>CloudPools 1.0</td>
<td>CloudPools 2.0</td>
<td>Support</td>
<td>Support</td>
<td>Support</td>
<td>Support</td>
</tr>
<tr>
<td>CloudPools 2.0</td>
<td>CloudPools 1.0</td>
<td>No Support</td>
<td>Support</td>
<td>No Support</td>
<td>Support</td>
</tr>
</tbody>
</table>
2.2 Non-disruptive upgrade support
When a cluster that has been using CloudPools 1.0 is upgraded to OneFS 8.2.0 or higher, a new CHANGEOVER process is initiated automatically after the upgrade commit. The process ensures a smooth transition from CloudPools 1.0 to CloudPools 2.0. CloudPools 2.0 is ready to use once the upgrade state is committed. For more information on upgrade states, refer to the white paper PowerScale Non-Disruptive Upgrade (NDU) Best Practices.

2.3 Snapshot efficiency
Prior to OneFS 8.2.0, CloudPools 1.0 supported archiving files with existing snapshots. However, CloudPools 1.0 had a limitation when archiving files that have existing snapshots: the copy-on-writes (CoW) process copied the entire contents of the file into the snapshot. Archiving files with existing snapshots therefore did not save space on the PowerScale cluster until the previously CoW-created snapshots expired. CloudPools 1.0 offered an option (Uncheck Archive files with snapshots in WebUI) to skip such files with snapshots. A user might have not chosen to archive files with snapshots if the previously CoW-created snapshots had long retentions, to avoid creating another copy on cloud storage where the retention period meant it would persist on PowerScale storage anyway.

CloudPools 2.0 eliminates CoW on the primary data source PowerScale cluster when archiving files with snapshots to the cloud. The file data is only stored in the cloud storage, which saves space on the PowerScale cluster. For more information on data CoW for snapshots, refer to the white paper Data Protection with Dell EMC PowerScale SnapshotIQ.

However, CloudPools 2.0 does not operate on the target cluster in a SyncIQ relationship. Therefore, in an environment with long snapshot retentions and an expectation that the same snapshots are maintained in both clusters, it is possible for storage usage on a target cluster to grow much larger than the storage on the primary cluster which has CloudPools enabled. For space efficiency, a user with requirements for long snapshot retentions on two clusters in a SyncIQ relationship might choose to utilize natively-tiered PowerScale archive storage, rather than CloudPools at the present time.

SnapshotIQ can take read-only, point-in-time copies of any directory or subdirectory within OneFS. A file in one directory can be either a regular file or a SmartLink file before creating a snapshot. A regular file can be truncated to a SmartLink file after archiving its file data to the cloud. A SmartLink file can be converted to a regular file after recalling its file data to the PowerScale cluster. When a snapshot is taken, it preserves the exact state of a file system at that instant. A file in the snapshot directory (/ifs/.snapshot) is a SmartLink file if the same file in the source directory is a SmartLink file. A file in the snapshot directory is a regular file if the same file in the source directory is a regular file. The earlier version of data can be accessed later in the snapshot directory.

The following scenarios address CloudPools 2.0 and snapshots. HEAD is the current version of a SmartLink file in the source directory.

- The file is already a SmartLink file in the source directory before creating a snapshot.
  - Scenario 1: Update HEAD.
  - Scenario 2: Update HEAD multiple times and a new snapshot is created between multiple updates.
  - Scenario 3: Read file data from a snapshot.
- The file is still a regular file in the source directory before creating a snapshot. Then, the regular file is archived to the cloud after a snapshot creation.
  - Scenario 4: Update HEAD.
  - Scenario 5: Read file data from a snapshot.

### 2.3.1 Scenario 1

When updating HEAD (SmartLink files in snapshot), a new SmartLink is generated for HEAD when updating HEAD and write-back to the cloud. Cache for HEAD will be empty once its own cache expires. For the workflow of updating a SmartLink file, refer to the section 1.3.4. The original version SmartLink file is still used for the next snapshot of HEAD. This does not cause the snapshot space to grow. Figure 7 shows the process of scenario 1 to update HEAD when SmartLink files are in the snapshot directory.

![Figure 7](image.png)

**Figure 7** Scenario 1: Update HEAD when SmartLink files are in the snapshot directory

### 2.3.2 Scenario 2

This scenario describes updating HEAD multiple times and a new snapshot is created between multiple updates (SmartLink files in snapshot). For example, a user updates HEAD (the first update) while a new (most recent) snapshot is created before the first update write-back is made to the cloud. Subsequently, another user updates (the second update) HEAD again after the new (most recent) snapshot is created. Now there are two snapshots: one is the next snapshot of HEAD, the other is the most recent snapshot of HEAD. When a snapshot is taken, it preserves the exact state of a file system at that instant. Therefore, data for the next snapshot of HEAD is the old data that is already archived to the cloud and its cache is empty. Data for the most recent snapshot is the new data (old data with the first update) and its cache is dirty before the new data write-back is made to the cloud. Data for HEAD is the latest data (old data with the first update and the
second update) and its cache is dirty before the latest data write-back is made to the cloud. A new version SmartLink is generated for the most recent snapshot after the new data (old data with the first update) write-back is made to the cloud (write-back in the snapshot). Also, a new version SmartLink is generated for HEAD after the latest data (old data with the first update and the second update) write-back is made to the cloud (write-back in HEAD). Cache for the most recent snapshot or HEAD becomes empty once its own cache expires. Now, all file data is only stored on the cloud and saves space on the PowerScale cluster. Users can read file data from its own SmartLink file at any time.

Figure 8 shows the process of scenario 2.

2.3.3 Scenario 3
This scenario describes reading file data from a snapshot (SmartLink files in snapshot). The file in the next snapshot and HEAD use the same version of SmartLink file when not updating HEAD after the snapshot is created. This is no different than reading the same file from HEAD or the next snapshot of HEAD. For the workflow of reading a SmartLink file, refer to the section 1.3.3. The same local data cache is used when reading the same file from HEAD and the next snapshot of HEAD simultaneously. This does not cause the snapshot space to grow. The file in the snapshot directory uses its version of SmartLink file when updating HEAD and performing a write-back to the cloud like in scenario 1 or scenario 2. Users can read earlier versions of file data in the snapshot directory. The snapshot space could grow temporarily for cache data and the grown space is released once its own cache expires.

2.3.4 Scenario 4
In this scenario, when updating HEAD (regular files in snapshot), a SmartLink file is used for HEAD and a regular file is used for the same file in the next snapshot of HEAD. A new SmartLink file is generated for HEAD when updating HEAD and performing a write-back to the cloud. The cache for HEAD be empty once its own cache expires. Meanwhile, OneFS enables the Block Allocation Manager Cache Manager (BCM) on the regular file in the next snapshot of HEAD. BCM contains the metadata of mapping to cloud objects for the regular file in the next snapshot of HEAD. This does not cause the snapshot space to grow.
Figure 9 shows scenario 4.

![Figure 9](image)

**Scenario 4: Update HEAD when regular files are in the snapshot directory**

**2.3.5 Scenario 5**

In this scenario, when reading file data from a snapshot (regular files in snapshot), file data is the same for HEAD (SmartLink file) and the same file (regular file) in the next snapshot of HEAD when not updating HEAD after the snapshot creation. File data is read from HEAD when reading the same file in the next snapshot of HEAD, and this does not cause the snapshot space to grow. The file in the next snapshot of HEAD is a regular file (enabled BCM) and has the earlier version of data when updating HEAD and performing a write-back to the cloud like in scenario 4. The earlier version of data is retrieved from the cloud by BCM. File data is stored on the PowerScale cluster when reading the earlier version of data from the regular file in the next snapshot of HEAD. The snapshot space grows and the grown space is not released unless the snapshot is deleted.

**Note:** In OneFS 8.2.0, CloudPools 2.0 supports write-back in a snapshot. Refer to scenario 3 (section 2.3.3) for details. However, CloudPools 2.0 does not support archiving and recalling files in the snapshot directory. Consider the case when there is already file data in a snapshot on a cluster running a OneFS release prior to OneFS 8.2.0. That data takes up storage space on the PowerScale cluster, and then the cluster is upgraded to OneFS 8.2.0. Because CloudPools 2.0 does not support archiving files in snapshots to the cloud, the storage space for this snapshot cannot be released when the cluster is upgraded.

If the SmartLink files are backed up by SyncIQ or NDMP, the mapping file data should be retrieved from the cloud using the backup copy of the SmartLink file. If the backup retention has not expired, the CDOs of the mapping file data cannot be deleted even though the snapshot has been deleted, because the CDOs of the mapping file data are still referenced by the SmartLink file backup. When the backup retention period has expired and the CDOs of the mapping file data are no longer used, the CDOs of the mapping file data is deleted. For more information on data retention, refer to section 1.1.3.5 on data retention. If SmartLink files...
are not backed up by SyncIQ or NDMP, the CDOs of the mapping file data are deleted after the snapshot is deleted.

Users can revert a snapshot or access snapshot data through the snapshots directory (/ifs/.snapshot). The main methods for restoring data from a snapshot are as follows:

- Revert a snapshot through the SnapRevert job.
- Restore a file or directory using Microsoft® Shadow Copy Client on Windows® or cp command on Linux.
- Clone a file from a snapshot (CloudPools does not support cloning a file from a snapshot).

For details on restoring snapshot data, refer to the administration guide OneFS 8.2.0 Web Administration Guide. CloudPools does not support cloning a file from a snapshot. The other two methods for restoring data from a snapshot in a CloudPools environment are described as follows.

When using the SnapRevert job to restore data from a snapshot, it reverts a directory back to the state it was in when a snapshot was taken. For example, there is a /ifs/test directory including a regular.txt regular file, and a smartlink.txt SmartLink file that has its file data archived to the cloud. A snap01 snapshot is created on the /ifs/test directory, and updates are made on the two files. The regular.txt file is then archived to the cloud and it is truncated to a SmartLink file. Then, the SmartLink file smartlink.txt is recalled and it is converted to a regular file. If the snapshot snap01 is restored, it overwrites the files in directory /ifs/test. The regular.txt file reverts back to a regular file and the smartlink.txt reverts back to a SmartLink file. The directory /ifs/test is reverted to the state it was in when snap01 was taken.

When using Microsoft Shadow Copy Client on Windows or the cp command on Linux, the file data is retrieved from the cloud through SmartLink files in a snapshot to create new regular files. That means extra space is required for the new regular files restored from a snapshot.

### 2.4 Sparse files handling

CloudPools 2.0 provides a new sparse file format for to improve handling of empty blocks. With this improvement, sparse zeros are not included in CloudPools operations, which reduces network utilization and saves space on the cloud target.

**Note:** No cloud objects are written when archiving full sparse files (fully empty blocks).

### 2.5 Quota management

In OneFS 8.2.0, quotas present actual space consumed on the PowerScale cluster.

For example, if there is a directory or user quota of 500 GB and it is reporting 400 GB used, 200 GB of files are archived from the PowerScale cluster to cloud. Moving data to the cloud reduces the quota’s measured node space consumption. In OneFS releases prior to 8.2.0, the quota is freed by the amount of data that has been archived to the cloud, and the quota shows 200 GB (400 to 200 GB) used out of 500 GB. That means the user or directory quota can exceed the set limit (500 GB). In OneFS 8.2.0, the application logical size integrated with CloudPools 2.0 measures the true capacity consumption even if data is archived from the PowerScale cluster to the cloud, and the quota shows 400 GB used out of 500 GB through the application logical size. That means the user or directory quota cannot exceed the set limit of 500 GB.

For more information on the new SmartQuota reporting capabilities in OneFS 8.2.0, refer to the white paper Storage Quota Management and Provisioning with Dell EMC PowerScale SmartQuotas.
2.6 Anti-virus integration

In OneFS releases prior to OneFS 8.2.0, SmartLink files were skipped for anti-virus scanning.

In OneFS 8.2.0, CloudPools 2.0 provides a configurable option for anti-virus scanning of SmartLink files. The file data is retrieved from the cloud and cached on the cluster for the scan only if the option is enabled. As shown in Figure 10, the Scan Cloudpool Files option is configured and verified using the command line.

Note: The Scan Cloudpool Files option is disabled by default, which means SmartLink files are skipped when scanning a directory which includes SmartLink files.

2.7 WORM integration

Dell EMC PowerScale SmartLock is an optional software feature of OneFS that enables SEC 17-a4 data compliance. In enterprise mode, individual directories can be set up as Write Once, Read Many (WORM) directories, and the data is immutable by everyone except the root account on the cluster once the files have been committed. A PowerScale cluster can also be set up in compliance mode where the root account on the cluster is removed and no one can change or delete data in WORM-locked folders.

Prior to OneFS 8.2.0, SmartLink files are not allowed in both enterprise and compliance modes. In OneFS 8.2.0, details about CloudPools 2.0 and SmartLock integration are listed below:

- **Compliance mode**: SmartLink files are not allowed in compliance mode.
- **Enterprise mode**: SmartLink files are allowed in enterprise mode.
  - Enterprise mode can be enabled on a directory with SmartLink files.
  - SmartLink files can be moved into an Enterprise mode directory which prevents modifying or deleting the SmartLink files.
  - SmartLink files can be recalled from the cloud to the PowerScale cluster once they are committed.
Best practices for PowerScale storage and Microsoft Azure

This section focuses on the considerations and best practices for configuring PowerScale CloudPools and Microsoft Azure.

3.1 PowerScale configuration

This section includes considerations and best practices for configuring PowerScale CloudPools.

3.1.1 CloudPools settings

CloudPools settings can be changed either on the CloudPools setting tab or on a per file pool policy from the OneFS WebUI. It is highly recommended to change these settings on a per file pool policy. The following list includes general considerations and best practices for CloudPools settings.

- **Encryption:** Encryption is an option that can be enabled either on the PowerScale cluster or on Microsoft Azure. The recommendation is to enable encryption on the PowerScale cluster instead of on the Microsoft Azure. If the average CPU is high (greater than 70%) on the PowerScale cluster, the encryption can be enabled on Microsoft Azure instead of on the PowerScale cluster. It is important to note that encryption adds an additional load on the PowerScale cluster. Encryption can also impact the CloudPools archive and recall performance. For more information on Azure Encryption, refer to Microsoft Azure documentation.

- **Compression:** Compression is an option that can be enabled on the PowerScale cluster, in which file data is compressed before sending it to Microsoft Azure. If network bandwidth is a concern, the recommendation is to enable compression on the PowerScale cluster to save network resources. It is important to note that compression adds an additional load on the PowerScale cluster which means it might take more time to archive files from PowerScale storage to Microsoft Azure.

- **Data retention:** The recommendation is to explicitly set the data retention for the file data being archived from the PowerScale cluster to Microsoft Azure. If the SmartLink files are backed up with SyncIQ or NDMP, the data retention defines how long the cloud objects remain on Microsoft Azure. Once the retention period has passed, the PowerScale cluster sends a delete command to Microsoft Azure. Microsoft Azure marks the associated cloud objects for deletion. The delete process is asynchronous and the space is not reclaimed until garbage collection completes. This is a low-priority background process, which may take days to fully reclaim the space depending on how busy the system is.

- **Local data cache:** If the storage space is limited on the PowerScale cluster, the recommendation is to set lower values for the Writeback Frequency and Cache Expiration. This reduces the time to keep file data in the local data cache and frees up storage space sooner on the PowerScale cluster.
3.1.2 File pool policy

File pool policies define what data will be archived from the PowerScale cluster to Microsoft Azure. The considerations are listed below:

- Ensure the priority of file pool policies is set appropriately. Multiple file pool policies can be created for the same cloud storage account. When the SmartPools job runs, it processes file pool policies in priority order.
- In terms of freeing up storage space on the PowerScale cluster, the recommendation is not to archive small files that are less than 32 KB in size.
- If the files need to be updated frequently, the recommendation is not to archive those files.
- OneFS supports a maximum of 128 file pool policies (SmartPools and CloudPools combined). The recommendation is not to exceed 30 file pool policies per PowerScale cluster.
- If the file pool policy is updated, it has no impact on the files already archived. It will only affect the files to be archived when the SmartPools job next runs.

3.1.3 Other considerations

Additional considerations include the following:

- **Deduplication**: CloudPools can archive deduped files from a PowerScale cluster to cloud storage. However, un-deduped files will be created when recalling those files from the cloud to the PowerScale cluster. For more information on deduplication within OneFS, refer to the white paper [Next Generation Storage Efficiency with Dell EMC PowerScale SmartDedupe](#).
- **Small file storage efficiency (SFSE)**: CloudPools and SFSE cannot work together. For PowerScale clusters using CloudPools, any SmartLink files cannot be containerized or packed. It is best practice to not archive small files that will be optimized using SFSE. The efficiencies gained from implementing SFSE for small files, outweigh the storage advantages gained from archiving them to the cloud using CloudPools. For more information on the Small File Storage Efficiency feature of OneFS, refer to the white paper [Dell EMC PowerScale OneFS Storage Efficiency](#).
- **Network proxy**: When a PowerScale cluster cannot connect to the CloudPool storage target directly, network proxy servers can be configured for an alternate path to connect to the cloud storage.
- **Cloud storage account**: Do not delete a cloud storage account that is in use by archived files. Any attempt to open a SmartLink file associated with a deleted account will fail. In addition, NDMP backup and restore and SyncIQ failover and failback will fail when a cloud storage account has been deleted.
- **OneFS upgrade (CloudPools 1.0 to CloudPools 2.0)**: Before beginning the upgrade, it’s recommended to check the OneFS CloudPools upgrade path showing in Table 3.

<table>
<thead>
<tr>
<th>Installed OneFS Version (CloudPools 1.0)</th>
<th>Upgrade to OneFS Version (CloudPools 2.0)</th>
<th>8.2.0</th>
<th>8.2.1 with May 2020 RUPs</th>
<th>8.2.2 with May 2020 RUPs</th>
<th>9.x</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.0.x or 8.1.x</td>
<td>Strongly discouraged</td>
<td>OK if needed but recommend 8.2.2</td>
<td>Strongly Recommended</td>
<td>Strongly Recommended</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Contact your Dell EMC representative if you plan to upgrade OneFS to 8.2.0. For the May 2020 RUPs, refer to the document [Current OneFS Patches](#).
In a SyncIQ environment with unidirectional replication, the SyncIQ target cluster should be upgraded before the source cluster because this allows the CloudPools-1.0-formatted SmartLink files to be converted into CloudPools-2.0-formatted SmartLink files through a post-upgrade SmartLink conversion process. Otherwise, SyncIQ policy need to be reconfigured to deep copy but this will cause archived file content to read from the cloud and replicated. In a SyncIQ environment with bi-directional replication, it is recommended to disable SyncIQ on both source and target clusters, upgrade both source and target clusters at the same time, and then re-enable SyncIQ on both source and target clusters once the OneFS upgrades have been committed on both source and target clusters. Depending on the number of SmartLink files on the target DR cluster and the processing power of that cluster, the SmartLink conversion process can take considerable time.

**Note:** No need to stop SyncIQ / Snapshot during the upgrade. All converted stub files must be re-sync’d by SyncIQ, it may take SyncIQ some time to catch up with all the changes.

To check the status of the SmartLink upgrade process, execute the command below, substituting the appropriate job number.

```
# isi cloud job view 6
```

ID: 6  
Description: Update SmartLink file formats  
Effective State: running  
Type: smartlink-upgrade  
Operation State: running  
Job State: running  
Create Time: 2019-08-23T14:20:26  
State Change Time: 2019-09-17T09:56:08  
Completion Time: -  
Job Engine Job: -  
Job Engine State: -  
Total Files: 21907433  
Total Canceled: 0  
Total Failed: 61  
Total Pending: 318672  
Total Staged: 0  
Total Processing: 48  
Total Succeeded: 21588652

**Note:** CloudPools recall jobs will not run while SmartLink upgrade / conversion is in progress.

For NANNON (Not All Nodes On Network) cluster, it's recommended to get the unconnected nodes connected to the network before executing the SmartLink conversion and disable SnapDelete until the SmartLink conversion is completed.

### 3.2 Microsoft Azure configuration

Before configuring PowerScale CloudPools on the PowerScale cluster, Microsoft Azure need to be configured properly. The following are the general considerations and best practices when configuring Microsoft Azure for CloudPools.
Best practices for PowerScale storage and Microsoft Azure

- **URI for CloudPools**: The blob endpoint is used as the URI for CloudPools. For example, if your general-purpose storage account is named mystorageaccount, the default blob endpoint for CloudPools is https://mystorageaccount.blob.core.windows.net. The endpoint is the URI for CloudPools.

- **Azure storage account**: Azure storage supports several types of storage accounts. For more details, refer to the document [Azure storage account overview](#) on the Microsoft website. The general-purpose V2 is recommended to use for CloudPools.

3.3 Protecting SmartLink files

SmartLink files are the sole means to access file data stored in Microsoft Azure, so it is important to protect them from accidental deletion.

This section discusses using PowerScale SyncIQ and NDMP to back up SmartLink files.

**Note**: SmartLink files cannot be backed up using a copy command, such as secure copy (scp).

3.3.1 SyncIQ

SyncIQ is CloudPools-aware, but consider the snapshot-efficiency guidance in section 2.4, especially where snapshot retention periods on the target cluster will be long.

SyncIQ policies support two types of data replication for CloudPools:

- **Shallow copy**: Replicates files as SmartLink files without file data from source PowerScale cluster to target PowerScale cluster.

- **Deep copy**: Replicates files as regular files or unarchived files from source PowerScale cluster to target PowerScale cluster.

For information on cross-version compatibility of CloudPools, refer to section 2.1 on NDMP and SyncIQ support.

SyncIQ, SmartPools, and CloudPools licenses are required on both the source and target PowerScale cluster. It is highly recommended to setup a scheduled SyncIQ backup of the SmartLink files. For more information on PowerScale SyncIQ, refer to the white paper [Dell EMC PowerScale SyncIQ: Architecture, Configuration, and Considerations](#).

When SyncIQ replicates SmartLink files, it also replicates the local cache state and unsynchronized cache data from the source PowerScale cluster to the target PowerScale cluster. Figure 11 shows the SyncIQ replication when replicating directories including SmartLink files and unarchived normal files. Both unidirectional and bi-directional replication are supported. Appendix A provides steps for failing over to a secondary PowerScale cluster and failing back to a primary PowerScale cluster.
Best practices for PowerScale storage and Microsoft Azure

Site 1

<table>
<thead>
<tr>
<th>PowerScale</th>
<th>CloudPool1</th>
<th>R/W</th>
<th>CloudPool2</th>
<th>R only</th>
</tr>
</thead>
</table>

Site 2

<table>
<thead>
<tr>
<th>PowerScale</th>
<th>CloudPool1</th>
<th>R only</th>
<th>CloudPool2</th>
<th>R/W</th>
</tr>
</thead>
</table>

SyncIQ Policy:
- Replicate all files
- Unarchived files
- SmartLink files

Users or applications access the CloudPools data via SMB or NFS

SyncIQ replication

**Note:** If encryption is enabled in a file pool policy for CloudPools, SyncIQ also replicates all the relevant encryption keys to the secondary PowerScale cluster along with the SmartLink files.

3.3.2 NDMP

NDMP is also CloudPools-aware and supports three backup and restore methods for CloudPools:

- **DeepCopy**: Backs up files as regular files or unarchived files; files can only be restored as regular files.
- **ShallowCopy**: Backs up files as SmartLink files without file data; files can only be restored as SmartLink files.
- **ComboCopy**: Backs up files as SmartLink files with file data; files can be restored as regular files or SmartLink files.

For information on cross-version compatibility of CloudPools, refer to section 2.1 on NDMP and SyncIQ support.

It is possible to update the file data and send the updated data to the cloud storage. Multiple versions of SmartLink files can be backed up to tapes using NDMP, and multiple versions of CDOs are protected on Microsoft Azure under the data retention setting. You can restore a specific version of a SmartLink file from tapes to a PowerScale cluster and continue to access (read/update) the file like before.

**Note:** If encryption is enabled in the file pool policy for CloudPools, NDMP also backs up all the relevant encryption keys to tapes along with the SmartLink files.
4 Reporting

This section describes reporting for CloudPools network stats and includes the following topics:

- CloudPools network stats
- Query network stats by CloudPools account
- Query network stats by file pool policy
- Query history network stats

4.1 CloudPools network stats

CloudPools network stats collect every network transaction and provide network activity statistics from connections to the cloud storage. The network activity statistics include bytes in, bytes out, and the number of GET/PUT/DELETE operations. CloudPools network stats are available in two categories:

- Per CloudPools account
- Per file pool policy

**Note:** CloudPools network stats do not provide file statistics, such as the file list being archived or recalled.

4.2 Query network stats by CloudPools account

Use the following command to check the CloudPools network stats by CloudPools account.

```
isi_test_cpool_stats -Q --accounts <account_name>
```

Figure 12 shows an example of current CloudPools network stats by CloudPools account.

![Network stats by CloudPools account]

4.3 Query network stats by file pool policy

Use the following command to check the CloudPools network stats by file pool policy.

```
isi_test_cpool_stats -Q --policies <policy_name>
```

Figure 13 shows an example of current CloudPools network stats by file pool policy.

![Network stats by file pool policy]

**Note:** The number of delete operations by policy is not displayed by the file pool policy.
4.4 **Query history network stats**

Use the following command to check the history CloudPools network stats.

```
isi_test_cpool_stats -q -s <number of seconds in the past to start stat query>
```

Use the `s` parameter to define the number of seconds in the past. For example, set it as 86,400 to query CloudPools network stats over the last day.

Figure 14 shows an example of CloudPools network stats over the last day.

```
hop-is1-p-1# isi_test_cpool_stats -q -s 86400
Account   Bytes-in  bytes-out  gets  puts  deletes
  testaccount | 419436000   | 419436000   |  4000 |  2001 |   8001
```

Figure 14  Network stats last day

Use the following command to flush stats from memory to database and get the latest CloudPools history network stats.

```
isi_test_cpool_stats -f
```
5 Commands and troubleshooting

This section describes CloudPools commands and troubleshooting methodologies.

5.1 Commands

This CloudPools operations and job monitoring commands discussed in this section include the following:

- CloudPools archive
- CloudPools recall
- CloudPools monitoring

5.1.1 CloudPools archive

Run the following command to archive files from a PowerScale cluster to the cloud on demand.

`isi cloud archive <file name> --recursive [true | false] --policy <policy name>`

Parameters:

- `<file name>`: File name to be archived
- `--recursive`: Whether the archive should apply recursively to nested directories
- `--policy`: Policy name to be used with archiving

Run either of the following two commands to check whether the file is a SmartLink file or not, as shown in Figure 15.

`ls -loh <file name>`

`isi get -DD <file name> | grep -i smartlink`

![Figure 15 SmartLink file](image)

5.1.2 CloudPools recall

Run the following command to recall files from the cloud to a PowerScale cluster on demand.

`isi cloud recall <files> --recursive [true | false]`

Parameters:

- `<file name>`: File name to be archived
- `--recursive`: Whether the archive should apply recursively to nested directories

5.1.3 CloudPools job monitoring

To check the CloudPools job status, use the following command.

`isi cloud jobs list`
To check the archive or recall file list status for a specific CloudPools job, use the following command. As shown in Figure 16, the job id can be found using the command `isi cloud jobs list`.

```
isi cloud jobs files list <job id>
```

![Figure 16](image)

File list of specific CloudPools job

**Note:** The output of the prior command only shows the file name and state for specific CloudPools job.

To perform additional actions, run the following commands:

- **Pause a CloudPools job:**
  ```
  isi cloud jobs pause <job id>
  ```

- **Resume a paused CloudPools job:**
  ```
  isi cloud jobs resume <job id>
  ```

- **Cancel a CloudPools job:**
  ```
  isi cloud jobs cancel <job id>
  ```

- **Check the file list state of writing updated data to the cloud (job id is 1), which is an internal CloudPools job and always running:**
  ```
  isi cloud jobs files list 1
  ```

## 5.2 Troubleshooting

This section describes various CloudPools troubleshooting methodologies, which include:

- CloudPools state
- CloudPools logs

### 5.2.1 CloudPools state

To check the CloudPools storage account state, use the following command:

```
isi cloud accounts view <cloudpools storage account name>
```

To check the CloudPool state, use the following command:

```
isi cloud pools view <cloud pool name>
```

To check the file pool policy state, use the following command:

```
isi filepool policies view <filepool policy name>
```
5.2.2 CloudPools logs

Check the CloudPools logs if needed. The location of CloudPools logs are as follows:

- Most normal daemon log is at /var/log/isi_cpool_d.log
- The log of IO to the cloud is at /var/log/isi_cpool_io_d.log
- Key management log is at /var/log/isi_km_d.log
- CloudPools job (Job Engine) log is at /var/log/isi_job_d.log
A Step-by-step configuration example

This section describes a step-by-step configuration example for CloudPools and Microsoft Azure and includes the following topics:

- Microsoft Azure
- PowerScale configuration
- SmartLink files and cloud data protection

A.1 Microsoft Azure configuration

This section describes the Microsoft Azure configuration for CloudPools.

The example Microsoft Azure configuration is designed to be a general guide when Microsoft Azure is used by CloudPools. It does not cover all details of Microsoft Azure configuration for other use cases. Consult the Microsoft Azure documentation for more details on Microsoft Azure configuration.

1. Make sure your Azure account is working properly.
2. Log in to the Azure portal using your own username and password.
3. In the Azure portal, follow the document Create a storage account on the Azure website to create a storage account.
4. Before configuring CloudPools, use the tool Azure Storage Explorer from Microsoft to get the URI (Blob Endpoint), username (Account Name) and passkey (Primary Key) for CloudPools, shown in Figure 17.
Step-by-step configuration example

A.2 PowerScale configuration

This section describes the CloudPools configuration on a PowerScale cluster, which includes the following:

- Verify licensing
- Cloud storage account creation
- CloudPool creation
- File pool policy creation
- Run SmartPools job for CloudPools
- SyncIQ policy creation

Figure 17  Azure Storage Account Information

Now all Microsoft Azure information is gathered for CloudPools.
A.2.1 Verify licensing

This section describes how to verify licensing on the PowerScale system.

1. Log in to the OneFS WebUI and navigate to **Cluster Management > Licensing** as shown in Figure 18.
2. Verify that the CloudPools and SmartPools license status is **Activated**.

![Figure 18 Verifying licenses](image)

A.2.2 Cloud storage account

This section describes how to create a cloud storage account on the PowerScale cluster.

1. Log in to the OneFS WebUI and navigate to **File System > Storage Pools**. Click **CloudPools** as shown in Figure 19.

![Figure 19 CloudPools](image)

2. Click the **Create a Cloud Storage Account** button from the **Create a Cloud Storage Account** page as shown in Figure 20. The minimum information for CloudPools and Microsoft Azure is as follows:
   - **Name or alias**: Type a name to identify the cloud storage account.
   - **Type**: Select Microsoft Azure.
   - **URI**: Type the URI to connect Microsoft Azure.
Step-by-step configuration example

- **User name (account name)**: Type the account name gathered on the Microsoft Azure portal.
- **Key (access key)**: Type the access key gathered on the Microsoft Azure portal.

![Create a cloud storage account](image)

Figure 20 Create a cloud storage account

3. Click **Connect account** to create a cloud storage account. This results in two buckets being created in Microsoft Azure; one will start with a `d` as a container to store the CDO’s, and the other will start with an `m` as a container to store the associated metadata.

### A.2.3 CloudPool

This section describes how to create a CloudPool for Microsoft Azure on the PowerScale cluster.

1. Log in to the OneFS WebUI and navigate to **File System > Storage Pools**. Click **CloudPools** as shown in Figure 19.
2. Click the **+ Create a CloudPool** button from the **Create a CloudPool** page as shown in Figure 21. The minimum information is as follows:

   - **Name**: Type a name to identify the CloudPool.
   - **Type**: Select Microsoft Azure.
   - **Account in CloudPool**: Select the cloud storage account.
Step-by-step configuration example

3. Click **Create a CloudPool** to create a CloudPool.

**A.2.4 File pool policy**

This section describes how to create a file pool policy on the PowerScale cluster.

1. Log in to the OneFS WebUI and navigate to **File System > Storage Pools**. Click **File Pool Policies** as shown in Figure 22.

![Create a CloudPool](Image)

Figure 21  Create a CloudPool

![Create a file pool policy](Image)

Figure 22  Create a file pool policy
2. Click the + Create a File Pool Policy button, from the Create a file pool policy page as shown in Figure 23 and Figure 24. The minimum information is as follows:

- **Policy Name**: Type a name to identify the file pool policy.
- **File Matching Criteria**: Define a logical group of files for CloudPools. See the section file matching criteria.
- **Move to cloud storage**: Select the specific CloudPool as the CloudPool storage target.
- **Data retention settings**: Set the data retention as your own. See the section Data retention settings.

![Create a file pool policy](image)

Figure 23  Create a file pool policy
Create a file pool policy (continued)

3. Click **Create policy** to create a file pool policy.

**A.2.5 Run SmartPools job for CloudPools**

This section describes how to run a SmartPools job for CloudPools on the PowerScale cluster.

1. Log in to the OneFS WebUI and navigate to **Cluster management > Job operations**. Click **Job types** as shown in Figure 25.

![Job types](image-url)
2. Select the **SmartPools** item and click **Edit** as shown in Figure 26.

![SmartPools job](image)

**Figure 26**  SmartPools job

3. From **Edit job type details** page as shown in Figure 27, you can perform the following:

- Enable or disable the job
- Set the priority of the job
- Set the impact policy
- Set the job schedule as manual or scheduled

![Edit job type details](image)

**Figure 27**  Edit job type details

4. Click **Start job** as shown in Figure 26 to run the file pool policy to archive files from the PowerScale cluster to Microsoft Azure. If you want to start a specific file pool policy job manually, refer to section 5 on **commands and troubleshooting**.

A.2.6  **SyncIQ policy**

This section describes how to create a SyncIQ policy on the PowerScale cluster.

1. Log in to OneFS WebUI and navigate to **Cluster Management > Licensing** as shown in Figure 18. Verify that the CloudPools, SmartPools, and SyncIQ license status is **Activated**.
2. Navigate to **Data Protection > SyncIQ > Policies** and click the + **Create a SyncIQ policy** button as shown in Figure 28 and Figure 29. The minimum information is as follows:

- **Policy name**: Type a name to identify the policy name.
- **Source root directory**: Type the directory name from source PowerScale cluster you want to replicate to the target PowerScale cluster.
- **Target host**: Type the IP or name of the target PowerScale cluster.
- **Target directory**: Type the directory name from the target PowerScale cluster you want to store the data replicated from the source PowerScale cluster.
- **Deep copy for CloudPools**: Select the type you want to use.

![Create SyncIQ policy](image)

Figure 28  Create SyncIQ policy
3. Click **Create policy** to create a SyncIQ policy.

### A.3 SmartLink files protection

This section describes an example of how to protect SmartLink files and cloud data. Ensure that you have already configured SyncIQ on the PowerScale clusters, which includes the following:

- Fail over to the secondary PowerScale cluster
- Fail back to the primary PowerScale cluster
A.3.1 Fail over to the secondary PowerScale cluster

This section describes the steps required to fail over to the secondary PowerScale cluster.

1. Log in to the secondary OneFS WebUI and navigate to Data Protection > SyncIQ. Click Local Targets on the policy that you want to fail over and select More > Allow Writes as shown in Figure 30. This will grant read/write access to the data on the primary PowerScale cluster being replicated to the secondary PowerScale cluster.

![Figure 30](image_url)  
Figure 30 Allow writes on secondary cluster

**Note:** If the primary PowerScale cluster is still online, stop all writes to the replication policy’s directory.

2. Check and change cloud access. Log in to the PowerScale clusters using SSH. To identify the CloudPools GUID, use the commands `isi cloud access list` and `isi cloud access view <GUID>`. Figure 31 shows the cloud access status on the secondary PowerScale cluster.

![Figure 31](image_url)  
Figure 31 Identify CloudPools GUID to be transferred

3. On the primary PowerScale cluster, remove the cloud write permission using the command `isi cloud access remove <GUID>` as shown in Figure 32. This disables the file pool policy, CloudPool, and cloud storage account on the primary PowerScale cluster.

![Figure 32](image_url)  
Figure 32 Remove Cloud write access on the primary PowerScale cluster
4. On the **secondary** PowerScale cluster, add the cloud write permission using the command `isi cloud access add <GUID>` as shown in Figure 33. This enables the file pool policy, CloudPool, and cloud storage account on the secondary PowerScale cluster.

![Figure 33 Add Cloud write access on the secondary PowerScale cluster](image)

**Note:** It is important to not allow write access to the CloudPools from more than one PowerScale cluster.

The SyncIQ failover is complete.

### A.3.2 Fail back to primary PowerScale cluster

This section describes the steps required to fail back to the primary PowerScale cluster.

1. Log in to the **primary** OneFS WebUI and navigate to **Data Protection > SyncIQ**. Click **Policies** on the policy that you want to failback and select **More > Resync-prep** as shown in Figure 34. This will create a SyncIQ replication mirror policy on the secondary PowerScale cluster.

![Figure 34 Resync prep SyncIQ policy on primary PowerScale cluster](image)
2. Log in to the **secondary** OneFS WebUI and navigate to **Data Protection > SyncIQ > Policies**. On the replication mirror policy that you want to failover and select **More > Start Job** as shown in Figure 35. This will sync any changes that have been written to the secondary PowerScale cluster back to the primary PowerScale cluster.

![SyncIQ](image_url)

**Figure 35**  Sync data from secondary PowerScale cluster to primary PowerScale cluster

3. Log in to the **primary** OneFS WebUI and navigate to **Data Protection > SyncIQ**. Click **Local Targets** on the policy that you want to failover and select **More > Allow Writes** as shown in Figure 36. This will grant read/write access to the replication directory back to the primary PowerScale cluster and change the secondary PowerScale cluster’s access to this directory as read only.

![SyncIQ](image_url)

**Figure 36**  Allow writes on primary PowerScale cluster to SyncIQ replication directory

**Note:** If the secondary PowerScale cluster is still online, stop all writes to the replication policy’s directory. Perform a final replication from the secondary PowerScale cluster to the primary PowerScale cluster to ensure both sites are in sync.
4. Check and change cloud access. Log in to the PowerScale clusters using SSH. To identify the CloudPools GUID, use the commands `isi cloud access list` and `isi cloud access view <GUID>`. Figure 37 shows the cloud access status on the secondary PowerScale cluster.

```
hop-isi-p-1# isi cloud access list
Name          GUID          Synced From    State
------------- ----------- ------------- ----------
hop-isi-n     006016694ac2126755c5a15e4a547aba6bb hop-isi-n permitted
hop-isi-p     006048509d1ce325875cd0033f35f8a983de (current) permitted
```

Figure 37  Identify GUID for CloudPools account and file pool policy

5. On the secondary PowerScale cluster, remove the cloud write permission using the command `isi cloud access remove <GUID>` as shown in Figure 38. This disables the file pool policy, CloudPool, and a cloud storage account on the secondary PowerScale cluster.

```
hop-isi-p-1# isi cloud access remove 006016694ac2126755c5a15e4a547aba6bb
Removing access to 006016694ac2126755c5a15e4a547aba6bb will disable the following CloudPool accounts and FilePool policies:
   testaccount (CloudPool Account)
   ecspolicy (FilePool Policy)
Are you sure? [yes/no]: yes
```

Figure 38  Remove cloud write access on the secondary PowerScale cluster

6. On the primary PowerScale cluster, add the cloud write permission using the command `isi cloud access add <GUID>` as shown in Figure 39. This enables the file pool policy, CloudPool, and cloud storage account on the primary PowerScale cluster.

```
hop-isi-n-1# isi cloud access add 006016694ac2126755c5a15e4a547aba6bb
Giving access to 006016694ac2126755c5a15e4a547aba6bb will enable the following CloudPool accounts and FilePool policies:
   testaccount (CloudPool Account)
   ecspolicy (FilePool Policy)
Are you sure? [yes/no]: yes
To ensure proper cleanup, a job must be run for each 33 enabled account to set an expiration date for all stale cloud files. Failure to set an expiration date will cause leaked data in the cloud resulting in additional costs from cloud service providers. Note that after the expiration date has passed, backups may no longer be able to restore deleted files. Expiration dates can be set later using the 'isi cloud restore-cont' command. To start expiration date jobs for applicable accounts, enter an expiration date now or 'default' to accept the default expiration date (2023-03-20): (date)/[default]/[cancel]:
```

Figure 39  Give the primary PowerScale cluster cloud write access

**Note:** It is important to not allow write access to the CloudPools from more than one PowerScale cluster.

The SyncIQ failback is complete.
Technical support and resources

Dell.com/support is focused on meeting customer needs with proven services and support.

Storage technical documents and videos provide expertise that helps to ensure customer success on Dell EMC storage platforms.

B.1 Related resources

Below is a list of documents and other assets that are referenced in this paper along with other resources that may be helpful in deployment of CloudPools on PowerScale.

OneFS CloudPools Administration Guide
OneFS Technical Overview
OneFS 8.2.2 CLI Administration Guide
OneFS 8.2.2 CLI Command Reference
OneFS 8.2.2 Web Administration Guide
Next Generation Storage Efficiency with Dell EMC PowerScale SmartDedupe
Dell EMC PowerScale OneFS Storage Efficiency
Dell EMC PowerScale SyncIQ: Architecture, Configuration, and Considerations
High Availability and Data Protection with Dell EMC PowerScale Scale-out NAS
Storage Quota Management and Provisioning with Dell EMC PowerScale SmartQuotas
PowerScale Non-Disruptive Upgrade (NDU) Best Practices
Data Protection with Dell EMC PowerScale SnapshotIQ
Microsoft Azure