Dell Storage Center OS 7.0 Data Reduction with Deduplication and Compression

Abstract

This paper provides a high-level overview of how data reduction works with the Dell™ Storage Center Operating System (SCOS).

January 2018

Revisions

Date	Revision	Comments
March 2014	1.0	Initial release for Storage Center OS (SCOS) 6.5.1
October 2014	2.0	Updated for SCOS 6.5.10
October 2015	3.0	Updated for SCOS 6.7
May 2016	4.0	Updated for SCOS 7.0 to include deduplication content
September 2016	4.1	Updated performance considerations
January 2017	4.2	Updated Data Progression information
January 2018	4.3	Updated miscellaneous technical details

Acknowledgements

Authors: Marty Glaser, Darin Schmitz

Table of contents

Executive summary

This paper provides a high-level overview of how data reduction works with the Dell™ Storage Center Operating System (SCOS). It provides technical details of the inner workings, configuration tips, and best practices to help alleviate common misconfigurations.

Audience

This document is intended for storage and server administrators that would like to learn more about Dell EMC SC Series data reduction. Users of this document are expected to have intermediate or advanced technical knowledge of SC Series array deployment, configuration, and best practices.

1 Introduction

Data reduction has become an increasingly important consideration in recent years due to the explosive growth of data in our digital-everything age. With the growing amount of data creation, the demands for storage capacity continue to increase rapidly for many administrators. In addition, old data that is archival in nature may have long retention requirements based on company policies, regulations, or service-level agreements. Traditionally, a common strategy is to archive old data to tape in order to free up disk space. However, this can negatively affect recovery time objectives (RTO), and comes with recovery risks that are associated with defective media.

With the advent of increasingly cheaper and higher capacity disks available for a tier 3 storage layer, administrators can store more archival data on the SAN, and leverage remote replication to ensure disaster recoverability of the data.

If we limit the scope of data reduction benefits to SAN storage, the benefits gained from deduplicating and compressing historical data are obvious. This is especially true when considering that, in most cases, about 80 percent of an organization's SAN data is going to be archival in nature. The most appealing benefit of data reduction is that it allows customers to realize higher storage densities as part of maximizing their storage investment.

The SC Series firmware has gradually rolled out data reduction features over multiple releases:

Prior to SCOS 6.5.1: SC Series SAN administrators had to rely on other means such as the OS or application layer to reduce data on the arrays.

SCOS 6.5.1: Data compression was included as a native feature that worked in concert with Data Progression on multi-tier systems to compress frozen inaccessible data at the block level on each volume.

SCOS 6.5.10: Data compression was extended to include support for SC Series arrays with only one tier of storage.

SCOS 6.7: Data compression was further enhanced to extend compression support to include both frozen accessible and frozen inaccessible data, and to support data compression on flash drives in all-flash arrays.

SCOS 7.0: Data deduplication with compression was introduced to further reduce the amount of data stored on an array.

2 Supported configurations

The SC Series data reduction with deduplication and compression feature is available at no cost, to customers that already have a current Dell storage support contract, to upgrade array firmware to a version that supports data reduction.

Data reduction requires the following configuration:

- SCOS version 7.0 or higher is required.
- SC4000, SC5000, SC7000, SC8000, or SC9000 Series controllers support data reduction. These controllers have the extra processor capacity to handle cyclic redundancy checks (CRC) and reduction overhead.

  The following restrictions apply to additional SC Series products:

  - Series 40 or SCv2000 Series controllers: Data reduction is not supported.
  - SCv3000 Series controllers: Only compression is supported, but not deduplication.

- A minimum of six managed SSD (flash) drives plus a hot spare are required. SSDs provide the read performance required for metadata generated by data deduplication.
- Data Progression must be enabled. Systems with only one tier of all-flash storage support data reduction, and do not require a Data Progression license in order for data reduction to work.

The system automatically classifies flash drives as tier 1 storage. Hybrid arrays containing both flash and spinning disks will have a minimum of two storage tiers.

Data reduction occurs on the lowest tier of each media type. For example, on systems with two all-flash tiers, data reduction occurs on tier 2. On hybrid systems with one tier of all-flash and one tier of spinning disks, data reduction will occur on both tiers simultaneously. See Table 1 for an overview on tiering with data reduction.

Table 1 Data reduction support matrix

Supported configuration	Deduplication with compression (SCOS 7.x)
All-flash array (1 tier)	✓ (Tier 1)
All-flash array (2 tiers)	✓ (Tier 2)
Hybrid flash array (2 tiers)	✓ (Tier 1 and Tier 3)
Hybrid flash array (3 tiers: 2 flash, 1 spinning)	✓ (Tier 2 and Tier 3)
All-spinning array	Not supported

3 Data reduction functionality

In order to understand how data reduction with deduplication and compression works in SC Series arrays, first it is important to understand how data is stored at the block level using data pages.

3.1 Data page states

Data on an SC Series array is stored as data pages that can transition between multiple states. These states effect whether or not a page is eligible for data reduction.

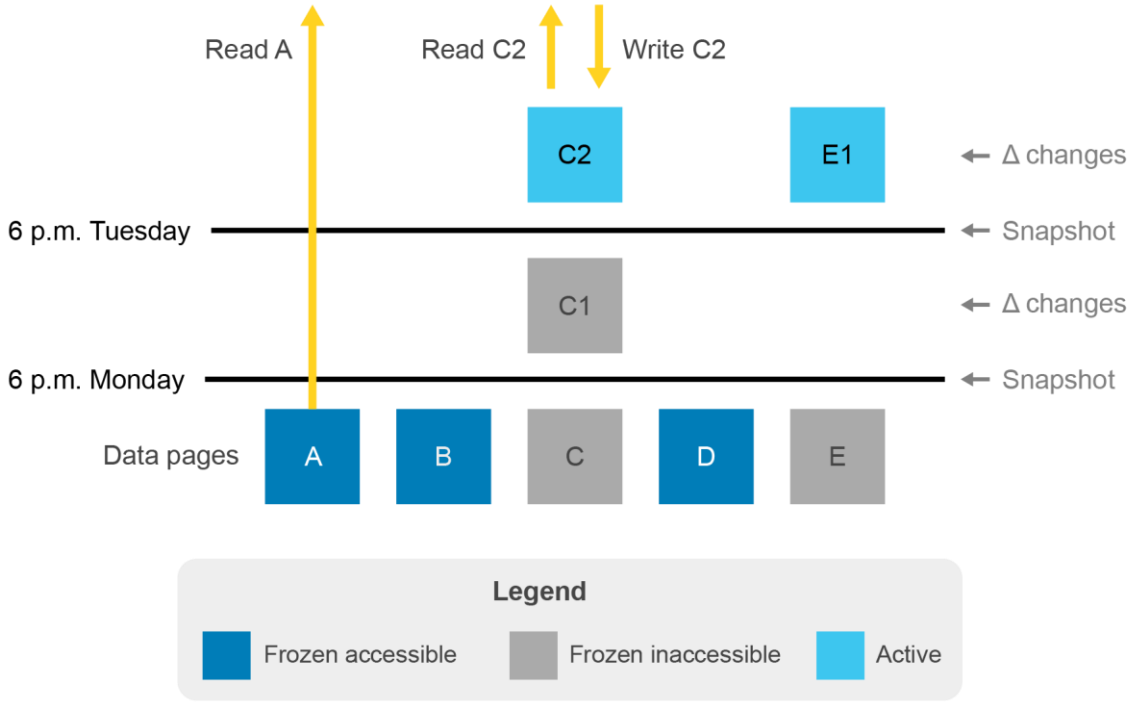

Figure 1 Data page states

Active data pages: Data is first written into an active data page.

Frozen accessible data pages: When a snapshot is taken, the active data page becomes frozen (non-writable by the host), but remains accessible for reads.

Frozen inaccessible data pages: When a host modifies the data contained in a frozen accessible data page, a new active page is allocated to supersede it, before the frozen page becomes a frozen inaccessible page.

Note: It is important to understand the concept of data page states because only frozen accessible and frozen inaccessible data pages are eligible for data deduplication and compression.

3.2 Data Progression and data reduction

Data reduction works in concert with Data Progression. As Data Progression is processing data, the array attempts to reduce eligible frozen data pages that are:

- Already in tier 3 storage (for arrays with tier 3 spinning disks)
- Data progressed from a higher tier down to tier 2 (in all-flash multi-tier arrays) or tier 3 storage
- In single tier all-flash systems as data progresses between page states

The data reduction process also runs during on-demand Data Progression, which is triggered immediately after a snapshot is taken of any volume on the system. Unlike the daily Data Progression cycle that analyzes all frozen data pages in the volume, an on-demand Data Progression cycle only processes data pages that have changed since the last snapshot.

Since SCOS has separate subsystems for both compression and deduplication with compression, if the frozen pages are eligible for reduction, Data Progression redirects them appropriately. For example, if the volume is set to only use compression, the page is forwarded to the compression subsystem, while if the volume is set to use deduplication with compression, data is sent to the deduplication subsystem.

3.2.1 Compression subsystem

When Data Progression forwards the frozen pages to the compression subsystem, it divides pages into 64 KB slices and then compresses each slice. The engine concatenates these compressed slices and sends them to the partial page manager (PPM) for organization and storage.

The data compression process automatically detects if the data in a page is worth compressing and then creates a few metadata tags that facilitate decompression when needed. If the metadata tags along with the newly compressed page require more space than the original compressed page, the data is stored uncompressed.

3.2.2 Deduplication with compression subsystem

The deduplication subsystem divides incoming pages into 4 KB slices and sends them to a computational thread pool for fingerprinting. The process then queries the deduplication dictionary to determine if the fingerprint matches any previously stored data. If the fingerprinted data has been previously stored, the engine does a block-level validation to ensure the dictionary data is correct. If the fingerprint has not been previously stored, the engine packages unique data into compression groups (up to 64 KB in size), where they are compressed. The engine then calculates a final map used to recreate the deduplicated page. From there, the compression group and its map are sent to the deduplication page manager (DPM) which organizes and stores the pages. In the final step, the engine updates all the back references, stores the metadata, and updates the deduplication dictionary with all the locations.

With both subsystems, before the Data Progression cycle completes, a defragmentation cycle runs to make sure all the data is optimally stored.

Data is changed from a reduced state to a hydrated state in the following situations:

- Data reduction is disabled on the volume so that the data can be decompressed and rehydrated during the next Data Progression cycle (starting at 7 p.m. each day by default).
- Data Progression moves the data from a data-reduced tier to a tier that will not reduce data (on multi-tier systems).
- When the data is replicated to another SC Series array, and it is decompressed and rehydrated in memory as it is transferred, but the data stays reduced on the source volume (see section 3.5 for more information).
- A copy of a data-reduced volume is made locally on the same array. The reduced data in the source volume stays reduced.

3.3 Data Progression maximum time limit

By default, Data Progression runs nightly at 7:00 p.m. and the maximum run time is set to unlimited. Depending on the amount of data the reduction process needs to ingest and process, the Data Progression cycle may require significant time to complete. If the data reduction process does not finish within the 24-hour Data Progression timeframe, any remaining data to be reduced will be processed in subsequent Data Progression cycles.

Note: Along with automated page movement and data reduction, Data Progression includes many processes that maximize storage on the SC Series array. Therefore, it is not recommended to set a maximum time limit on Data Progression because these processes may not be allowed enough time to run.

3.4 Data reduction and view volumes

When data reduction is enabled on an SC Series volume, due to inheritance it is also enabled on the view volumes created from the volume snapshots. In other words, view volumes share data reduction settings from the parent (or source) volume. Likewise, when disabling data reduction on a parent volume, it is also disabled on any dependent view volumes. For performance considerations with data reduction and view volumes, see section 6.

3.5 Data reduction and volume replication

When a volume that contains reduced data is replicated to another SC Series array, the reduced pages are rehydrated and decompressed in memory before they are transferred. If the target array supports data reduction, it can be enabled on the destination volume, and the eligible data will be reduced during the next Data Progression cycle.

It is important to note that when creating a replication, there is an option for deduplication (see Figure 2). The replication deduplication works separately from the data reduction deduplication process by analyzing volume snapshots and only replicating changed data. Even with the replication deduplication option enabled, replication of a data-reduced volume still requires the data to be rehydrated and decompressed before the data is replicated.

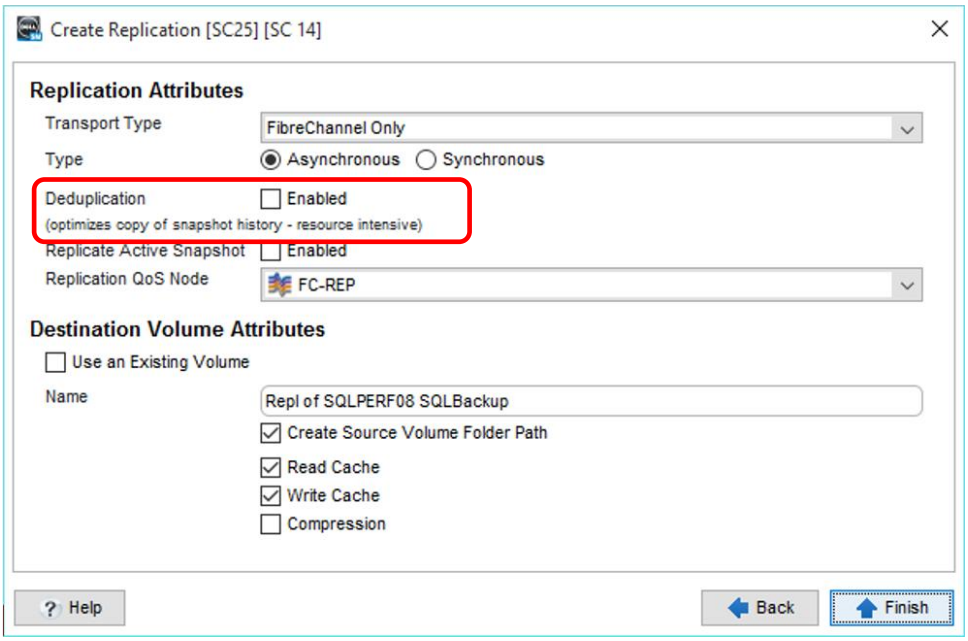

Figure 2 Replication deduplication

3.6 Changing the controller ownership of a data-reduced volume

SC Series bindings prevent changing the controller ownership of a volume that has reduced data (for example, from one controller head to another in a dual pair). If the volume has reduced data, it has to be rehydrated and decompressed before changing controller ownership. Reduced data is tracked per controller, which is why it must be rehydrated and decompressed before changing the ownership.

Complete the following steps to change controller ownership for a data-reduced volume:

1. Make sure that there is sufficient storage capacity to accommodate the rehydrated and decompressed data.
2. Disable data reduction on the desired volume.
3. Allow Data Progression to run (by default, this runs at 7:00 p.m. daily). Allow sufficient time for the entire volume to be rehydrated. Volumes larger in size may take multiple Data Progression cycles.
4. Change ownership to the preferred controller.
5. Re-enable data reduction on the volume (if desired).
6. Allow Data Progression to run overnight. This reduces the eligible data.

Note: As a work-around, create a copy of the reduced volume (assuming there is adequate storage capacity) and assign the copy to the other controller.

Note: Controller failover is not affected when volumes have reduced data. During a failover, the other controller in the pair continues to service I/O to all volumes regardless of their data reduction state or controller ownership.

4 Manage data reduction

Data reduction can be enabled or disabled on multiple volumes simultaneously, or on a volume-by-volume basis. It can also be paused system-wide, on multiple volumes simultaneously, and volume-by-volume. This is useful in cases where an administrator desires to pause all data-reduction-related processor overhead and I/O as a part of troubleshooting, or to give priority to another process or operation temporarily.

4.1 Enable, pause, and disable data reduction on individual volumes

Using the Dell Storage Manager (DSM) client, data reduction can be enabled on an individual volume when it is created or after it has been created by editing the volume settings. To enable data reduction on a volume, click the **Data Reduction Profile** drop-down list and select either **Compression** or **Deduplication with Compression** as shown in Figure 3.

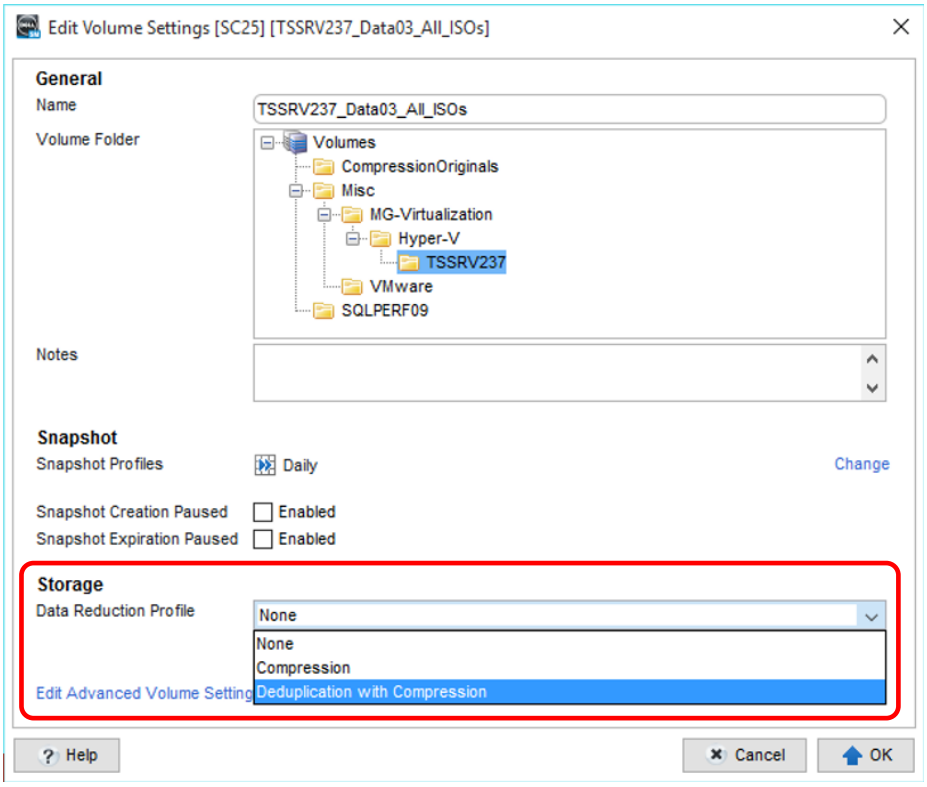

Figure 3 Volume data reduction settings

When enabled on a volume, data reduction processes all snapshot pages (frozen accessible and frozen inaccessible) by default. To change the data reduction input to use only frozen inaccessible pages, click **Edit Advanced Volume Settings**. Click the **Data Reduction Input** drop-down list and select **Inaccessible Snapshot Pages** as shown in Figure 4.

In essence, the Data Reduction Input chooses between the highest deduplication savings (All Snapshot Pages), or the highest performance (Inaccessible Snapshot Pages).

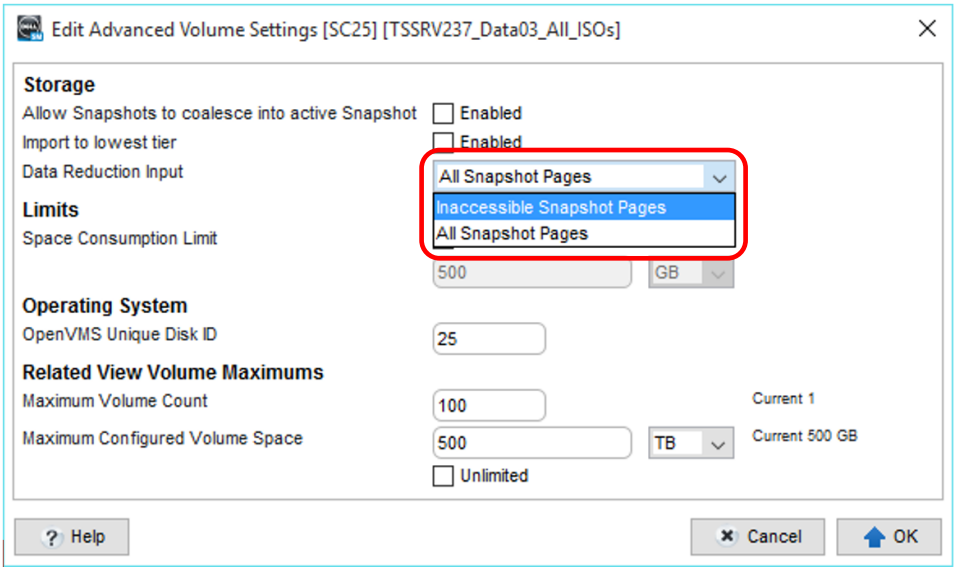

Figure 4 Advanced volume settings

To pause data reduction on individual volumes, check the **Data Reduction Paused** check box as shown in Figure 5. Pausing data reduction stops new data pages from being reduced during the next Data Progression cycle, but does not rehydrate any of the data reduced presently.

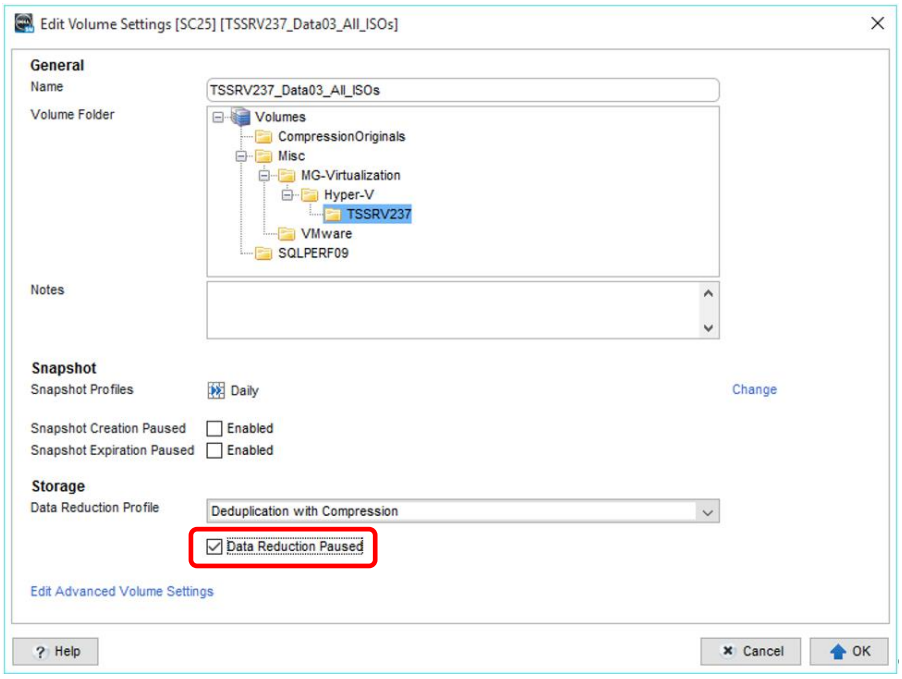

Figure 5 Pause data reduction

To disable data reduction on a volume, select **None** in the **Data Reduction Profile** drop-down list that is shown in Figure 3. The data will be rehydrated and decompressed the next time Data Progression runs. Data is rehydrated and decompressed to the same tier. Therefore, it is important to verify that there is sufficient storage capacity before rehydrating and decompressing a volume. If there is insufficient capacity in tier 3, rehydrated and decompressed data is written to tier 2 or tier 1. If the system detects that rehydrating and decompressing a volume will put the array into conservation mode (critically low on space), the data will stay reduced.

4.2 Pause data reduction system-wide

To pause data reduction for all volumes on an SC Series array, edit the properties and select the **Pause Data Reduction** checkbox as shown in Figure 6.

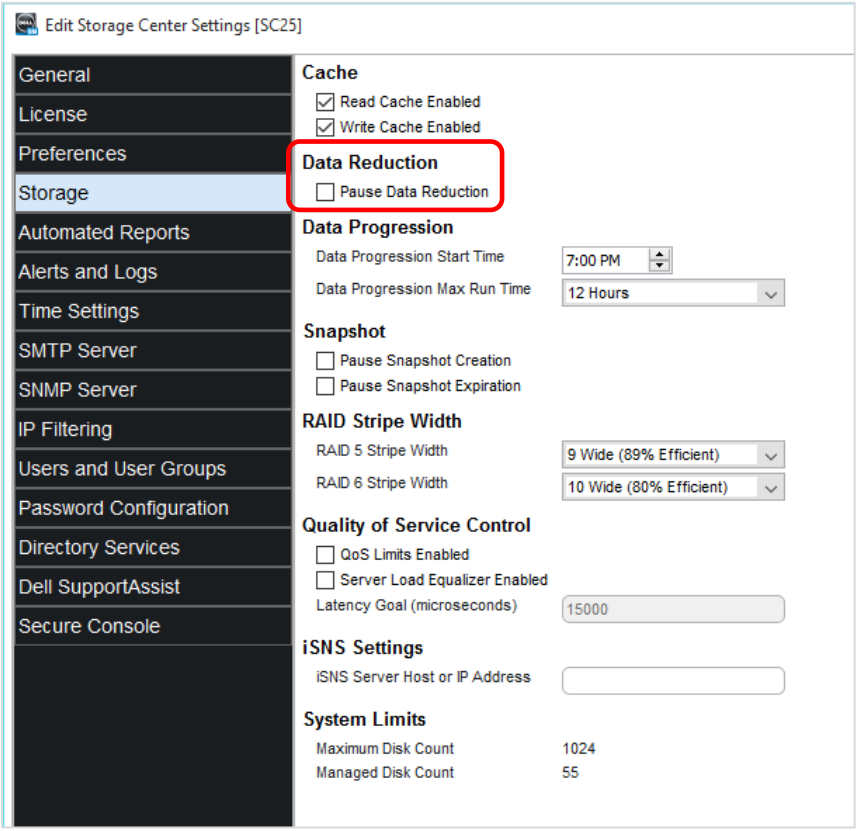

Figure 6 Pause data reduction

Pausing data reduction ensures that no new data is reduced, but does not rehydrate any of the data reduced presently.

4.3 Enable, pause, and disable data reduction on multiple volumes

To enable data reduction on multiple volumes, complete the following steps:

1. Start the DSM client and log in.
2. In the **Storage** tab, right-click the **Volumes** tree and select **Edit Multiple Volumes**.
3. Expand the volume folders to select individual volumes or select the volume folder to include all of those volumes and click **Next**.

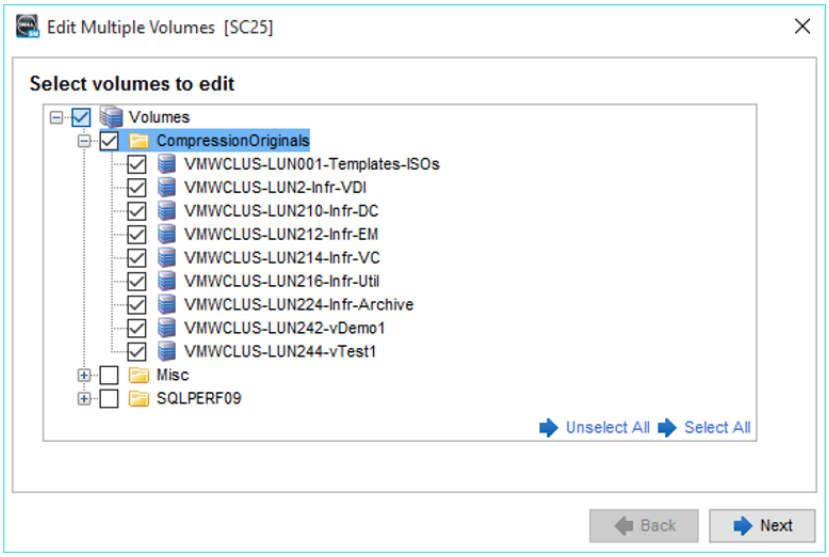

Figure 7 Edit multiple volumes

4. Check the box to apply a data reduction profile, and if desired, select the data reduction input source as shown in Figure 8. Click **Next**.

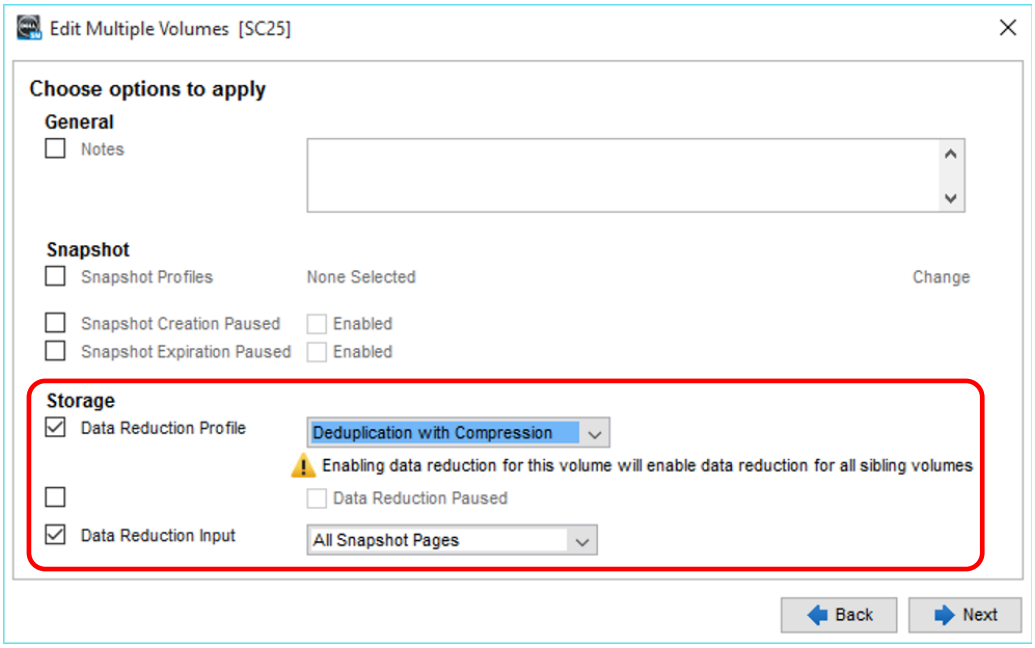

Figure 8 Multiple volume options

5. Review the settings for the selected volumes and click **Next**.
6. Click **Finish**.

To disable data reduction on multiple volumes, select **None** from the **Data Reduction Profile** drop-down list as shown in Figure 9.

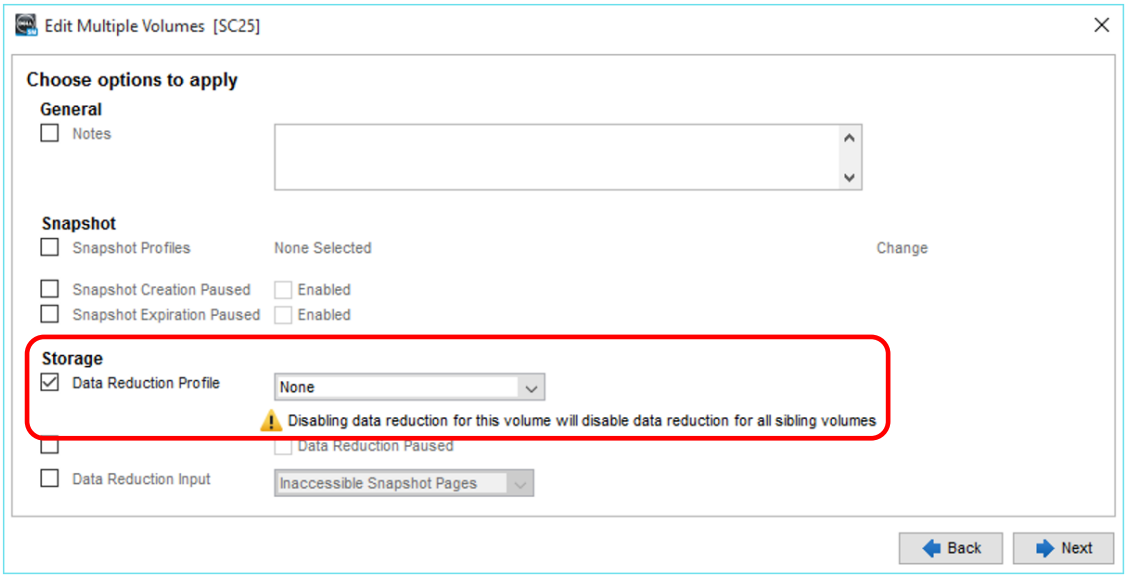

Figure 9 Disable data reduction

To pause data reduction for multiple volumes, select the **Data Reduction Pause** checkboxes as shown in Figure 10.

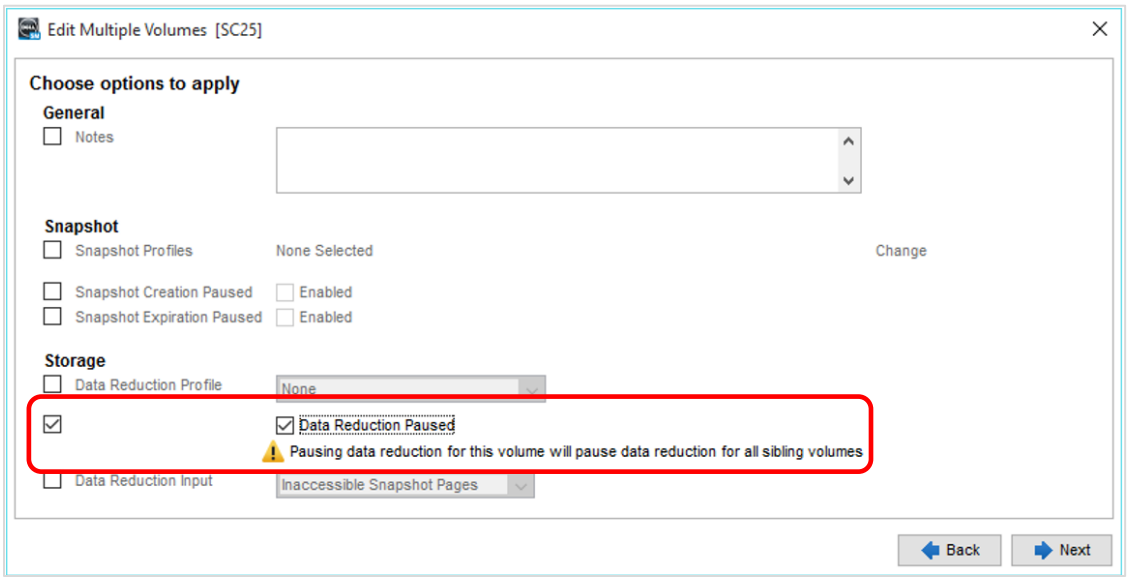

Figure 10 Pause data reduction

5 View data reduction results

SC Series arrays report data reduction results as a percent of overall disk space saved on a volume. Data reduction results can be viewed in DSM.

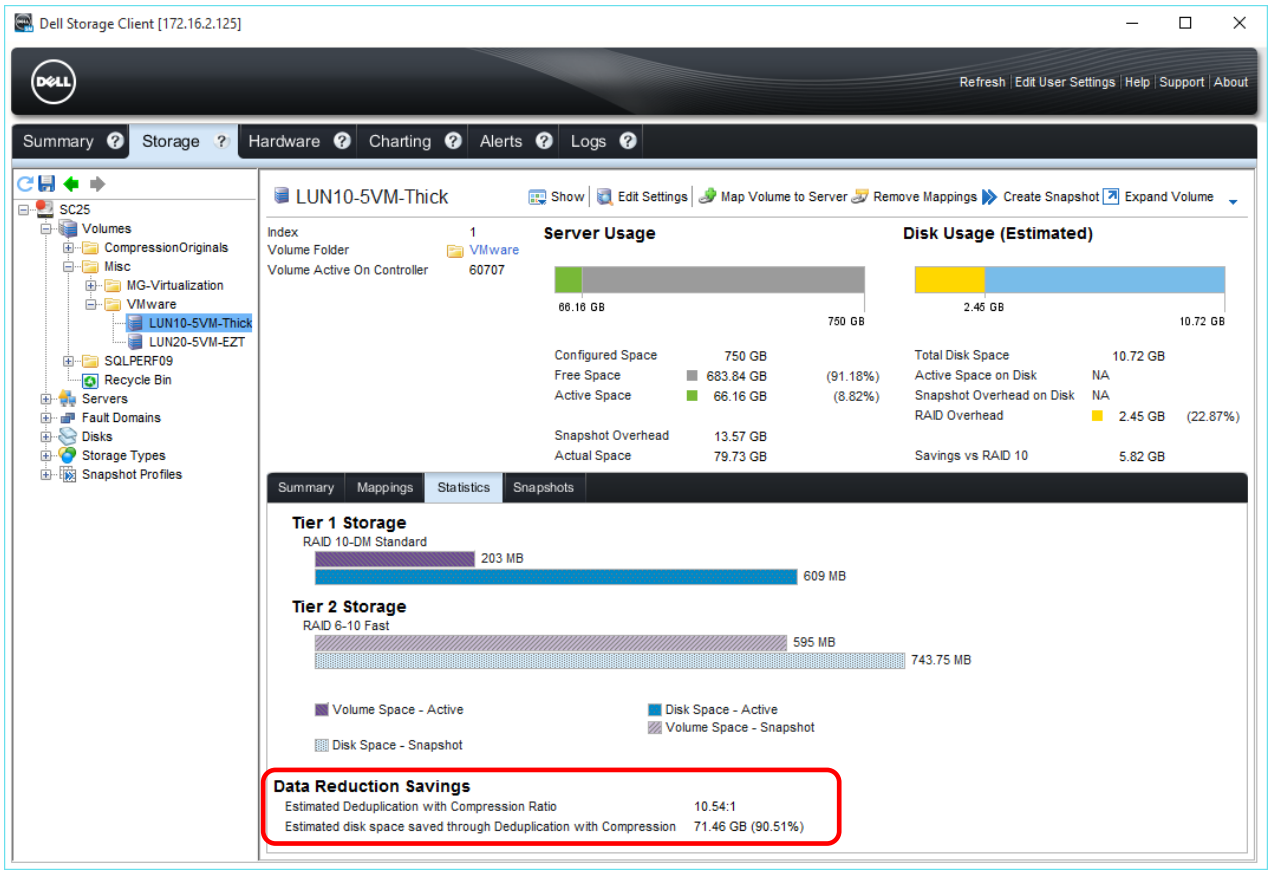

Figure 11 View data reduction savings with the DSM client

To view the data reduction savings for a volume:

1. Log in to DSM and click the **Storage** tab.
2. Expand the **Volumes** tree and select the desired volume.
3. Click the **Statistics** tab to view the data reduction savings for the volume.

To view overall data reduction savings on the system, expand the **Storage Types** tree and click the disk folder name. Data reduction savings display in the **Summary** tab, as shown in Figure 12.

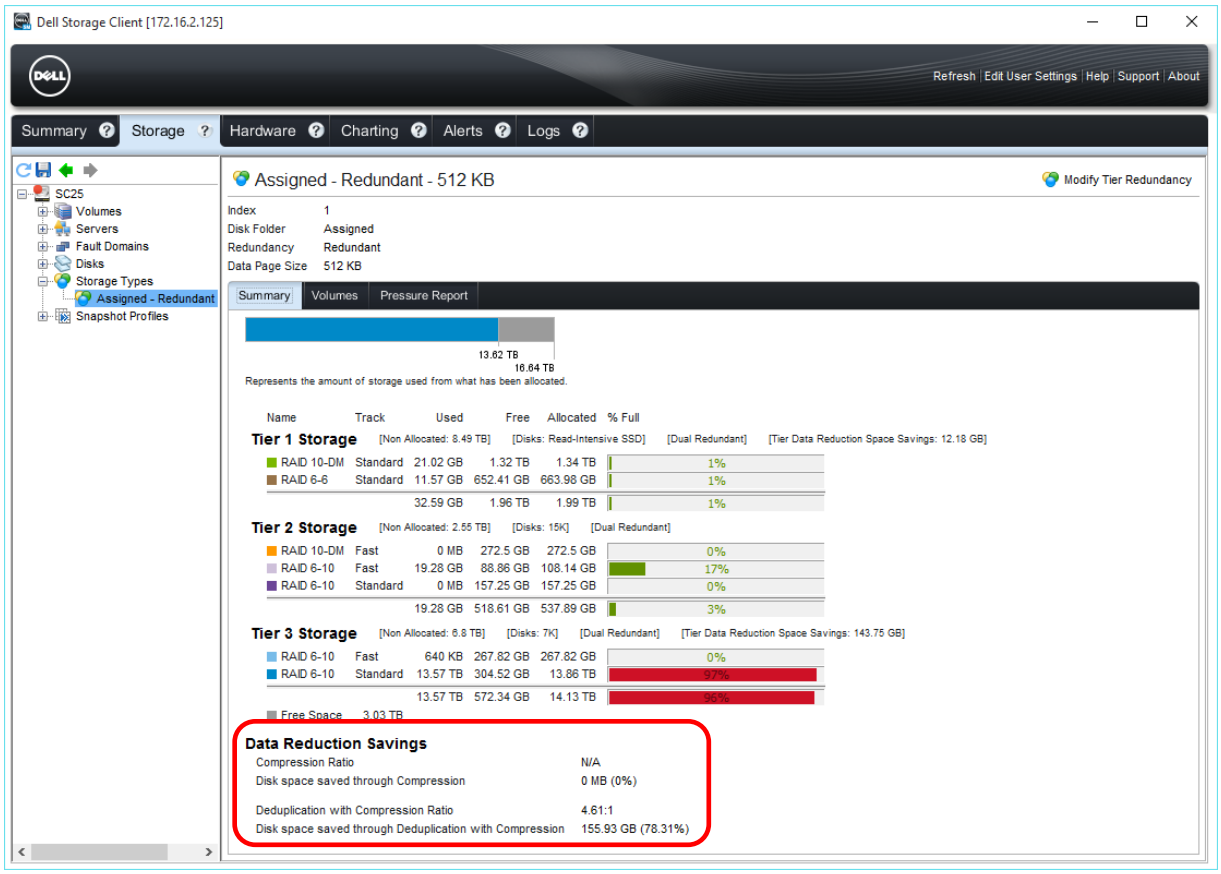

Figure 12 Overall data reduction savings

Note: Data reduction results will not display until the entire volume has been processed during Data Progression.

6 Performance considerations

While the data reduction process runs on dedicated processor cores, there will be a performance impact, generally in terms of read latency. In cases where there may be concerns with Data Reduction effecting performance, temporarily pause all data reduction operations in DSM as shown in Figure 6.

When data reduction is initially enabled on an existing SC Series array that contains a large quantity of eligible data, it may require significant time for data reduction to complete. In this case, it is recommended to enable data reduction on only a few volumes at a time and allow reduction to complete before enabling it on additional volumes.

Since on-demand Data Progression will attempt to reduce delta changes after a snapshot is taken, it is recommended to stagger snapshot creation, as well as expirations.

When considering any performance impact data reduction has on volumes, it is important to adhere to recommended queue depth settings for the host(s) that the volumes are mapped to. For OS-specific best practices and recommendations, refer to the SC Series Technical Documents page.

When adding flash drives to an existing system to enable deduplication, remember that in certain configurations, the flash drives will become a new tier 1. Storage profiles may need to be modified on volumes where administrators intend the data to be directly ingested into tier 3. For example, if flash drives are added to an existing system of all 7k drives, the flash drives will become the new tier 1, and the 7k drives the new tier 3.

When utilizing snapshot profiles or Replay Manager to manage volume snapshots, snapshots should be created with a standard creation method. The standard creation method takes snapshots in series for all volumes associated with the snapshot. Avoid using the parallel snapshot creation method because this method takes snapshots simultaneously for all volumes associated with the snapshot. Simultaneous snapshots of data-reduced volumes can have a negative impact on performance. If possible, create snapshots of data-reduced volumes individually (rather than part of a larger set).

Data reduction cannot be enabled as a way to recover space if an SC Series system is in conservation mode (when it is critically low on disk space). This is because data reduction in and of itself requires some overhead in the form of temporary free disk space while it runs.

The default Data Reduction Input will reduce All Snapshot Pages, which may introduce additional read latency. Volumes that cannot tolerate the additional latency should be set to reduce only the Inaccessible Snapshot Pages.

When data reduction runs on data that has already been reduced by the OS or application, additional data reduction may not be realized on the volume. Determine if enabling data reduction in both places would provide additional benefit for the given environment and data type. The advantages of enabling data reduction at the array level include:

- The ability to pause all data reduction globally with a single check box, as opposed to having to do so server by server, or application by application.
- The ability to offload the processor overhead required for data reduction from the OS or application to dedicated processor cores on SC Series arrays.

The amount of data reduction realized on eligible SC Series data pages is largely dependent of the type of data being reduced. Volumes with files that are text-based typically achieve a higher reduction ratios than audio and video files, which typically are already reduced in size by codecs.

A Technical support and resources

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

Dell TechCenter is an online technical community where IT professionals have access to numerous resources for Dell software, hardware, and services.

Storage Solutions Technical Documents on Dell TechCenter provide expertise that helps to ensure customer success on Dell storage platforms.