Redbooks
ibm.com/redbooks

Introduction and Implementation of Data Reduction Pools and Deduplication

Jon Tate
Dionysios Kalofonos
Carsten Larsen
Atif Syed
Kendall Williams

Storage

IBM

Redbooks

International Technical Support Organization

Introduction and Implementation of Data Reduction Pools and Deduplication

November 2018

SG24-8430-00

Note: Before using this information and the product it supports, read the information in “Notices” on page vii.

First Edition (November 2018)

This edition applies to IBM Spectrum Virtualize V8.1.3.2 (minimum recommended software level) and the associated hardware and software detailed within. Note that the screen captures included within this book might differ from the generally available (GA) version, because parts of this book were written with pre-GA code.

Contents

Notices

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

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

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

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

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

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

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

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

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

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

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

COPYRIGHT LICENSE:

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

Trademarks

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

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

Easy Tier®
FlashCopy®
IBM®
IBM Cloud™
IBM Spectrum™
IBM Spectrum Virtualize™
Real-time Compression™
Redbooks®
Redbooks (logo)  ®
Storwize®
System Storage®

The following terms are trademarks of other companies:

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

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

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

Preface

Continuing its commitment to developing and delivering industry-leading storage technologies, IBM® introduces Data Reduction Pools (DRP) and Deduplication powered by IBM Spectrum™ Virtualize. These are innovative storage features that deliver essential storage efficiency technologies and exceptional ease of use and performance, all integrated into a proven design.

This IBM Redbooks® publication describes Data Reduction Pools (DRP) and Deduplication. It is intended for experienced storage administrators who are fully familiar with IBM Spectrum Virtualize™, SAN Volume Controller, and the IBM Storwize® family of products. For more information about those products, and to supplement your understanding of this book, we suggest you see the following publications:

Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1, SG24-7933

Implementing the IBM Storwize V7000 with IBM Spectrum Virtualize V8.1, SG24-7938

IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines, SG24-7521

Authors

This book was produced by a team of specialists from around the world working at IBM UK's Hursley and Manchester labs.

Jon Tate is a Project Manager for IBM System Storage® SAN Solutions at the ITSO, San Jose Center. Before joining the ITSO in 1999, he worked in the IBM Technical Support Center, providing Level 2/3 support for IBM mainframe storage products. Jon has 32 years of experience in storage software and management, services, and support. He is an IBM Certified IT Specialist, an IBM SAN Certified Specialist, and is Project Management Professional (PMP) certified. He is also the UK Chairman of the Storage Networking Industry Association (SNIA).

Dionysios Kalofonos is a Software Engineer for IBM Systems Storage, IBM Manchester Labs, UK and is working on the SAN Volume Controller and Storwize family of products. Before joining IBM he co-founded a software company, and worked in academia as a researcher.

Carsten Larsen is an IBM Certified Senior IT Specialist working for the Technical Services Support organization in IBM Denmark and delivering consultancy services to IBM clients within the storage arena. Carsten joined IBM in 2007 when he left HP where he worked with storage arrays and UNIX for 10 years. While working for IBM, Carsten obtained several Brocade and NetApp certifications. Carsten is the author of several IBM Redbooks publications.

Atif Syed, at the time of writing, was an IBM Spectrum Virtualize Performance Engineer in IBM Hursley labs, UK. He has 14 years of experience in enterprise infrastructure. Before joining the performance team in 2018, he worked as a level 3 support engineer for five years. From 2004 to 2012 he worked as a business partner consultant for Enterprise Infrastructure in Europe, Middle East, and South East Asia.

He holds a degree in Computer Science from UCP, Pakistan. His areas of expertise include UNIX, Linux, Storage, and backup software. He is an IBM Certified UNIX/Linux, Storage, and TSM backup software.

Kendall Williams III is a Subject Matter Expert and Project Field Engineer working with the IBM Spectrum Virtualize Storage Support Family at IBM. He holds a Bachelors of Science in Information Technology, with a concentration in DB Management Systems and Architecture. His areas of expertise include client performance analysis and copy services support for production environments. Kendall joined IBM in 2012, and has since become an advocate for some premier IBM clients.

Thanks to the following people for their contributions to this project:

Ben Sasson, Carlos Fuente, Jon Parkes, Andy Martin
IBM Hursley, UK

James Whitaker, Vanessa Howling
IBM Manchester, UK

Sergey Kubin
IBM Russia

Now you can become a published author, too

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time. Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us.

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form:

  ibm.com/redbooks

- Send your comments in an email:

  redbooks@us.ibm.com

- Mail your comments:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

Architecture

This section offers a high-level description of Data Reduction Pools and deduplication architecture.

1.1 What Data Reduction Pools are

Data Reduction Pools (DRP) represent a significant enhancement to the storage pool concept. This is because the virtualization layer is primarily a simple layer that runs the task of lookups between virtual and physical extents. Now with the introduction of data reduction technology, compression, and deduplication, it has become more of a requirement to have an uncomplicated way to stay *thin*.

Data Reduction Pools increase existing infrastructure capacity usage by employing new efficiency functions and reducing storage costs. The pools enable you to automatically de-allocate (not to be confused with *deduplicate*) and reclaim capacity of thin-provisioned volumes containing deleted data. In addition, for the first time, the pools enable this reclaimed capacity to be reused by other volumes.

With a new log-structured pool implementation, Data Reduction Pools help deliver more consistent performance from compressed volumes. Data Reduction Pools also support compression of all volumes in a system, potentially extending the benefits of compression to all data in a system.

Traditional storage pools have a fixed allocation unit of an extent, and that itself does not change with Data Reduction Pools. However, features like Thin Provisioning and IBM Real-time Compression™ (RtC) use smaller allocation units and manage this allocation with their own metadata structures. These are described as the *Binary Trees* or *Log Structured Arrays* (LSA).

In order to “stay thin”, you need to be able to reclaim capacity that is no longer used, or in the case of an LSA (where all writes go to new capacity) *garbage collect* the old overwritten data blocks. This also needs to be done at the smaller allocation unit size (KB) per extents.

Figure 1-1 shows the DRP mirroring structure.

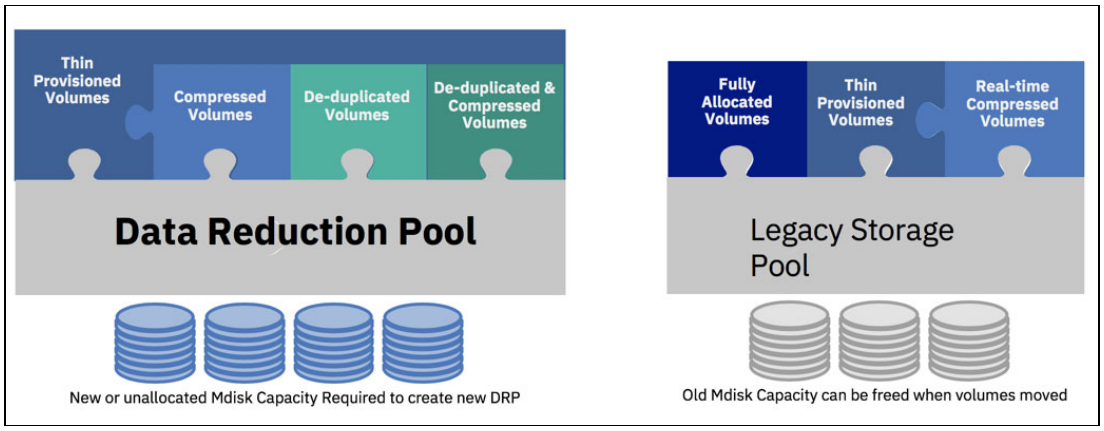

Figure 1-1 New Data Reduction Pool volume mirroring structure

Note: Use volume mirroring to clone data to a new DRP, because DRP does not support **`migrate`** commands.

1.1.1 DRP Volume types

DRP technology enables you to create five types of volumes:

- Fully allocated

  This type provides no storage efficiency, but the best performance, and is available for migration.

- Thin

  This type provides storage efficiency, but no compression or deduplication.

- Thin and Compressed

  This type provides storage efficiency with compression, and this combination provides the best performance numbers.

- Thin and Deduplication

  This type provides storage efficiency, but without compression.

- Thin, Compressed, and Deduplication

  This type provides storage efficiency with maximum capacity savings.

With storage efficiency, DRP thin and compressed volumes provide the best performance numbers. This is due to the new compression implementation. This implementation provides better load balancing and consistent performance when compared to the RACE implementation. This feature is also the second-best performer to fully allocated volumes, followed by thin, compressed, and deduplication volumes when it comes to storage efficiency.

Figure 1-2 shows the types of volumes in the DRP pools.

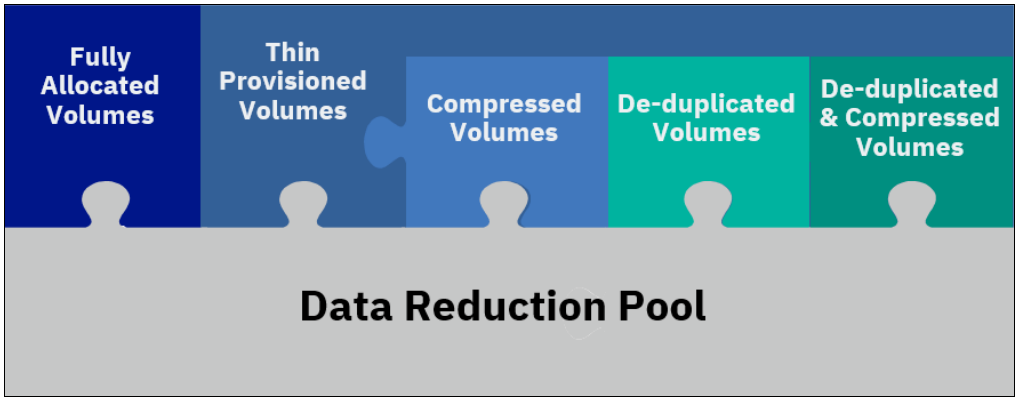

Figure 1-2 Volume types

There are four main characteristics that make up the IBM Data Reduction Pool design:

- Fine-grained allocation of data blocks
- The ability to free back unused (unmapped, or overwritten) capacity at a fine grain
- Give consistent, predictable performance
- Optimize performance for solid-state storage, such as Flash

A Data Reduction Pool, at its core, uses a log-structured array (LSA) to allocate capacity. Therefore, the volume that you create from the pool to present to a host application consists of a directory that stores the allocation of blocks within the capacity of the pool.

All writes for Data Reduction Pools take place at the upper cache layer to the host. Reads have to go through the lower cache layer. The heart of the new DRP functionality is in the new implementation of the Log Structured Array. This includes lower cache, virtualization, IBM Easy Tier®, and RAID. LSA understands what works best for each of these components.

A log structured array allows a “tree” like directory to be used to define the physical placement of data blocks independent of size and logical location. Each logical block device has a range of Logical Block Addresses (LBAs), starting from `0` and ending with the block address that fills the capacity. When written, an LSA enables you to allocate data sequentially and provide a directory that provides a lookup to match the LBA with the physical address within the array.

Note: LSA always appends new data to the end of the array. When data is overwritten, the old location and capacity utilized needs to be marked as `free`. **`UNMAP`** functions can also request that you “free” no longer needed capacity. Compression overwrites can result in a different capacity being used. Deduplication might find new duplicates when data is re-written.

Figure 1-3 shows an IBM Spectrum Virtualize I/O stack structure.

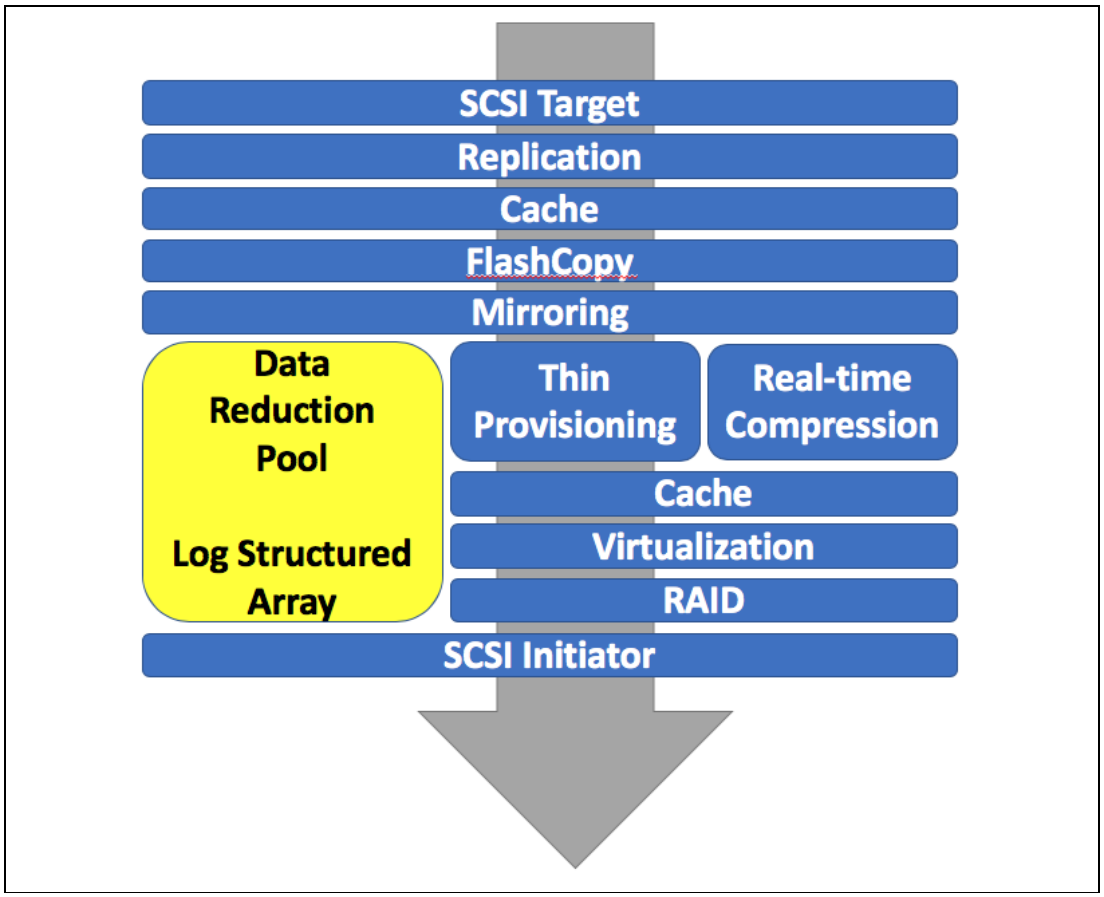

Figure 1-3 Data Reduction Pools/LSA is located in the IBM Spectrum Virtualize I/O stack

1.1.2 What is in a Data Reduction Pool

As shown in Figure 1-4, the user sees a sample of four volumes in a Data Reduction Pool. Internally, there are four directory volumes, one journal volume (per I/O group), one customer data volume (per I/O group), and one reverse lookup volume (per I/O group).

Figure 1-4 shows the view of DRP.

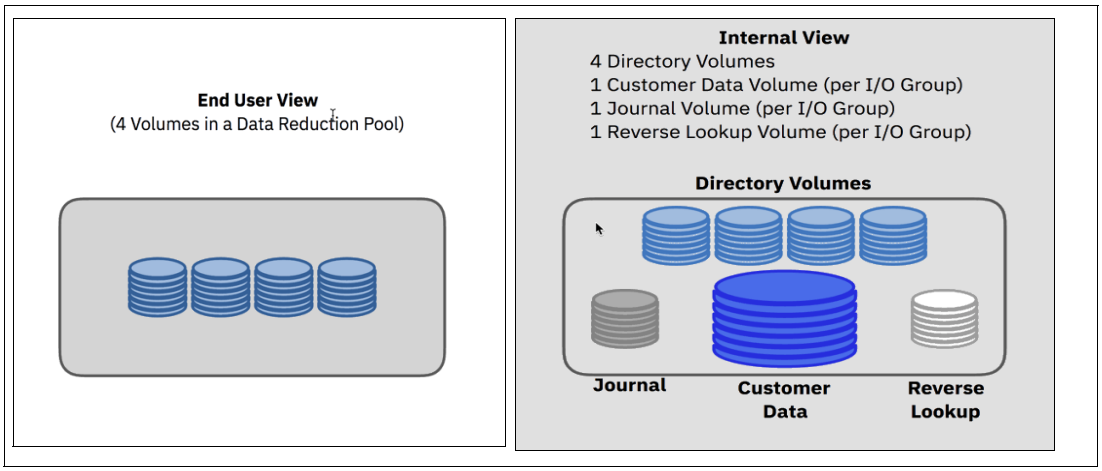

Figure 1-4 Both the front-end and back-end view of DRP

Each Internal volume type has very specific I/O patterns with its own percentage used of the total capacity of the pool, as shown in Table 1-1.

Table 1-1 I/O Patterns per Internal Volumes

Customer Data volumes	Directory volumes	Journal volumes	Reverse Lookup
98% of Pool Capacity	1% of Pool Capacity	Less than 1% of Pool Capacity	Less than 1% of Pool Capacity
Large sequential write pattern and short random read pattern	Short 4 KB random read and write pattern	Large sequential write I/O and only read for recovery scenarios (T3 and so on)	Short, semi-random read/write pattern

1.1.3 Allocation block size

The allocation size of these blocks is now 8 KB. Previously, thin provisioned volumes used 32 KB and RACE Compression write of 32 KB of compressed data. Here are some key reasons behind the 8 KB allocation:

- **UNMAP** requests as small as 8 KB can be catered for.
- The addressability of data in the pool is at an 8 KB (uncompressed) boundary, compared to 32 KB compressed with previous RACE compression.
- All random read requests are of 8 KB size (or less if compressed), which is ideal for Flash Storage.
- With a common metadata access size served by lower cache, performance is much more consistent.

Figure 1-5 shows the DRP Compression I/O Amplification, where *K* is kilobytes.

I/O Type	Space Allocated	I/O Amplification including Size
8K Read	Unallocated	0.50 (1x4K metadata + no data I/O)
	Allocated	1.00 (1x4K metadata + 1x4K data read)
32k Read	Unallocated	0.125 (1x4K metadata + no data I/O)
	Allocated	0.625 (1x4K metadata + 1x16K data read)
8K Write	Unallocated	1.50 (2x4K metadata + 1x4K data write)
	Allocated	1.50 (2x4K metadata + 1x4K data write)
32k Write	Unallocated	0.75 (2x4K metadata + 1x16K data write)
	Allocated	0.75 (2x4K metadata + 1x16K data write)

Figure 1-5 I/O Amplification of space allocated (assumes 50% Compression rate)

Note: Writes to already allocated space will drive the need for Garbage Collection (GC). The cost of GC depends on the amount of valid data in the extent that has not been over-written.

1.2 RACE versus DRP

RACE uses variable input/fixed output, intermittently having to wait or pause to see if more I/O is coming for a volume, as shown in Figure 1-6 on page 7. The RACE minimum block size to read from the backend is 32 KB, and RACE pushes at least 4 - 8 times more data through decompress hardware than DRP for true random workload. In contrast, DRP Compression uses fixed input/variable output.

In Figure 1-7 on page 7, the DRP maximum block size to read from the backend is 8 KB (typically 4 KB or less though). With the use of 8 KB input sizes, there is a small loss in compression ratio but a gain in lower latency when all workload is put into the same predictable small block size. Host I/O that is put in these small grain block size in DRP enables the following functionality:

- Provide fine-grained allocation of block data
- Free unused capacity at a fine grain
- Give consistent, predictable performance
- Optimize performance for Flash storage

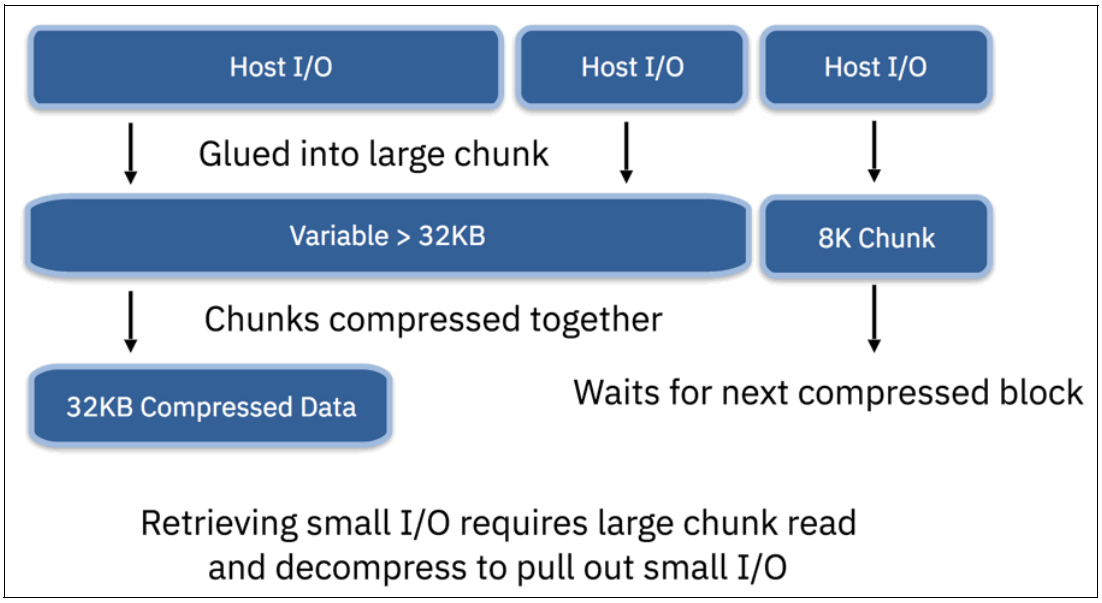

Figure 1-6 RACE compression I/O stack

Figure 1-7 shows the DRP compression I/O stack.

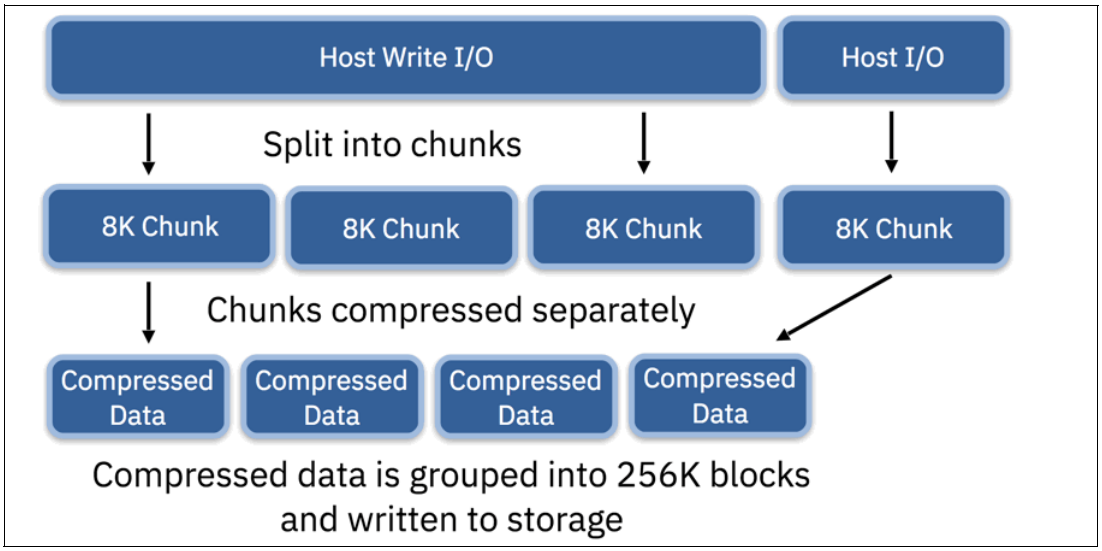

Figure 1-7 DRP compression I/O stack

Here are some of the key differences in DRP when compared to RACE:

- CPU
  - Data reduction uses the same threads as the main I/O process
  - No separate compression CPU utilization
  - No dedicated CPU cores for compression
- Memory
  - Data reduction shares memory with the main I/O process
  - 1 GB memory taken from cache when data reduction is enabled
- Compression hardware
  - Shared with existing RtC compression and compression for IP replication
  - New DRP compression achieves up to 4.8 GBps per node (compression card limit)

1.2.1 Benefits

There are many advantages to Data Reduction Pools, including:

- Designed to be highly scalable to support hardware with more cores and more memory
- Tightly integrated compression shares available cores with other processes for greater efficiency
- Optimization for flash storage through conversion of random write I/Os into larger sequential writes
- No limit on the number of compressed volumes enables greater use of compression (up to 5x as many volumes) and so more compression benefit and reduced storage cost
- Up to 3x better throughput for compressed data, enabling its use with a wider range of data types.
- Ability to release and reuse storage in response to server needs, reducing overall storage required
- Designed for future data reduction technologies
- Separation of metadata and user data improves cache effectiveness
- Compression integrated within I/O stack
- Shared resource design
- Active/Active: Mirrored non-volatile metadata means significantly improved failover/failback response times due to no revalidation of metadata
- No limit on the number of compressed volumes
- Space reclamation: Unmap
- Designed for deduplication
- Smaller 8 KB chunk means less compression bandwidth for small I/Os
- Metadata and user data separated, better use of cache prefetch/destage
- On average 1.8 x I/O amplification on host I/O: Much more predictable latency compared to RACE
- Able to use maximum compression bandwidth
- Compretimator support

Table 1-2 shows DRP disk limitations.

Table 1-2 DRP Disk Limitations

Extent Size	Volume size	4 I/O groups
1 GB	128 TB	512 TB
2 GB	256 TB	1 PB
4 GB	512 TB	2 PB
8 GB	1 PB	4 PB

1.3 Data Reduction Pools and unmap

DRPs support end-to-end unmap functionality. Space that is freed from the hosts is a process called unmap. A host can issue a small file unmap (or a large chunk of unmap space if you are deleting a volume that is part of a data store on a host), and either of these results in the freeing of all the capacity allocated within that unmap. Similarly, deleting a volume at the DRP level frees all of the capacity back to the pool.

When a Data Reduction Pool is created, the system monitors the pool for reclaimable capacity from host unmap operations. This capacity can be reclaimed by the system and redistributed into the pool. Create volumes that use thin provisioning or compression within the Data Reduction Pool to maximize space within the pool.

1.4 Data Reduction Pools with Easy Tier

A DRP uses an LSA as mentioned above. RACE Compression has used a form of LSA since its introduction in 2011, and this means that there is normal garbage collection that needs to be done all the time. An LSA always appends new writes to the end of the allocated space, even if data already exists, and the write is an over-write, the new data is not written in place. Rather, the new write is appended at the end and the old data is marked as needing garbage collected. This itself allows for the following functionality:

- Writes to a DRP volume are always sequential, so we can build all of the 8 KB chunks into a larger 256 KB chunk and destage the writes from cache, either as full stride writes, or as large 256 KB sequential stream of writes.
- This should give the best performance both in terms of RAID on backend systems, and on Flash, where it is easier for the Flash device to also garbage collect on a larger boundary.
- We can start to record metadata about how frequently certain areas of a volume are over-written.

We can then bin sort the chunks into a heat map in terms of rewrite activity, and then group commonly rewritten data onto a single extent. This is so that EasyTier will operate correctly for not only read data, but write data when data reduction is in use. Previously, writes to compressed volumes held lower value to the Easy Tier algorithms, because writes were always to a new extent, so the previous heat was lost.

Now, we can maintain the heat over time and ensure that frequently rewritten data gets grouped together. This also aids the garbage collection at the virtualize level and also on Flash, where it is likely that large contiguous areas will end up being garbage collected together.

1.5 Garbage collection

DRP has built in services to enable garbage collection (GC) of unused blocks. This means that lots of smaller unmaps end up enabling a much larger chunk (extent) to be freed back to the pool. If the storage behind IBM San Volume Controller (SVC) supports unmap, we pass an unmap command to the backend storage (again equally important with today's Flash backend systems) especially so when they themselves implement some form of data reduction.

Trying to fill small holes is very inefficient: too many I/Os would be needed to keep reading and rewriting the directory. So, GC waits until an extent has many small holes. Move the remaining data in the extent, compact, and rewrite. When we have an empty extent, it can be freed back to the virtualization layer (and back end with **UNMAP**) or start writing into the extent with new data (or rewrites).

The reverse lookup metadata volume tracks the extent usage, or more importantly the holes created by overwrites or unmaps. GC looks for extents with the most unused space. After a whole extent has had all valid data moved elsewhere, it can be freed back to the set of unused extents in that pool, or it can be reused for new written data.

1.6 Data Reduction Pools with deduplication

Deduplication can be configured with thin-provisioned and compressed volumes in Data Reduction Pools for added capacity savings. The deduplication process identifies unique chunks of data, or byte patterns, and stores a signature of the chunk for reference when writing new data chunks.

If the new chunk's signature matches an existing signature, the new chunk is replaced with a small reference that points to the stored chunk. The same byte pattern might occur many times resulting in the amount of data that must be stored being greatly reduced.

Duplicate matches are found using SHA1 hashes created for each 8 KB align region of client data to a deduplicated copy. The matches are detected when the data is written. For Data Reduction Pools, deduplication data can work in two separate ways. It can be grouped into 256 KB blocks and written to storage, or can be passed as 8 KB chunks, and compressed first.

Deduplication has specific I/O characteristics in the handling of data and data copies. When a matching fingerprint is found, the metadata is updated to point to the metadata of the existing copy of the data. Each copy of the data can have up to 255 8 KiB virtual chunks referring to it. Each virtual 8 KiB chunk can track up to 3 versions of data. I/O performance takes precedence over finding duplicate copies of data. Host I/Os smaller than 8 KiB do not attempt to find duplicates.

2

Planning

This chapter describes general guidelines and important considerations for using the Data Reduction Pools (DRP) functionality available for the IBM Spectrum Virtualize family of products.

This chapter assumes that the reader is already familiar with IBM Spectrum Virtualize Real-time Compression and Easy Tier. Information on these technologies can be found in many sources, including the following publications:

- *IBM Real-time Compression in IBM SAN Volume Controller and IBM Storwize V7000*, REDP-4859
- *Implementing the IBM Storwize V7000 with IBM Spectrum Virtualize V8.1*, SG24-7938
- *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933
- *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521

Recommended software level: The minimum recommended software level is V8.1.3.2, which is available from IBM Fix Central and contains a number of critical fixes.

2.1 Data Reduction Pools benefits

Data Reduction Pools are a new type of storage pool that implement several techniques, such as thin-provisioning, compression, and deduplication, to reduce the amount of physical capacity required to store data. Savings in storage capacity requirements translate into reduction in the cost of storing the data.

The cost reductions that are achieved through software can facilitate the transition to all Flash storage. Flash storage has lower operating costs, including lower power consumption, cheaper cooling, and higher density. However, the cost of Flash storage is still higher than Nearline Disk. With other technologies, such as DRP, the cost difference can be reduced to a point where an all-Flash solution is feasible.

The first benefit of DRP is in the form of storage savings due to deduplication. The deduplication process identifies unique data patterns, and stores a signature of the data for reference when writing new data. If the signature of the new data matches an existing signature, the new data is not written to disk, but instead a reference to the stored data is written. The same byte pattern can occur many times, resulting in the amount of data that must be stored being greatly reduced.

The second benefit of DRP comes in the form of performance improvements due to compression. Where deduplication aims to identify the same data elsewhere in the storage pool and create references to the duplicate data instead of writing extra copies, compression tries to reduce the size of the host data that is written. Compression and deduplication are not mutually exclusive, one or both or neither features can be enabled. If the volume is deduplicated *and* compressed, data is deduplicated first, and then compressed. Therefore, deduplication references are created on the compressed data stored on the physical domain.

DRP offers a new implementation of data compression that is integrated into the I/O stack. The new implementation makes better use of the system resources and, similar to RACE, uses hardware accelerators for compression. As with RACE, the new implementation uses the same Lempel-Ziv (LZ) based Real-time Compression and decompression algorithm. However, in contrast to RACE compression, DRP compression operates on smaller block sizes, resulting in additional performance gains. For more information about these performance gains, see Chapter 5, “Best practices and guidelines” on page 43.

The third benefit of DRP is performance improvements due to Easy Tier. The metadata of Data Reduction Pools do not fit in RAM, so they are stored on disk, in metadata volumes that are separate from data volumes. The metadata volumes of Data Reduction Pools are small and frequently accessed. They are very good candidates for promotion through Easy Tier. In contrast, data volumes are large, but because metadata is stored separately, Easy Tier is able to accurately identify frequently used data. Performance gains are expected as Easy Tier promotes metadata and customer data to the fastest available storage tier.

2.2 Planning considerations

This section provides detailed considerations directly related to DRP.

2.2.1 Data Reduction Pool characteristics

Data Reduction Pools can contain generic (thick) volumes. The space for these volumes is claimed at creation time. Thin-provisioned volumes, on the other hand, delay allocating storage until it is first written. Therefore, unused space in volumes presented to host servers does not actually use any physical storage capacity. To create a thin-provisioned volume, the user specifies both the real and virtual capacity of the volume when it is created. The real capacity defines how much physical storage is initially associated with the volume, and the virtual capacity defines how large the volume will appear to the host server.

A Data Reduction Pool makes use of metadata, labeled as *overhead* in Figure 2-1. Even when there are no volumes in the pool, some of the space in the pool is used to store the metadata. The space allocated to metadata is relatively small. Regardless of the type of volumes that the pool contains, metadata are always stored separately from customer data.

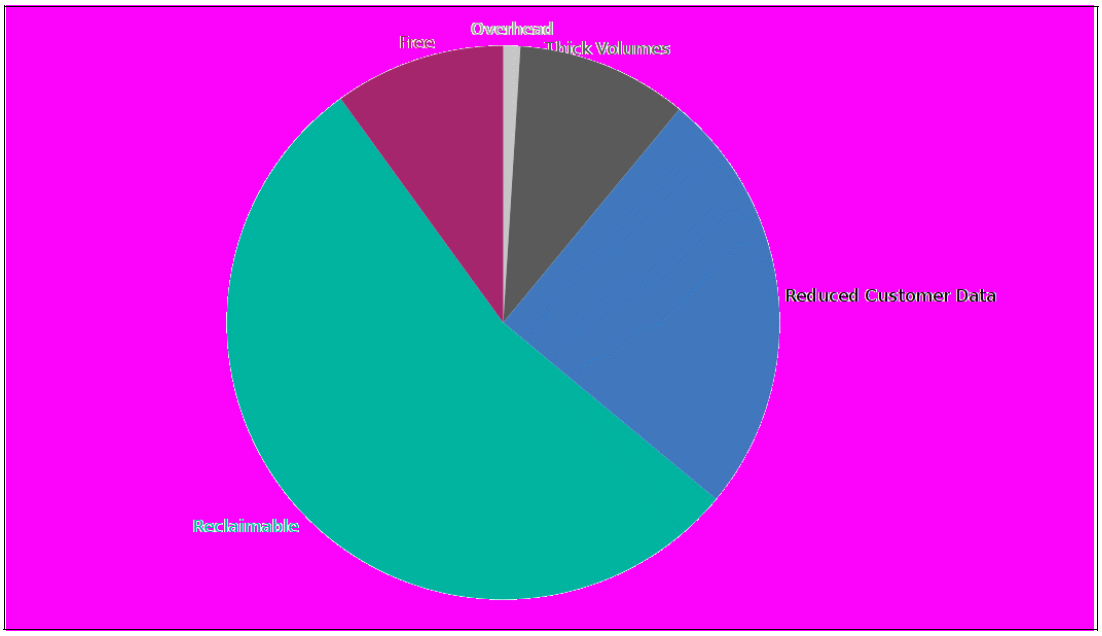

Figure 2-1 Example of capacity use in a Data Reduction Pool

In Data Reduction Pools, the maintenance of the metadata results in *I/O amplification*. I/O amplification occurs when a single host-generated read or write I/O results in more than one back-end storage I/O request due to advanced functions. A read request from the host results in two I/O requests, a directory lookup and a data read. A write request from the host results in three I/O requests, a directory lookup, a directory update, and a data write.

The *reclaimable* space refers to garbage collection. As a result of compression and deduplication, overwriting host writes does not always use the same amount of space that the previous data was using. This leads to these writes always having to occupy new space on back-end storage while the old data is still in its original location. The primary task of garbage collection is to track all of the regions that have been invalidated, and to make this capacity usable for new writes.

For garbage collection, stored data is divided into regions. As data is overwritten, a record is kept of which areas of those regions have been invalidated. Regions that have many invalidated parts are potential candidates for garbage collection. When the majority of a region has invalidated data, it is fairly inexpensive to move the remaining data to another location, therefore freeing the whole region.

Because garbage collection needs to move data in order to free regions, it is suggested that you size pools so that the occupied capacity is 85% of the size of the pool. This practice ensures that there will always be some free space for garbage collection. For more information, see Chapter 5, “Best practices and guidelines” on page 43.

2.2.2 Migration

A Data Reduction Pool is a new type of pool that supports thin provisioning, compression, and deduplication. Any or all of these features can be enabled. A volume created in a Data Reduction Pool can be either fully allocated or thin provisioned. It can be compressed or uncompressed. Deduplication can be enabled or not.

Even though data can be migrated regardless of the nature of the new volumes, the type of these volumes determines the migration strategy used. RACE and DRP compression cannot coexist within the same MDisk group, or some systems do not have enough hardware resources to run RACE and DRP compression at the same I/O group. In these cases, additional resources (storage or hardware) are required to migrate the data to Data Reduction Pools.

Migration strategy for uncompressed volumes

This approach works on all platforms because there is no restriction to using thin volumes in Data Reduction Pools. In this case, the following migration strategy is advised:

1. Create a new Data Reduction Pool and add storage to it.
2. Add a volume copy (from the volume whose data is migrated) to the Data Reduction Pool.

Migration strategy for hardware compression

Data reduction compression cannot be used at the same time as RACE compression within the same MDisk group, but it can be used in the same I/O group if there are hardware compression capabilities. In this case, the following migration strategy is advised:

1. Create a new Data Reduction Pool and add storage to it.
2. Add a volume copy (from the volume whose data is migrated) to the Data Reduction Pool with compression enabled.

Migration strategy for software compression for Storwize V5030

Storwize V5030 does not support hardware compression. The platform is limited to two I/O groups, and it cannot be clustered with another node type that supports hardware compression.

This is the only platform that supports Data Reduction Pools with software compression. However, the platform does not have the CPU resources to run RACE and DRP compression within the same I/O group at the same time. In this case, a two-step migration approach is required.

Decompress and recompress

The first approach is to use a thin Data Reduction Pool, or a regular pool to migrate the data, as shown in the following steps:

1. Create a new Data Reduction Pool or a regular pool.
2. Add a thin copy in the new pool, and use the `autodelete` flag.
3. Add a compressed copy within the new pool, and use the `autodelete` flag.
4. Delete the original MDisk group.

The downside of this approach is that volumes cannot be migrated one at a time, because all RACE volumes need to be removed before a compressed volume in a Data Reduction Pool is created. For this approach, the available storage to decompress all of the RACE volumes needs to be available.

Migrate data across I/O groups

For this approach, two I/O groups are needed, one containing the RACE volumes, and the other the DRP volumes, as shown in the following steps:

1. Move all of the RACE volumes under a single I/O group.
2. Create a Data Reduction Pool under the I/O group that has no RACE volumes.
3. Use intracluster Global Mirror or Metro Mirror to mirror the compressed volumes from the I/O group with RACE to the I/O group with DRP.
4. Use intracluster Global Mirror or Metro Mirror to mirror any uncompressed volumes that need migration.

The downside of this approach is that it requires the hardware to set up a second I/O group.

Migration strategy for copy services

These are the migration strategies for copy services.

Volume mirroring

Systems that already make use of volume mirroring need to have the replication broken during the migration process. When the migration process is complete, the replication can be re-established. This is particularly important for systems with stretched clusters.

Global Mirror/Metro Mirror/Hyperswap

Systems with dual site replication do not have to break the replication during the migration. The exception is Storwize V5030 systems. Storwize V5030 systems that use Global Mirror, Metro Mirror, or Hyperswap do have to break the replication during the migration process. When the migration is complete, replication can be re-established.

2.2.3 Limits

The following limits apply to the previously described scenarios.

Maximum number of volumes

As of IBM Spectrum Virtualize version 8.1.3, 12 internal copy volume IDs are allocated to each Data Reduction Pool. Table 2-1 lists the maximum number of (compressed) volumes per Data Reduction Pool.

Table 2-1 Maximum number of compressed volumes per DRP

Product	Platforms	Volume Copies	Volumes
SVC	ALL	10,000	9,988
V9000	ALL	10,000	9,988
V7000	ALL	10,000	9,988
V5000	V5030	8,192	8,180
	V5020	8,192	8,180
	V5010	4,096	4,084
	V5000	8,192	8,180

Maximum number of DRPs

The maximum number of Data Reduction Pools that can be created in a cluster is 4.

Maximum volume copy size

There is a limit on the maximum size of a customer data volume at 128,000 extents per I/O group. This places a limit on the maximum physical capacity in a pool before data reduction, as shown in Table 2-2 (note that it includes garbage collection (reclaimable capacity) as well).

Table 2-2 Maximum physical capacity in a pool before data reduction

Extent Size	1 DRP - 1 I/O group	1 DRP - 4 I/O groups	4 DRP - 4 I/O groups
1 GB	128 TB	512 TB	2 PB
2 GB	256 TB	1 PB	4 PB
4 GB	512 TB	2 PB	8 PB
8 GB	1 PB	4 PB	16 PB

If the pool reaches an out of space capacity, the whole pool goes offline.

2.2.4 Compression support

Table 2-3 summarises the hardware types that support RACE compression, DRP compression, and RACE plus DRP compression at the same time.

Table 2-3 Hardware types and compression supported

Hardware type	RACE compression support	DRP compression support	RACE plus DRP compression support
Storwize V3500, V3700, V5000, V5010, V5010E, V5020	No	No	No
Storwize V5030, V5030E (16 GB)	No	No	No
Storwize V5030E (32 GB)	No	Yes	No
Storwize V5030 (32 GB)	Yes	Yes	No
SVC DH8 or SV1 nodes with compression hardware, V7000 Gen2	Yes	Yes	Yes
Storwize V5100, V7000 Gen3, FS9100	No	Yes	No
Storwize V7000 Gen 1, CF8, CG8	Yes	No	No
SVCasSW	Dual CPU only	Dual CPU only	No

3

Implementation

This chapter describes data reduction, requirements, Data Reduction Pool (DRP) creation, volume creation, volume migration, and DRP limitations.

Recommended software level: The minimum recommended software level is V8.1.3.2, which is available from IBM Fix Central and contains a number of critical fixes.

3.1 About data reduction

Using data reduction reduces the amount of physical storage required and therefore saves money. Although the price of flash storage is decreasing, there is still a 10x difference in cost between flash and nearline disk.

Using compression and deduplication, it is common to achieve 3x to 5x data reduction ratios. With the addition of thin provisioning, the data reduction ratio is often up to 10x.

In addition, flash storage has lower operating costs: lower power, cheaper to cool, and higher density. Data reduction can therefore be used to facilitate a transition to all-flash storage.

Data reduction with the IBM Spectrum Virtualize family of products includes the following techniques:

- Thin provisioning

  Capacity is allocated on-demand as storage is first written to
- Compression

  Data is compressed before being written to storage
- Deduplication

  Duplicates of data are detected and are replaced with references to the first copy

Data Reduction Pools

Data Reduction Pools increase existing infrastructure capacity usage by leveraging new efficiency functions. Data Reduction Pools enable you to automatically deallocate and reclaim capacity of thin-provisioned volumes containing deleted data. The pools then enable this reclaimed capacity to be reused by other volumes. With the new log-structured pool implementation, Data Reduction Pools help deliver more consistent performance from compressed volumes. Data Reduction Pools also support compressing all volumes in a system, potentially extending the benefits of compression to all data in a system.

Requirements

Compression and deduplication have the following software and hardware requirements:

- Enabled Compression license
- Code level V8.1.3.2 or later is needed for Data Reduction Pools
- Code level V8.1.3.2 or later is needed for deduplication
- Nodes must have at least 32 GB memory to support deduplication
- Nodes that have more than 64 GB memory can use a bigger deduplication fingerprint database, which might lead to better deduplication

Random Access Compression Engine (RACE) compression and DRP compressed volumes can coexist in the same I/O group. However, deduplication is not supported in the same I/O group as RACE compressed volumes.

Review or enable the compression license

The compression license can be reviewed and enabled from the menu **Settings** → **System** →**Licensed Functions**, as shown in Figure 3-1.

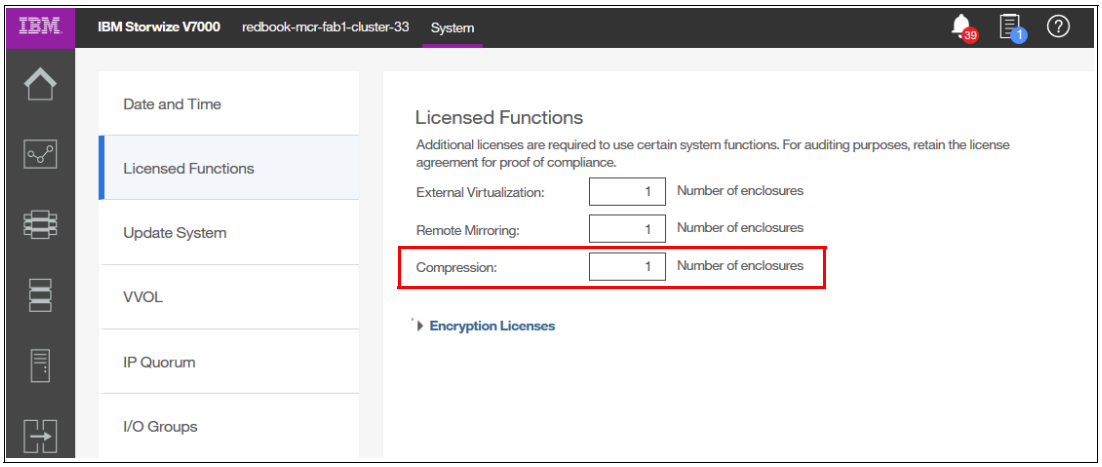

Figure 3-1 Review compression license

As shown in Figure 3-1, we have enabled a compression license for a system with a single enclosure. If you add enclosures to the system, you must add the compression license for additional enclosures.

Volume compression

Volume compression with the IBM Spectrum Virtualize family of products has traditionally been done with RACE software and compression acceleration cards. From V8.1.2 of IBM Spectrum Virtualize, software technology is moving away from RACE and towards Data Reduction Pools. Data Reduction Pools use the same hardware as RACE, but the compression software has changed to reflect today's requirements for faster and more efficient storage.

Deduplicated volumes

A deduplicated volume or volume copy can be created in a Data Reduction Pool. When you implement deduplication, you must consider your specific requirements in the storage environment.

Deduplication is a type of compression that eliminates duplicate copies of data. Deduplication of user data occurs within a storage pool, and only between volumes or volume copies that are marked as deduplicated. However, there is no requirement for all nodes in a system, and therefore all I/O groups, to support deduplication. You can create deduplicated volumes in an I/O group when no compressed volumes or volume copies are in regular storage pools (when RACE compression is in use on that I/O group). DRP compressed volumes can coexist with RACE compressed volumes in the same I/O group.

You can migrate any type of volume from a regular storage pool to a Data Reduction Pool. You can also migrate any existing RACE compressed volume to a Data Reduction Pool. After you migrate a volume to a Data Reduction Pool, you can then create a deduplicated volume.

Note: RACE compression and DRP compressed volumes can coexist in the same I/O group, however deduplication is not supported in the same I/O group as RACE compressed volumes.

3.2 Creating Data Reduction Pools

Data Reduction Pools has been available since V8.1.2 of IBM Spectrum Virtualize software. Deduplication was added in version V8.1.3.

Recommended software level: The recommended software level is V8.1.3.2, which is available from IBM Fix Central and contains a number of critical fixes.

In the following section, we demonstrate how to create a Data Reduction Pool and add storage to it.

Figure 3-2 shows an IBM Storwize V7000 *Pools* view where a single standard pool exists.

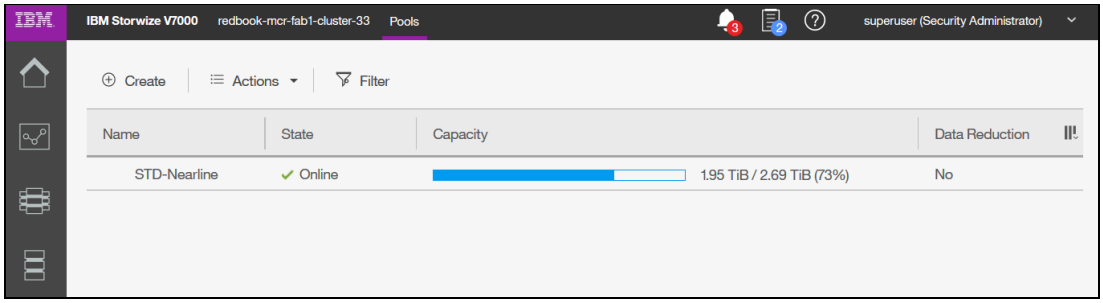

Figure 3-2 Storage Pools view

Complete the following steps:

1. Click **Create** to open the Create Pool wizard. We provide the name `DRP-10KSAS` to reflect that this is a Data Reduction Pool containing SAS-drives 10K RPM speed. We also select the Data reduction **Enable** check box.

   Figure 3-3 shows the Create Pool wizard.

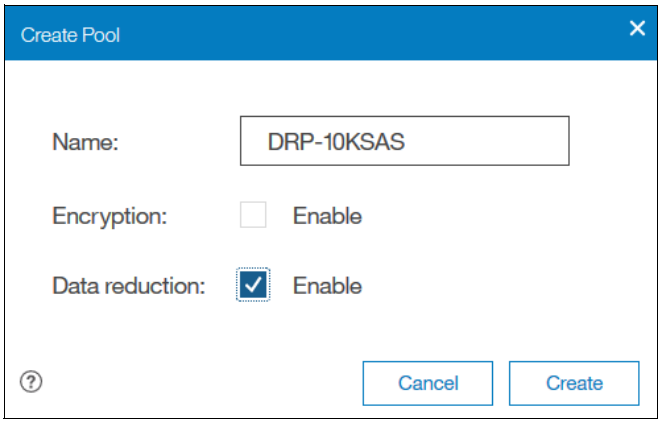

Figure 3-3 Create Pool wizard

2. Click **Create** to create the new Data Reduction Pool.

Figure 3-4 shows the command for creating the Data Reduction Pool.

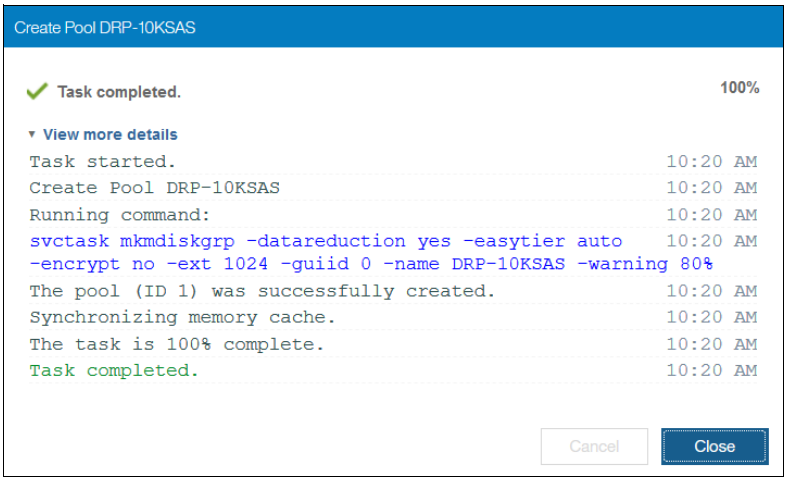

Figure 3-4 CLI command runs

We see that the extent size currently defaults to 1.00 GiB. If a different extent size is required, either the GUI or the CLI can be used to create the storage pool.

Figure 3-5 shows the storage pools view with the newly created Data Reduction Pool. Currently no storage is allocated to the pool.

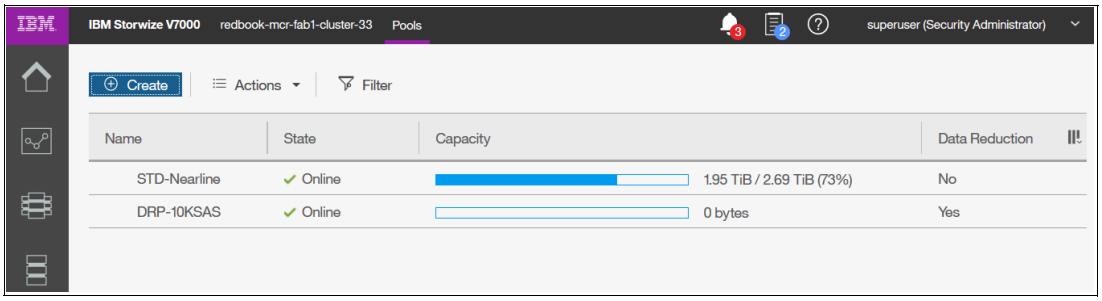

Figure 3-5 Storage Pools view including a Data Reduction Pool

3. Figure 3-6 shows how to initiate the Assign Storage wizard. You either right-click the storage pool or click **Actions** →**Add Storage**.

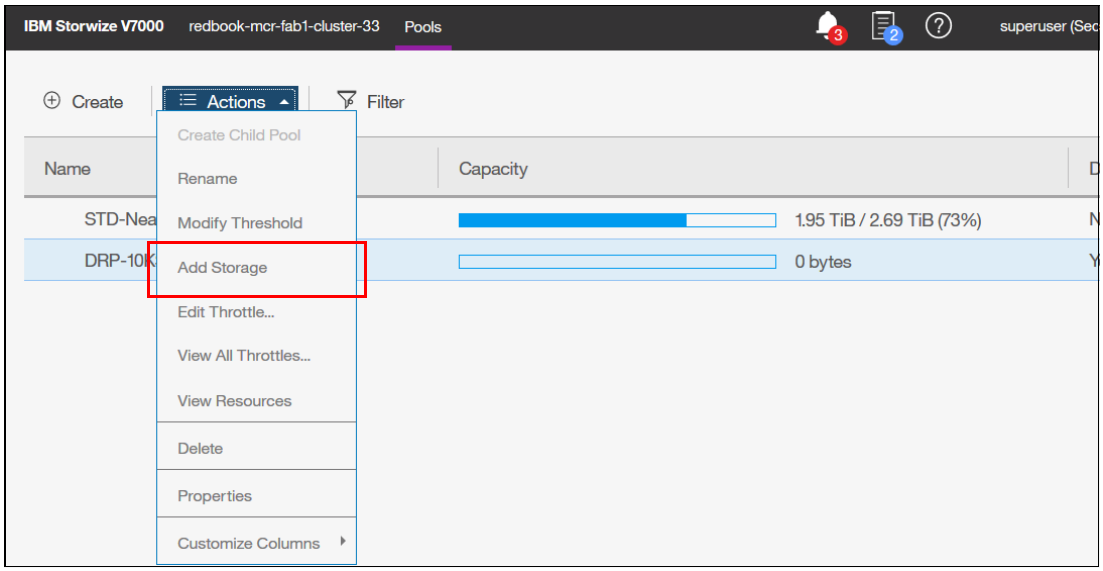

Figure 3-6 Initiate the Assign Storage wizard

4. Figure 3-7 shows the Assign Storage to Pool window. In order to bypass defaults and to prevent volume format from initiating (only applicable for thick provisioned volumes), click **Internal Custom** →**Assign** in the **Advanced** section.

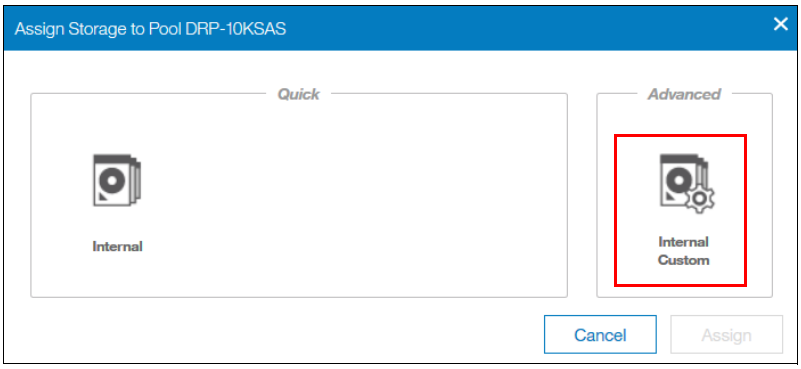

Figure 3-7 Assign storage

5. Figure 3-8 shows the Drive Assignment window. Select **Drive Class**, the number of **Drives** that you want to assign, and the preferred **RAID** type. IBM suggests Distributed RAID-6 (DRAID6) with at least one spare drive for redundancy and better performance.

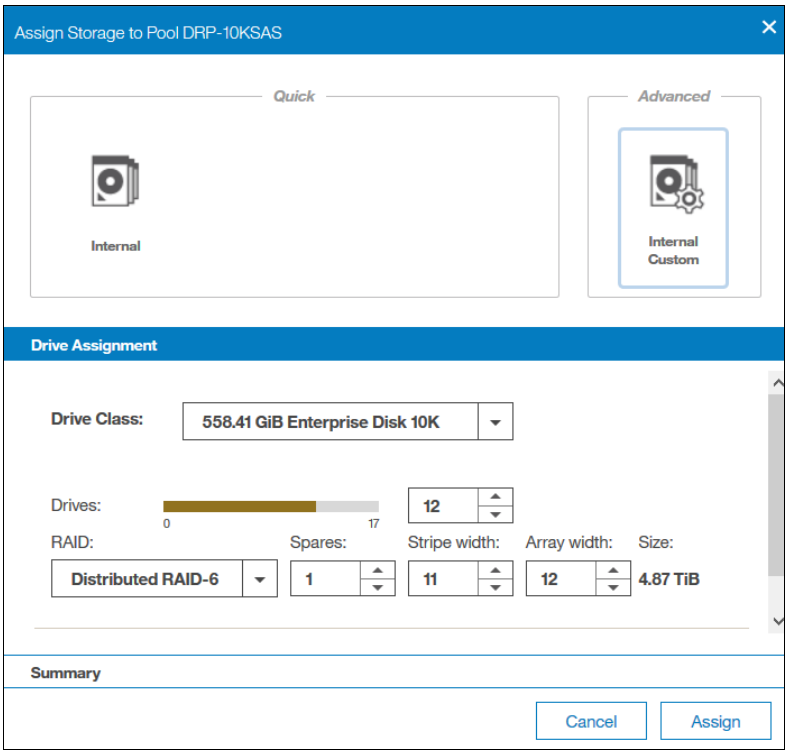

Figure 3-8 Drive assignment

6. Click **Assign** to create the new MDisk in the Data Reduction Pool.

   Figure 3-9 shows how the CLI runs the drive assignment command.

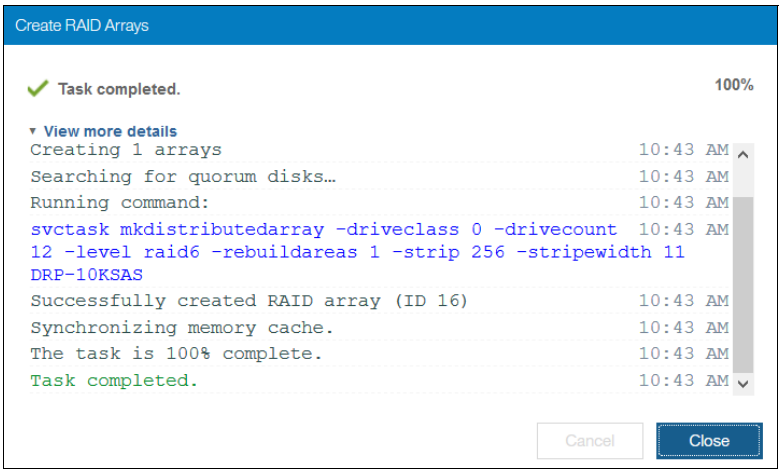

Figure 3-9 CLI runs drive assignment

Figure 3-10 shows the Pools window after adding drives to the data reduction storage pool.

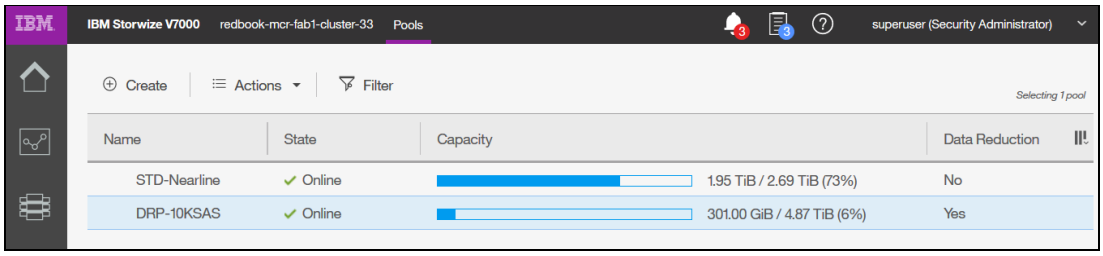

Figure 3-10 Capacity added

Data Reduction Pools require space for internal purposes, such as garbage collection, which is used for cleaning up data extents. Cleaning up extents frees space and improves performance. In Figure 3-10, 6% is allocated from the Data Reduction Pool immediately after assigning storage and before any volumes are created within the pool.

Figure 3-11 shows the **Pools** →**MDisks by Pools** window, from which storage pools can also be created and managed. This view enables you to view MDisks and properties for them, including which drives are in the MDisks.

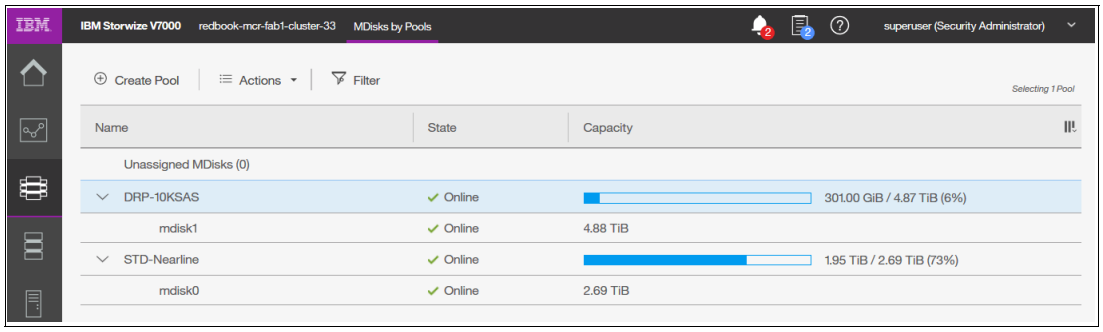

Figure 3-11 MDisks by Pools

Figure 3-12 shows properties for the Data Reduction Pool. It shows that data reduction is enabled and the default extent size is 1.00 GiB.

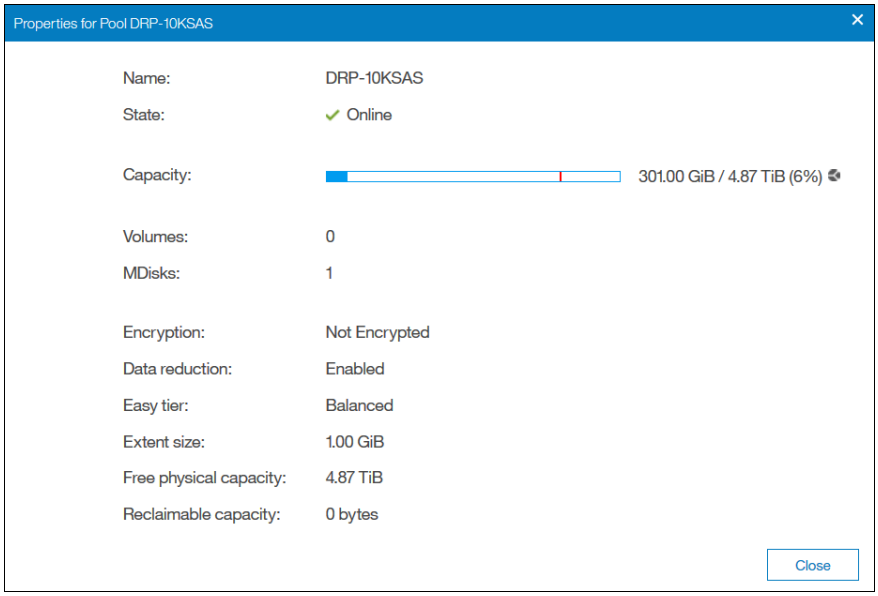

Figure 3-12 Data Reduction Pool properties

Extent size

A larger extent size increases the total amount of storage that the system can manage. A smaller extent size provides more fine-grained control of storage allocation. Extent size does not affect performance.

When choosing the extent size, consult the extent table in the V8.1.3.x Configuration Limits and Restrictions for IBM System Storage SAN Volume Controller page, to ensure that the extent size chosen is big enough for the expected amount of MDisk capacity in the pool:

http://www.ibm.com/support/docview.wss?uid=ssg1S1012364

For example, a DRP with an extent size of 2048 MB can have a maximum of 256 TB of used capacity within a single I/O group.

The minimum extent size for a Data Reduction Pool with IBM Spectrum Virtualize code V8.1.3 and later is 1024 MB. This is enforced in the GUI and by the CLI.

Note: For fully allocated volumes or migration between standard pools, volume migration between storage pools can happen only if storage pools have the same extent size.

You cannot migrate a thin-provisioned volume of any kind into or out of a DRP regardless of extent size.

Capacity threshold

Data Reduction Pools are more sensitive to over-subscription than standard storage pools because they use more advanced features. The default warning threshold is 80%, just like the standard pool. This limit represents the preferred practice and leaves space for “housekeeping purposes”. A Data Reduction Pool might experience performance degradation if it is filled with more than 85% of valid data (garbage is additional to this).

3.3 Creating volumes in Data Reduction Pools

The main purpose of Data Reduction Pools is to be a fast and efficient container for volumes with data reduction capability. Data Reduction Pools can also contain standard thick provisioned volumes.

Figure 3-13 shows the **Volumes** →**Volumes** view before adding volumes into Data Reduction Pools.

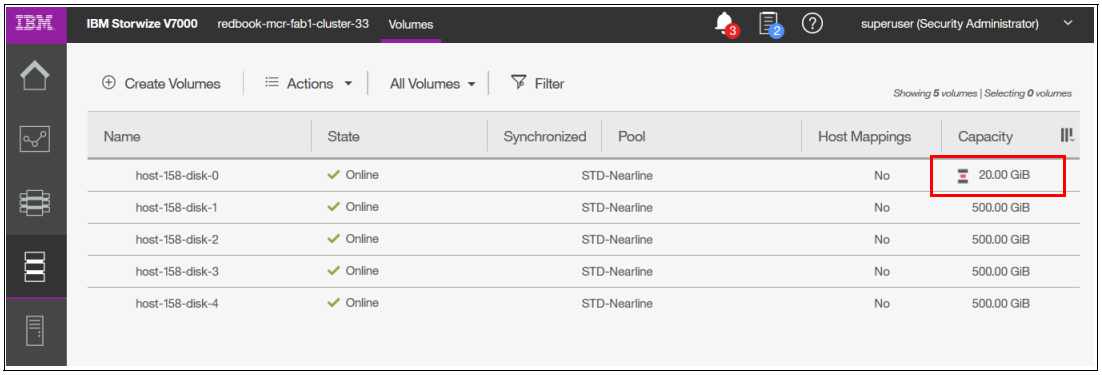

Figure 3-13 Volumes view

So for a given I/O group, compressed volumes in a standard storage pool cannot coexist with deduplicated volumes in Data Reduction Pools. Hence the compressed volume in the standard pool must be moved into the Data Reduction Pool, or converted into a non-compressed volume before we can create deduplicated volumes in our newly created Data Reduction Pool.

The reason for this limitation is a restriction in memory and compression adapters where two technology types cannot share the same hardware.

Reminder: Compressed volumes in a standard storage pool cannot coexist with deduplicated volumes in Data Reduction Pools

Once the standard storage pool has no compressed volumes you will be able to create compressed/deduplicated volumes in the Data Reduction Pool.

Figure 3-14 on page 29 shows the Create Volume wizard. Under the Capacity Savings menu, you can select **None**, **Thin Provisioned**, or **Compressed**. If you select **Compressed**, the **Deduplicated** check box is available and can then also be selected.

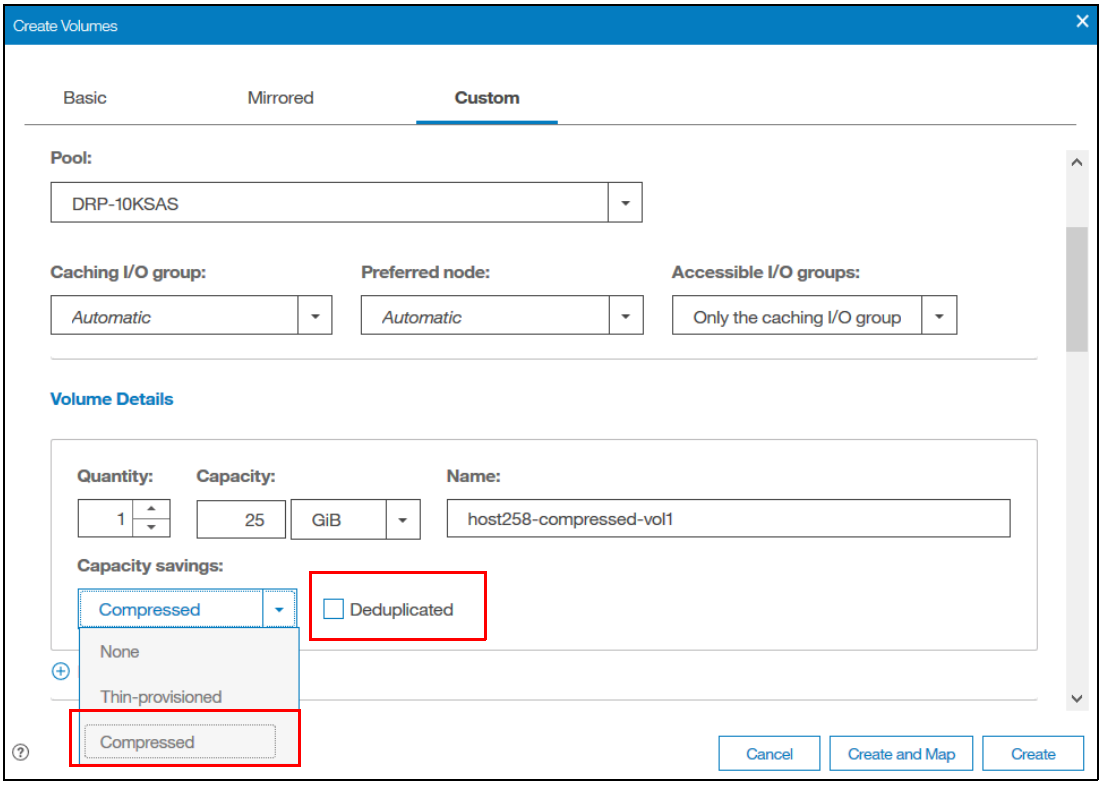

Figure 3-14 Create compressed volume

Figure 3-15 shows the Volumes window. Real capacity and Used capacity have been selected to be displayed, but show `Not Applicable` for thin provisioned and compressed volumes. Real capacity, Used capacity, and Compression savings show `Not Applicable` for volumes with capacity savings. Only fully allocated volumes display those parameters.

Note: When using and reviewing volumes in Data Reduction Pools, be aware that there is no volume copy-level reporting on used capacity. Data reduction savings can be viewed only on a storage pool basis.

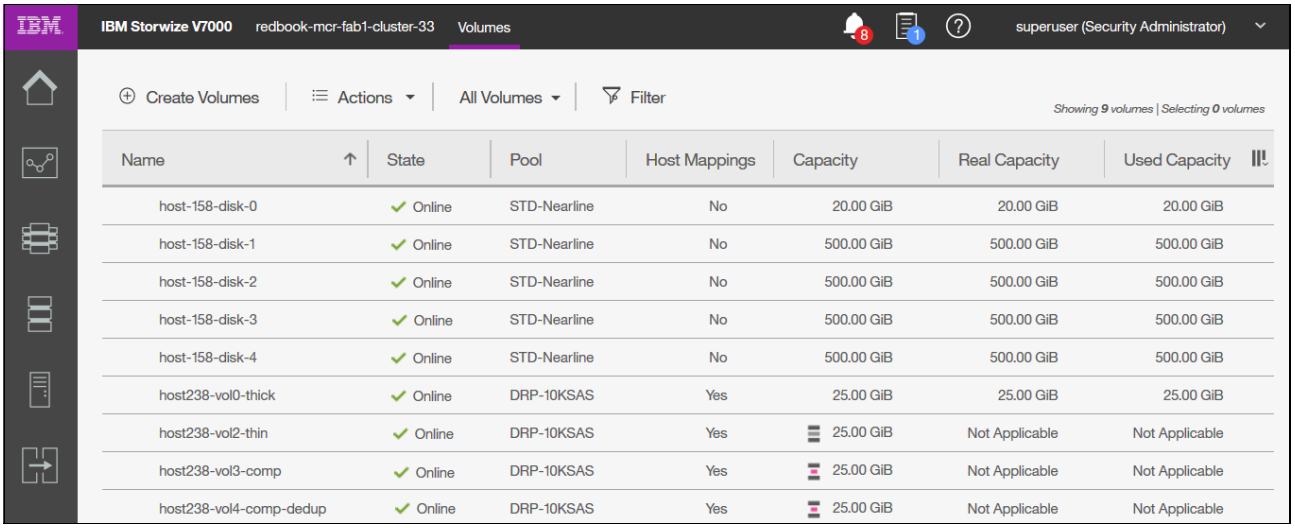

Figure 3-15 Volumes window

Easy Tier in Data Reduction Pools

IBM Easy Tier is a function that automatically and non-disruptively moves frequently accessed data from slower drives to faster drives and vice-versa, placing “hot” and “cold” data in a faster or slower tier of storage depending on data activity.

Easy Tier can be managed at the storage pool level ,and it can be enabled and disabled at the volume level for traditional volumes. However, due to changed volume allocation in Data Reduction Pools, Easy Tier cannot be disabled on volumes in Data Reduction Pools.

3.4 Volume migration

In the following section, we give an example of migrating volumes from a standard storage pool to a Data Reduction Pool. Reasons for such a migration include compressing a non-compressed volume, or compressing and deduplicating a non-compressed volume. Data Reduction Pools are more effective and better performing for data reduction purposes. If the goal is to reduce and improve performance for compressed data, migrating data away from standard storage pools is a common task for storage administrators.

Migrating a volume from a standard pool to a DRP

We will now demonstrate how to migrate a volume from a standard pool into a Data Reduction Pool.

Figure 3-16 shows the Volumes window.

Name	State	Pool	Host Mappings	Capacity	Capacity Savings
host-158-disk-0	✓ Online	STD-Nearline	No	20.00 GiB	Compressed
host-158-disk-1	✓ Online	STD-Nearline	No	20.00 GiB	None

Figure 3-16 Volumes window

Figure 3-17 shows the Add Volume Copy window. To migrate a volume, complete the following steps:

1. The default for the original Copy 0 is the pool where the volume is currently located. Select the pool for Copy 1, which is **DRP-10KSAS**.
2. Under Capacity Savings, select **Compressed**. The Deduplicated check box is not available because, as mentioned in 3.3, “Creating volumes in Data Reduction Pools” on page 28, compressed volumes in standard pools cannot coexist with deduplicated volumes in Data Reduction Pools.

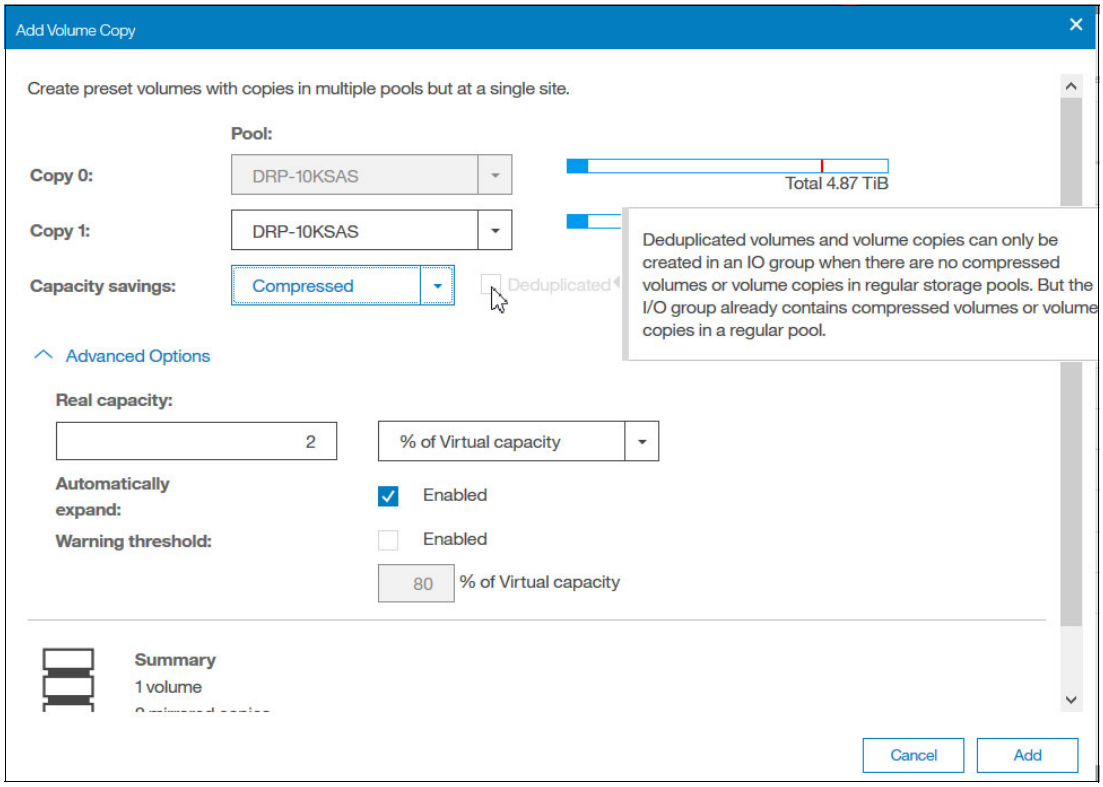

Figure 3-17 Add Volume Copy

3. Click Add to create the compressed volume copy.

   Figure 3-18 shows the resulting volumes window. We now have two volume copies, and one is compressed and in a Data Reduction Pool. The volume copy creation might not have finished creation, so its progress can be viewed in the Show Running Tasks window.

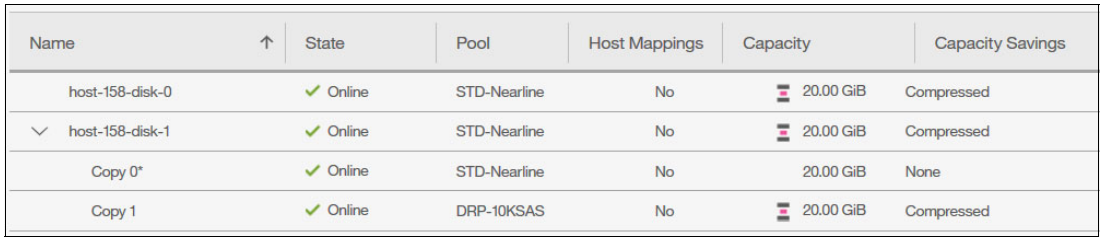

Figure 3-18 Volumes window

4. When the volume copy creation finishes, the original volume copy can be deleted by right-clicking Copy 0 and selecting **Delete**, as shown in Figure 3-19.

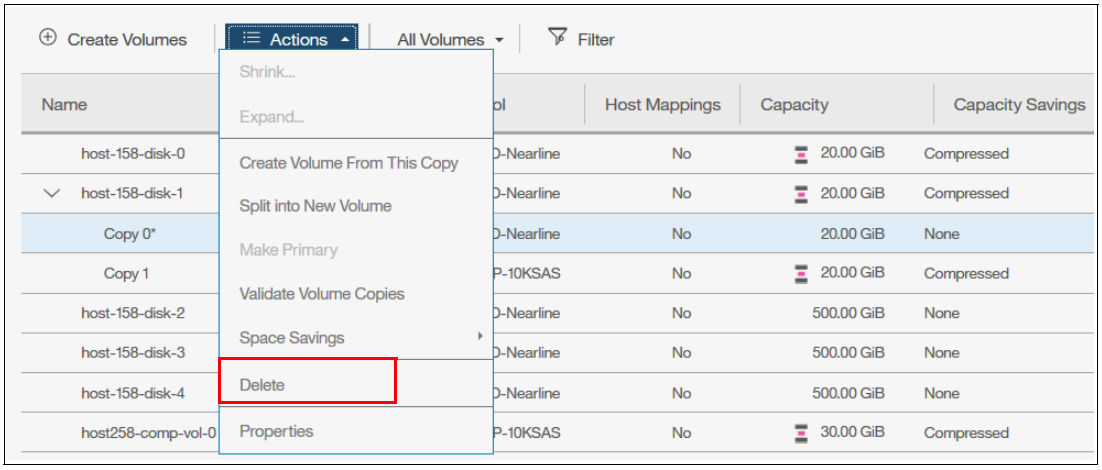

Figure 3-19 Delete volume copy

When this action has completed, the volume has now been migrated to a compressed volume in the Data Reduction Pool. The resulting Volumes window is shown in Figure 3-20.

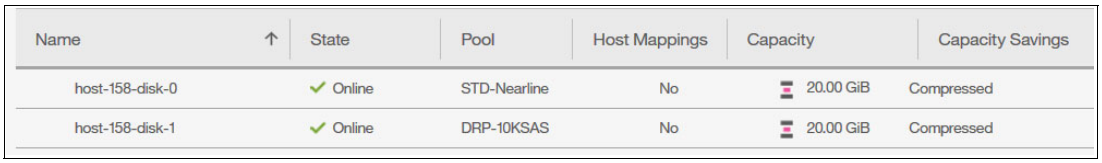

Figure 3-20 Resulting Volumes window

Mirror sync rate

The **`mirror sync rate`** parameter controls the data copied per second in volume copy relationships. This setting affects the initial rate of formatting and the rate at which the volume copies resynchronize after loss of synchronization.

Table 3-1 Relationship between the sync rate value and the data copied per second

User-specified sync rate attribute value	Data copied/sec
1 - 10	128 KB
11 - 20	256 KB
21 - 30	512 KB
31 - 40	1 MB
41 - 50	2 MB
51 - 60	4 MB
61 - 70	8 MB
71 - 80	16 MB
81 - 90	32 MB
91 - 100	64 MB

The default mirror sync rate value is 50. For creating mirror copies in larger volumes, it might be necessary to increase the mirror sync rate for the mirror copy to complete to a synchronized state within a reasonable time.

To review or change the mirror sync rate,complete the following steps:

1. Go to the menu **Volumes** →**Volumes**.
2. Select the volume to be modified and either right-click or click **Actions**, followed by **Modify Mirror Sync Rate**, as shown in Figure 3-21.

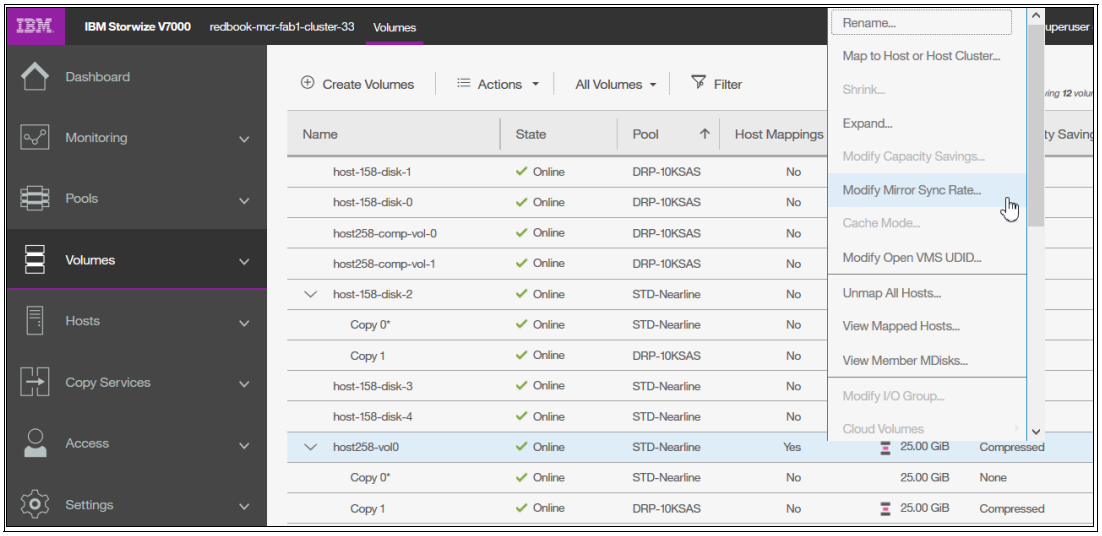

Figure 3-21 Modify Mirror Sync Rate

3. The default mirror sync rate value is 50. Change the value to 100 and click **Modify**, as shown in Figure 3-22.

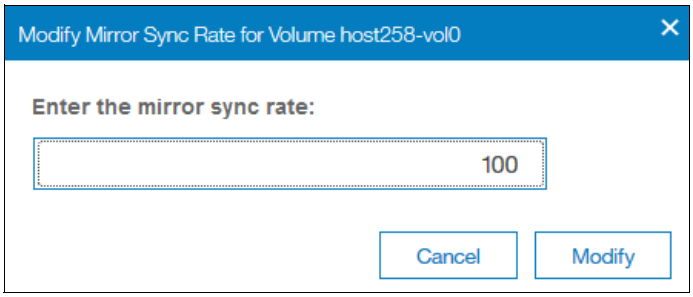

Figure 3-22 Select the new mirror sync rate value

The CLI runs, as shown in Figure 3-23.

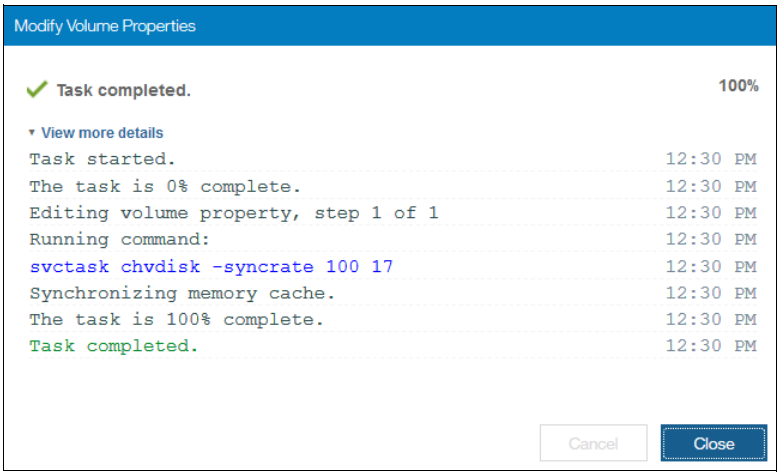

Figure 3-23 CLI runs

Figure 3-24 shows Running Tasks of a newly created volume mirror relationship where a 500 GB standard volume creates a mirror into a compressed volume. The volume uses the default mirror sync rate of 50, which corresponds to 2 MBps. By viewing running tasks, you can see that the remaining copy time to completion is more than 70 hours.

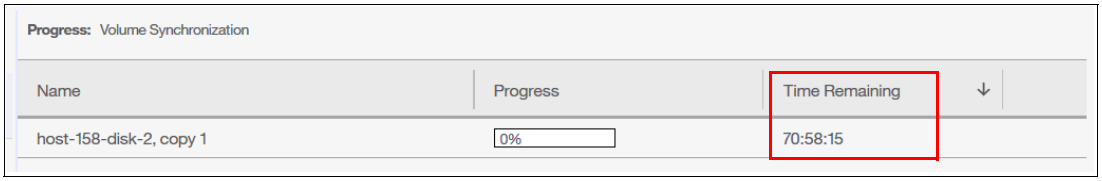

Figure 3-24 Running tasks remaining time using default mirror sync rate

If you change the mirror sync rate value to the maximum, the mirror will complete much quicker. Figure 3-25 shows that with a mirror sync rate value of 100, the Time Remaining is now only two hours, which corresponds to 64 MBps.

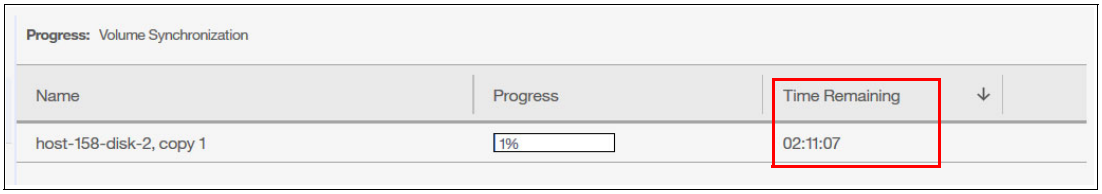

Figure 3-25 Running tasks remaining time using mirror sync rate value 100

Considerations for modifying mirror sync rate

Obviously, use a higher mirror sync rate for the volume copy if you want volume mirror to go quicker. The fastest speed is 64 MBps. Volume mirror workload is sequential, and it might not cause a performance problem if a higher rate is used. However, if many mirror copies are active doing an initial sync at the same time, or if resync is needed due to site or hardware problems, it might have a performance impact if all volumes are set to the maximum mirror sync rate.

The default copy synchronization rate is set to 2 MBps. in order to avoid overloading the managed disks or backend storage system. Use the IBM Spectrum Virtualize built-in performance monitor, or IBM Spectrum Control, to monitor MDisk latency while building mirror relationships. This practice can help you determine if a higher mirror sync rate can be chosen, or if more copies can be active simultaneously.

One approach to create mirror copies into compressed copies can be to use the maximum mirror sync rate while building the mirror copies, and then reduce to the default when synchronization is obtained.

For more information about the **addvdiskcopy** command and mirror sync rate, see the following website:

https://ibm.co/2ztGdUu

Migrating a compressed volume to a compressed volume in a DRP

Consider a RACE compressed volume that is in a standard storage pool, and you want to migrate this to a Data Reduction Pool. The way to migrate such a volume is first to create a compressed mirror volume copy in the Data Reduction Pool, and then delete the original volume copy following the same directions as shown in “Migrating a volume from a standard pool to a DRP” on page 30.

4

Estimator and sizing tools

This chapter provides information about specific tools used for sizing environments for Data Reduction Pools and deduplication.

This chapter describes the tools and how to use them as it pertains to Data Reduction Pools and deduplication.

4.1 Evaluate compression savings using comprestimator

Comprestimator is a comprehensive tool that is used to monitor the performance of clusters, and to measure the level of compression and whether it is a good candidate for compression.

Before using IBM Data Reduction Pools (DRP), it is important for storage administrators to analyze individual volumes or all volumes that are being considered for potential compression savings. This helps you determine if the workload that you are analyzing is a good candidate for DRP.

Note: Data Reduction Pools have metadata processing that takes place in the background. Consider this when planning for sizing, although the resource use is minimal.

IBM Comprestimator is an integrated GUI and CLI host-based utility that estimates the space savings achieved when using compressed volumes for block devices. This utility provides a quick and easy view that shows the benefits of using compression. The utility performs read-only operations and therefore has no effect on the data that is being stored on the device.

If the compression savings prove to be beneficial in your environment, volume mirroring can be used to convert volumes to compressed volumes, then add them to the Data Reduction Pools.

When using the CLI, use the **analyzevdisk** command shown in Example 4-1 to run volume analysis against a single volume..

Example 4-1 The analyzevdisk command

```
IBM_Storwize:redbook-mcr-fab1-cluster-33:superuser>svctask analyzevdisk -h

analyzevdisk

Syntax

>>- analyzevdisk -- --+----------+-- --+- vdisk_id ---+--------><
                      '- -cancel-'     '- vdisk_name -'




For more details type 'help analyzevdisk'.

IBM_Storwize:redbook-mcr-fab1-cluster-33:superuser>svctask analyzevdisk vol1
IBM_Storwize:redbook-mcr-fab1-cluster-33:superuser>
```

When using the CLI use the **analyzevdiskbysystem** command shown in Example 4-2 to run volume analysis against the entire system..

Example 4-2 The analyzevdiskbysystem command

```
M_Storwize:redbook-mcr-fab1-cluster-33:superuser>svctask analyzevdiskbysystem -h

analyzevdiskbysystem

Syntax
```

```
>>- analyzevdiskbysystem -- --+----------+-- ------------------><
                              '- -cancel-'



For more details type 'help analyzevdiskbysystem'.

IBM_Storwize:redbook-mcr-fab1-cluster-33:superuser>svctask analyzevdiskbysystem
vol1
IBM_Storwize:redbook-mcr-fab1-cluster-33:superuser>
```

Note: The **analyzevdisk** and **analyzevdiskbysystem** commands return back to the command prompt.

If you want to see the results of the volumes that you are analyzing, run the commands shown in Example 4-3.

Example 4-3 The lsvdiskanalysis and lsvdiskanalysis progress commands

```
IBM_Storwize:redbook-mcr-fab1-cluster-33:superuser>lsvdiskanalysisprogress
vdisk_count pending_analysis estimated_completion_time
12          12               180622091300
IBM_Storwize:redbook-mcr-fab1-cluster-33:superuser>
```

As stated previously, the utility can be used from the IBM Spectrum Virtualize V8.1 GUI. Figure 4-1 on page 40 shows how to start a complete system analysis on compression estimates. Go to **Volumes** →**Actions** →**Space Savings** →**Estimate Compression Savings.**

A window prompt displays, showing the estimated time of completing the estimate on compression savings.

This process can be monitored by adding the volumes that show compression savings in the volume table, as shown in Figure 4-1.

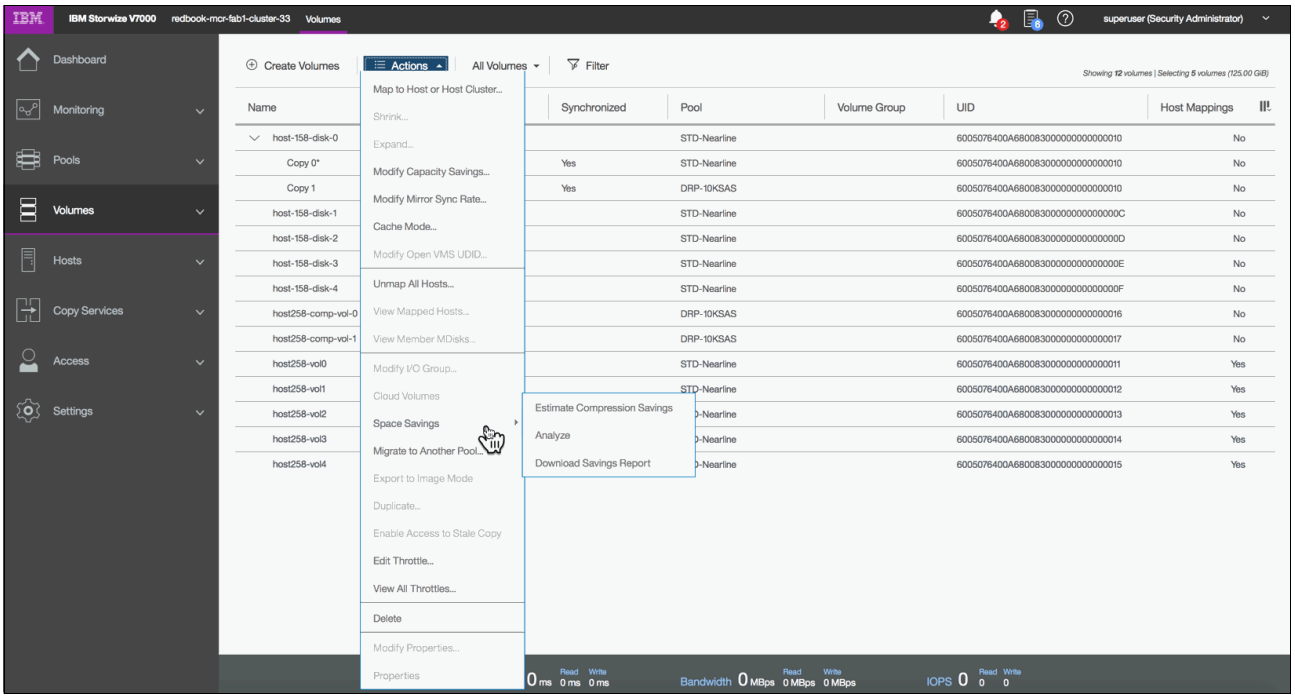

Figure 4-1 Estimate compression savings

If you are using a version of IBM Spectrum Virtualize that is before V7.6, or if you would like to estimate the compression savings of a different device, the IBM Comprestimator utility can be installed on a host that is connected to the device that needs to be analyzed. More information and the latest version of this utility can be found at the following website:

`https://www14.software.ibm.com/support/customercare/sas/f/comprestimator/home.html`

There are a few preferred practices for using Comprestimator:

- Run the Comprestimator utility before implementing an IBM Spectrum Virtualize solution and before implementing the IBM Data Reduction Pools technology.
- Download the latest version of the Comprestimator utility if you are not using one that is included in your IBM Spectrum Virtualize solution.
- Use Comprestimator to analyze volumes that contain as much active data as possible, rather than volumes that are nearly empty or newly created. This will ensure more accuracy when sizing your environment for compression and Data Reduction Pools.

Note: Comprestimator can run for a long period (a few hours) when it is scanning a relatively empty device. The utility randomly selects and reads 256 KB samples from the device. If the sample is empty (that is, full of null values), it is skipped. A minimum number of samples with actual data are required to provide an accurate estimation. When a device is mostly empty, many random samples are empty. As a result, the utility runs for a longer time as it tries to gather enough non-empty samples that are required for an accurate estimate. If the number of empty samples is over 95%, the scan is stopped.

4.2 Evaluating workload using Data Reduction Estimator Tool

Data Reduction Estimator Tool (DRET) is a CLI-operated execution file that runs from the host to analyze a given block device. It provides a report of what it would expect the deduplication savings to be from data written to the disk. There are no additional adjustments or requirements to run this tool on your IBM Spectrum Virtualize solution.

DRET displays its accuracy best with sequential workloads, and with volumes that contain the most active data in a storage environment. DRET is accurate at analyzing a range of workloads to identify the range at which the capacity of a volume can be used. Furthermore, it is able to complete these tasks within a reasonable amount of time, as shown in Table 4-1.

Table 4-1 Estimated time per block size

Volume Size	Time Taken (minutes)
15 GB	1
35 GB	1
100 GB	41
1 TB	32

When using DRET to analyze a block device used by a file system, all underlying data in the device is analyzed, regardless of whether this data belongs to files that were already deleted from the file system. For example, you can fill a 100 GB file system and make it 100% used, then delete all the files in the file system making it 0% used. When scanning the block device used for storing the file system in this example, DRET accesses the data that belongs to the files that are already deleted.

Important: The preferred method of using DRET is to analyze volumes that contain as much active data as possible, rather than volumes that are mostly empty of data. This increases the accuracy level and reduces the risk of analyzing old data that is already deleted, but may still have traces on the device.

More information and the latest version of this utility can be found at the following website:

http://www14.software.ibm.com/webapp/set2/sas/f/dretool/home.html

4.3 Evaluating workload using Disk Magic

Disk Magic for IBM is a performance analysis and hardware configuration planning tool for IBM storage disk subsystems that runs on Microsoft Windows. It is licensed by IBM from IntelliMagic for exclusive use by IBMers and IBM Business Partners in the selling of new or upgraded IBM disk storage solutions.

Proper initial sizing greatly helps to avoid future sizing problems. Disk Magic is one such tool that is used for sizing and modeling storage subsystems for various open systems environments and various IBM platforms.

It provides accurate performance and capacity analysis and planning for IBM Spectrum Virtualize products, other IBM storage solutions, and other vendors' storage subsystems. Disk Magic provides in-depth environment analysis, and is an excellent tool to estimate the performance of a system that is running Data Reduction Pools (DRP).

For IBM Business Partners that would like more information about Disk Magic, and to obtain the latest version, go to the following website:

http://www.ibm.com/partnerworld/wps/servlet/ContentHandler/SSPQ048068H83479I86

5

Best practices and guidelines

This chapter describes the preferred practices and guidelines for using IBM Spectrum Virtualize Data Reduction Pool (DRP) technology. The suggestions are based on the newly revised architecture of the product.

This chapter provides use cases and preferred practices for the DRP technology and its performance and efficiency gains.

This chapter is intended for experienced IBM SAN Volume Controller (SVC) and IBM Storwize administrators and consultants. This book requires advanced knowledge of the SVC and Storwize products.

Recommended software level: The minimum recommended software level is V8.1.3.2 which is available from IBM Fix Central and contains a number of critical fixes.

5.1 RACE and DRP overview

Flash and solid-state drive (SSD) technology are constantly improving performance, and this technology provides low latency for application workloads.

Although Flash technology is cheaper than before, it has also reduced TCO because it requires less cooling and rack space, but it is currently still more expensive than traditional spinning disks. For this reason, storage administrators optimize the amount of data stored on Flash storage to drive the TCO even lower.

IBM Data Reduction Pool technology is developed to optimize the Flash workload to provide cost savings by storing less data but at the same time providing stable and predictable performance.

5.1.1 DRP performance benefits

The SVC and Storwize first RACE implementation was software-only. Later, dedicated hardware compression was introduced. The RACE implementation was created at a time when spinning disks were implemented in the majority of storage infrastructures. Today, SSD/Flash technology use is more widespread, and data storage and capacity optimization are one of the key goals in the industry.

Initially, the original RACE implementation was done using a software-only approach on the Storwize platform. As the product matured, we were able to use hardware accelerators and Flash/SSD back-end technology to increase performance.

Data Reduction Pool technology is targeted to leverage the hardware accelerator and Flash technology. The architecture of DRP compression does not dedicate any hardware to compression software, as significantly improved load balancing and Flash optimized I/O paths provide an optimal performance for a compressed workload.

5.1.2 DRP performance improvements

These are some of the major DRP improvements:

- Multithreaded approach improved latency and throughput
- All memory and cores available to all functions, with no dedicated cores or memory
- Up to 3x throughput can be achieved for systems with hardware accelerators
- New compression and deduplication implementation to leverage new hardware

Note: Storwize V5030 does not have a dedicated compression accelerator card, so it uses software compression only.

Data Reduction Pool I/O handling

Data Reduction Pool technology uses a fixed block size (also known as grain size) of 8 KB. This block size was chosen specifically to optimize the workload for Flash storage where small block I/O leverages low latency flash performance. Host I/O is acknowledged by the upper cache layer and then processed in 8 KB chunks.

Figure 5-1 shows DRP compression.

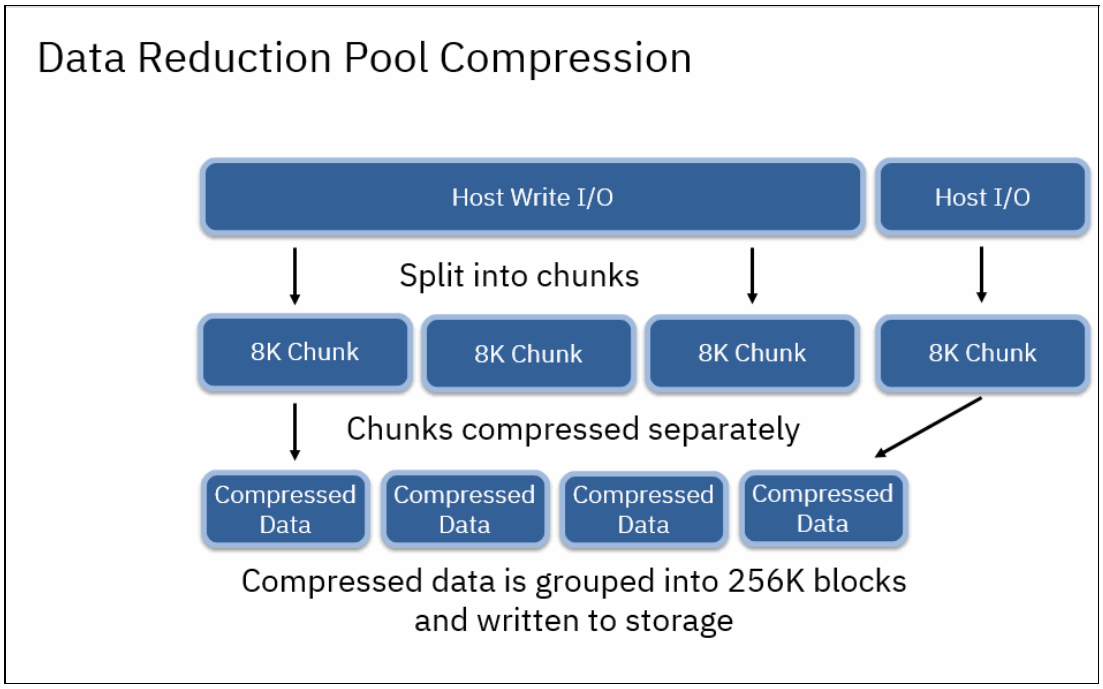

Figure 5-1 DRP compression

DRP technology introduces an implementation of compression that provides a stable performance irrespective of the I/O size and workload pattern, compared to the RACE implementation where certain workload types might have reduced performance. The compression workload provides predictable performance no matter what the workload type is. This consistent performance encourages users to compress workloads on Flash, and it provides efficiency, cost savings, and better performance.

DRP technology is ideal for Flash storage with compression. For example, with a compression pattern of 2:1 or higher, host I/O is “sliced” into 8KB equal chunks. In addition, each 8 KB chunk will be compressed to 4 KB or less, and is the optimal block size for leveraging flash performance.

The new compression implementation can produce up to 4x throughput for compressed workloads, with consistent performance.

DRP deduplication with compression provides the best storage efficiency. This combination deduplicates and then compresses the data, reducing the storage capacity usage. It is advised to use this option when the workload is compressible and it has a high duplicate ratio identified by the DRET tool. For more information, see 4.2, “Evaluating workload using Data Reduction Estimator Tool” on page 41.

DRP Deduplication I/O handling

Host I/O is processed in 8 KB chunks, and then an in-memory calculation is performed for deduplication. Host write I/O for the deduplication volume is acknowledged back to the host by the upper cache layer. Deduplication is performed after acknowledging the write to the host, so the host performance is not impacted by deduplication.

The host read may be a cache hit or a read from the back-end depending upon the workload type.

Figure 5-2 shows DRP deduplication.

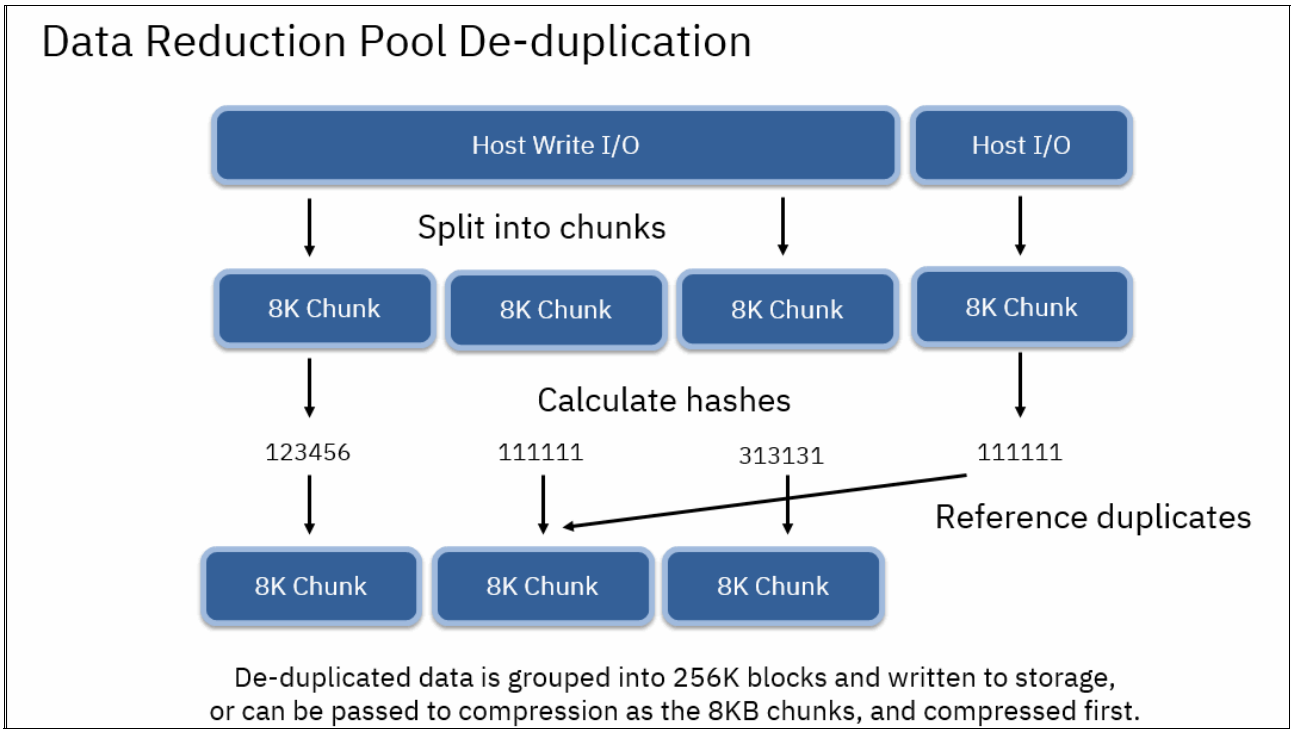

Figure 5-2 DRP deduplication

DRP volume types

As a reminder from previously, DRP technology enables you to create five types of volumes:

- Fully allocated

  This type provides no storage efficiency but the best performance, and is available for migration.

- Thin

  This type provides storage efficiency but no compression or deduplication.

- Thin and Compressed

  This type provides storage efficiency with compression, and this combination provides the best performance numbers.

- Thin and Deduplication

  This type provides storage efficiency but without compression.

- Thin, Compressed, and Deduplication

  This type provides storage efficiency with maximum capacity savings.

With storage efficiency, DRP thin and compressed volumes provide the best performance numbers. This is due to the new compression implementation. This implementation provides better load balancing and consistent performance when compared to the RACE implementation. This feature is also the second best performer to fully allocated volumes, followed by thin, compressed and deduplication volumes when it comes to storage efficiency.

Figure 5-3 shows the types of volumes in the DRP pools.

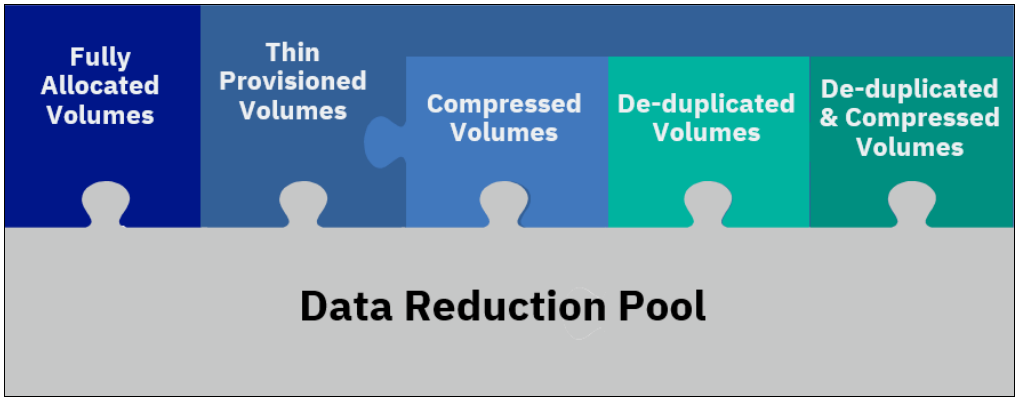

Figure 5-3 Volume types

We will discuss these volume combinations from a performance perspective.

Users should select the volume type that meets their business objectives and leverages good performance from DRP technology. Each of the above volume combinations provides certain benefits for storage efficiency and have different performance characteristics.

DRP technology allows the user to create a fully allocated volume. Fully allocated volumes provide the best performance but do not provide the best storage efficiency. This option is allowed for migrations, and also for applications that require maximum performance with lowest possible latency.

The current suggestion is to use DRP compressed volumes. There are two exceptions to this:

- If the workload has higher than a 30% deduplication ratio, use DRP deduplicated volumes instead.
- If the hardware platform is Storwize V5030, size for compressibility and performance. For more details, see Chapter 4, “Estimator and sizing tools” on page 37.

Configuring your system for best DRP performance

SVC/Storwize storage pools are designed to separate workload or to mitigate against hardware failures using a failure domain. Standard pools can get similar performance numbers whether they are using single or multiple pools. However, DRP performance gets a boost by increasing the numbers of pools in the system. Storage infrastructure should be designed in such a way that the workload is separated by workload type in different failure domains to achieve maximum DRP performance.

SVC/Storwize system are multi-core machines where the system load balances the workload utilizing all of the available resources to provide better performance. With DRP technology, the best performance numbers can be achieved using four Data Reduction Pools. Four DRP storage pools enables the optimal use of node canister hardware resources, such as CPU cores.

Using standard pools (non-DRP), the storage administrator has to cater for the performance of each volume at an application level. However, using DRP technology, storage administration is simplified because the storage administrator only monitors the storage pool for capacity and performance, and the volumes no longer require monitoring for capacity. More DRP pools enable the system to load balance the resource more evenly, which provides good performance.

Size your workload and capacity using the tools described in Chapter 4, “Estimator and sizing tools” on page 37. An application workload with a 2:1 or higher compression ratio leverages the performance benefits from DRP compression, but to enhance the performance further using up to four pools provides the best performance numbers.

5.1.3 Garbage collection

DRP technology is built to leverage SSD/Flash technology. DRP technology is based on log structured array (LSA), where nothing is ever overwritten. Storage is allocated on demand and if a host or application rewrites a region of a disk, the new write is stored in a new location on the back end. DRP technology invalidates the old data on the back-end disk to be garbage collected. For more information, see Chapter 1, “Architecture” on page 1.

Garbage collection keeps track of extent usage in pools, and extents with the most unused space are garbage collected. Under normal circumstances, garbage collection works in a trickle mode, clearing up space and optimizing the back-end storage allocation. In a case where the pool is over provisioned and close to getting full, 85 - 100% of physical capacity, garbage collection has to work harder and it will choose the least-used storage pool extents to clear space.

The entire environment should be sized for performance and capacity before implementing DRP technology. DRP compression enables the user to compress the data, and it lowers storage costs. This could also lead to the user over-provisioning, and in some cases potentially an out-of-space condition. If a system runs out of space on the back-end disks, the DRP pool goes offline.

A DRP storage pool over-provisioning condition could create a performance issue when the physical space in the pool is between 85 - 100%. In this scenario, host I/O and garbage collection can lead to disk I/O contention, which could potentially overload the back end or running out of space completely.

Another use case where garbage collection may have to work harder is when a user deletes a large amount of data from the host, or one or more volumes are deleted. If this happens, the reclaimable capacity in the DRP pool goes up, and garbage collection needs to run and reclaim space. If the storage system is not back-end disk limited, garbage collection should not impact the performance of the system.

Users can check and monitor the amount of data that needs garbage collection. Monitor a Data Reduction Pool using the **lsmdiskgrp** CLI command or the GUI. Look for the new DRP attribute called **reclaimable_capacity**. This can be viewed using either the CLI or the GUI view.

When using the CLI, issue the following command:

```
svcinfo lsmdiskgrp <pool_id>
```

In this case, <*pool_id*> is the numeric id of the Data Reduction Pool. Also, view the amount of data that needs garbage collection in **reclaimable_capacity** output.

Example 5-1 shows the output from the **svcinfo lsmdiskgrp 0** command, with the reclaimable capacity in bold.

Example 5-1 svcinfo lsmdiskgrp 0

```
svcinfo lsmdiskgrp 0
id 0
name Group0
```

```
status online
mdisk_count 16
vdisk_count 64
capacity 208.00TB
extent_size 1024
free_capacity 203.81TB
virtual_capacity 5.00TB
used_capacity 3.91TB
real_capacity 3.91TB
overallocation 2
warning 0
easy_tier auto
easy_tier_status balanced
tier tier0_flash
tier_mdisk_count 16
tier_capacity 208.00TB
tier_free_capacity 204.06TB
tier tier1_flash
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
tier tier_enterprise
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
tier tier_nearline
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
compression_active no
compression_virtual_capacity 0.00MB
compression_compressed_capacity 0.00MB
compression_uncompressed_capacity 0.00MB
site_id
site_name
parent_mdisk_grp_id 0
parent_mdisk_grp_name Group0
child_mdisk_grp_count 0
child_mdisk_grp_capacity 0.00MB
type parent
encrypt no
owner_type none
owner_id
owner_name
data_reduction yes
used_capacity_before_reduction 5.00TB
used_capacity_after_reduction 1.83TB
overhead_capacity 2.08TB
deduplication_capacity_saving 0.00MB
reclaimable_capacity 1.04GB
physical_capacity 93.13TB
physical_free_capacity 85.33TB
shared_resources no
```

This can also be achieved using the GUI by selecting **Pool View** →**Select DRP Pool** → **Action** →**Properties** →**View More Details**, and reclaimable capacity is shown in Figure 5-4.

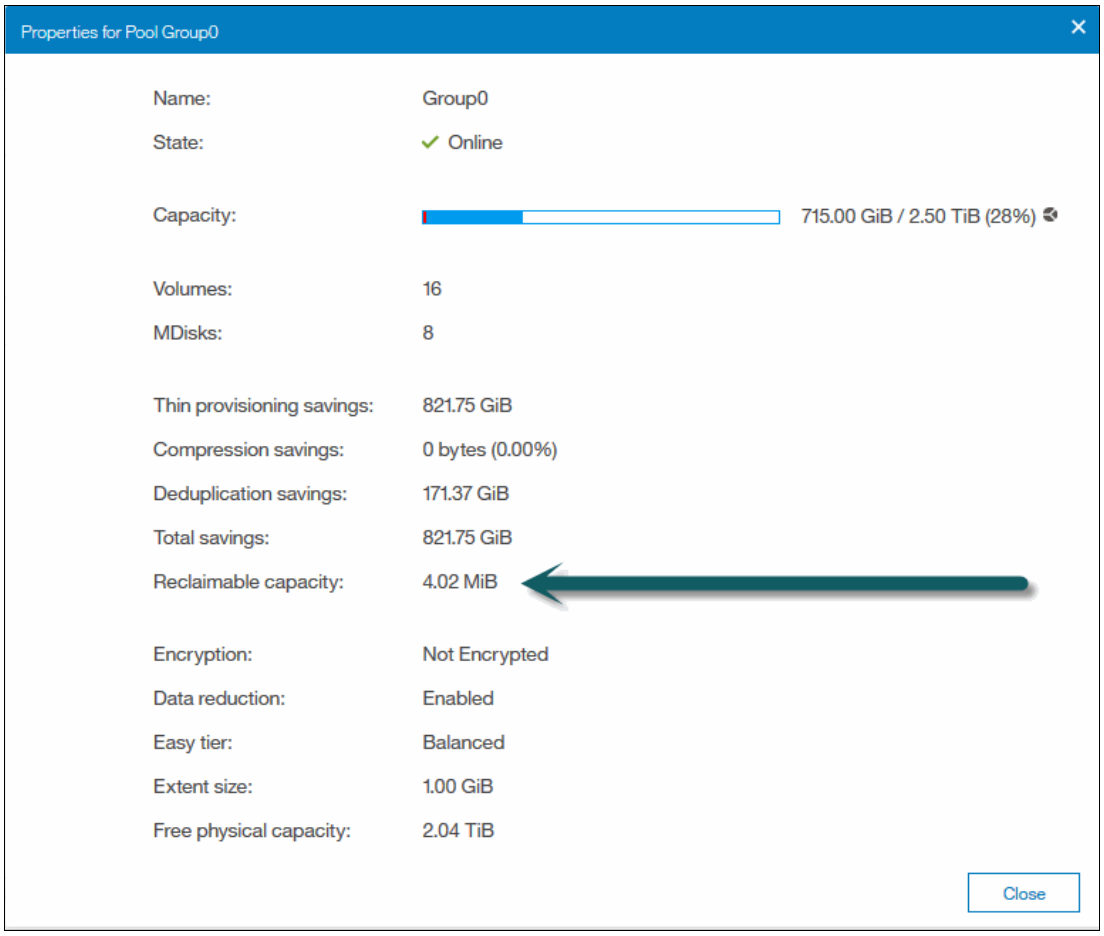

Figure 5-4 Reclaimable capacity

The reclaimable capacity is the amount of data that will be garbage collected. As long as there is more than 15% of the total capacity free, the user should not experience any performance degradation.

5.1.4 Inter-node ports

IBM Spectrum Virtualize SVC/Storwize hardware platforms consist of at least two and up to a maximum of 8 nodes in a clustered system. The performance of the system relies on the timely communication of nodes in the cluster to provide good performance.

The Storwize family uses a dedicated interconnect to communicate to the partner node canister in the same enclosure. In a single I/O group configuration, the Storwize family code only uses the dedicated link unless the link (hardware) is broken, and FC ports are not used for inter-node connectivity. For a cluster with two or more control enclosures, inter-node communication between two control enclosures is done on FC ports as it is with the SVC.

SVC nodes use FC ports for connectivity. Prior to the DRP/Deduplication release, the advice was to use at least two FC ports for inter-node communication.

The two FC port inter-node suggestion remains the same for DRP users unless the host/application using DRP has a high write throughput (above 2 GBps host write workload)

or high write IOPS workload. Specifically, for deduplication users with a high write workload (above 1 GBps host workload) we suggest using four FC ports dedicated to inter-node traffic.

For more information about the original two FC port recommendation, see the following publication:

IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines, SG24-7521

5.1.5 SVC with Storwize family back-end controllers

Because both systems will be running the same code, there is no performance or functionality gain available by using DRP compression or deduplication on both systems together. Therefore, it is suggested to use DRP technology at the SVC layer.

5.1.6 SVC with external back-end controllers

If external back-end controllers are used and virtualized behind SVC, again the advice is to use DRP compression on the front-end SVC layer. Careful planning is required for the back-end controller, which also can compress or deduplicate, and you should consider using a 1:1 compression and deduplication ratio. This is to avoid running out of physical capacity.

5.1.7 Storwize family with external back-end controllers

If Storwize is used to virtualize any other back-end controller behind it, use DRP at the Storwize layer. This combination provides storage efficiency and good, stable performance.

5.2 When to use DRP Compression

DRP pools use a new compression algorithm implementation. The suggestion for optimal performance is to use compression on all platforms and on all workloads (except Storwize V5030, as noted previously).

There is no performance penalty for writing non-compressible data, and for application workloads with a 2:1 or higher compression ratio there is significant capacity saving with no performance overhead. This eliminates a lot of planning work, and simplifies capacity savings for any storage solution.

Storwize V5030 does not have hardware accelerator cards, and it uses only software compression on the six core hardware platform. To use compression on Storwize V5030, use the same RACE guidance for compressibility, nodes CPU usage. See the following publications for guidelines on using compression on Storwize V5030:

- *IBM Real-time Compression in IBM SAN Volume Controller and IBM Storwize V7000*, REDP-4859
- *Implementing the IBM Storwize V5000 Gen2 (including the Storwize V5010, V5020, and V5030) with IBM Spectrum Virtualize V8.1*, SG24-8162

5.2.1 When to use DRP deduplication

Application workloads with more than a 30% deduplication ratio should consider using the DRP deduplication feature. A workload with a 2:1 or higher deduplication ratio will benefit

from capacity savings. In addition, the deduplication feature should be used with compression, which provides the best capacity savings.

To identify the deduplication ratio, use the DRET tool. For more details, see 4.2, “Evaluating workload using Data Reduction Estimator Tool” on page 41.

5.3 Performance monitoring

IBM advises performance monitoring of the system, which enables the system administrator to monitor the production workload and storage solution performance. IBM provides Spectrum Control which is an enterprise class performance monitoring solution for SVC, Storwize and other storage products.

There is another alternative IBM Cloud™ offering known as *Storage Insights*, which can be used for performance monitoring of SVC and Storwize solutions. For more details, contact your IBM sales representative or Business Partner. Alternatively, visit the following link for more information:

https://www.ibm.com/us-en/marketplace/analytics-driven-data-management/resources

5.3.1 Total port data rate by port

Users should monitor the port data rates using the following guidelines:

- Although FC ports are full duplex, setting a warning level near the *single direction line speed* is advised to prevent ports from becoming saturated. Specifically, we suggest a warning level of 1200 MBps for 8 Gb ports, and 2100 MBps for 16 Gb ports.
- Exceeding this limit requires further review. Workload can be balanced or moved to other available resources (ports, nodes, or I/O Groups), or additional resources might be needed to accommodate additional growth.

5.3.2 Total SVC node data rate

The suggested monitoring limit for the front-end data rate is 1800 MBps per node. This number is based on 75% utilization of 8 Gb ports shared for host and backend traffic.

CPU utilization percentage: maximum per SVC node

Users should monitor CPU usage using the following guidelines:

- While a weighted average CPU utilization for the cluster (subsystem) is available, it’s best to monitor utilization per node.
- Exceeding the node average of 70% might be impacting I/O performance for some workloads.
- CPU utilization per core should also be monitored to ensure that no single core exceeds 90%.

Port to local node send response time: maximum per node

Users should monitor port to local node send response time using the following guidelines:

- The port to local node response time is critical to ensuring stable performance of the cluster. Values greater than 0.5 ms impact the overall performance of the cluster to varying degrees.

- Poor port to local node send response time is generally only an issue when the ports utilized for local node traffic are shared with other traffic (host/storage or replication). But even with dedicated ports, elevated port to local node send response time can still be an issue for other conditions, such as port saturation or degraded link quality due to hardware conditions.

Buffer credit depletion: maximum per SVC port

Users should monitor buffer credit depletion using the following guidelines:

- Buffer credit depletion exceeding 20% that is sustained for longer than 20 minutes can negatively impact performance for all workloads that are sharing that port.
- Short duration spikes in credit depletion that occur because of spikes in workload are normal FC flow control events, and are not indicative of a problem.
- 16 Gbps adapters do not support buffer credit counters. A replacement metric called *Port Send Delay ms/op* is available in SVC V7.8.1 and later, and supported in IBM Spectrum Control V5.2.14 and later.

Volume and back-end response times

Each environment can have a unique workload pattern and require a certain desired performance from the given workload. For example, a video streaming storage solution might require a 10 ms response time from volumes and 5 - 10 ms latency is not an issue. Another environment that hosts financial software, for example an OLTP database, might require <1 ms response times. In these sorts of environments, we recommend that clients engage with IBM Professional Services.

6

Troubleshooting

We advise the reader to consult the following publications for common procedures to maintain the IBM Spectrum Virtualize environment, and to ensure that all of the necessary logs are kept should they need to be analyzed by IBM Support:

- *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521
- *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933
- *Implementing the IBM Storwize V7000 with IBM Spectrum Virtualize V8.1*, SG24-7938

6.1 Troubleshooting and fix procedures

The management GUI of IBM Spectrum Virtualize is a browser-based GUI for configuring and managing all aspects of your system. It provides extensive facilities to help troubleshoot and correct problems. This section explains how to effectively use its features to avoid service disruption of your system.

Figure 6-1 shows the menu to start the **Monitoring** menu for **System** information, viewing **Events**, or seeing real-time **Performance** statistics.

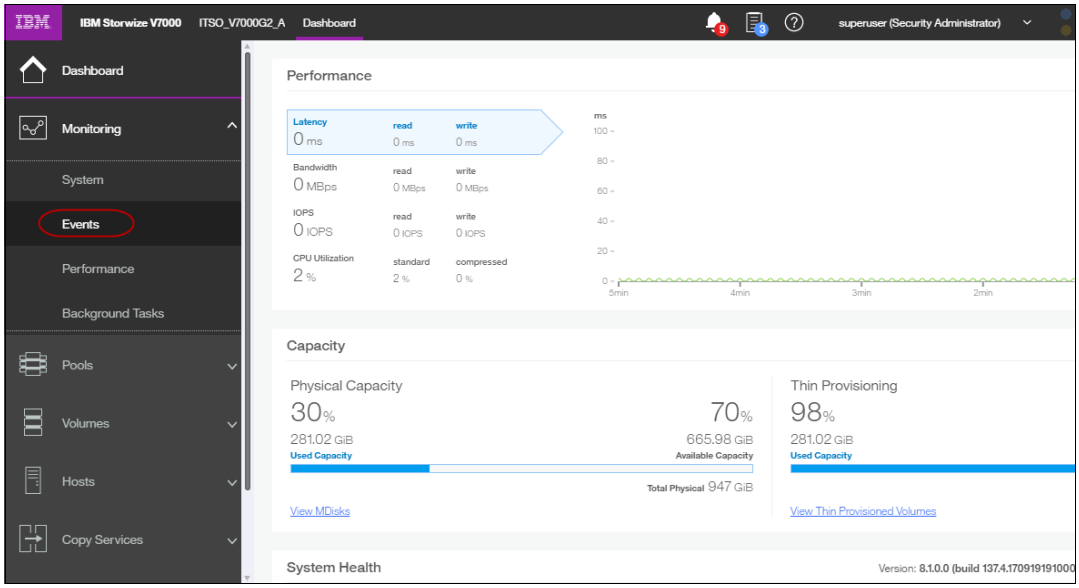

Figure 6-1 Monitoring options

Use the management GUI to manage and service your system. Select **Monitoring** →**Events** to list events that should be addressed, and list maintenance procedures that walk you through the process of correcting problems. Information in the Events window can be filtered in several ways:

- Recommended Actions

  Shows only the alerts that require attention. Alerts are listed in priority order and should be resolved sequentially by using the available fix procedures. For each problem that is selected, you can perform these tasks:

  - Run a fix procedure
  - View the properties

- Unfixed Messages and Alerts

  Displays only the alerts and messages that are not fixed. For each entry that is selected, you can perform these tasks:

  - Run a fix procedure
  - Mark an event as fixed
  - Filter the entries to show them by specific minutes, hours, or dates
  - Reset the date filter
  - View the properties

- Show All

  Displays all event types whether, they are fixed or unfixed. For each entry that is selected, you can perform these tasks:

  – Run a fix procedure
  – Mark an event as fixed
  – Filter the entries to show them by specific minutes, hours, or dates
  – Reset the date filter
  – View the properties

Some events require a certain number of occurrences in 25 hours before they are displayed as `unfixed`. If they do not reach this threshold in 25 hours, they are flagged as `expired`. Monitoring events are below the coalesce threshold, and are usually transient.

Important: The management GUI is the primary tool that is used to *operate* and *service* your system. Real-time *monitoring* should be established by using SNMP traps, email notifications, or syslog messaging in an automatic manner.

6.1.1 Managing event log

Regularly check the status of the system using the management GUI. If you suspect a problem, first use the management GUI to diagnose and resolve the problem.

Use the views that are available in the management GUI to verify the status of the system, the hardware devices, the physical storage, and the available volumes by completing these steps:

1. Click **Monitoring** →**Events** to see all problems that exist on the system (Figure 6-2).

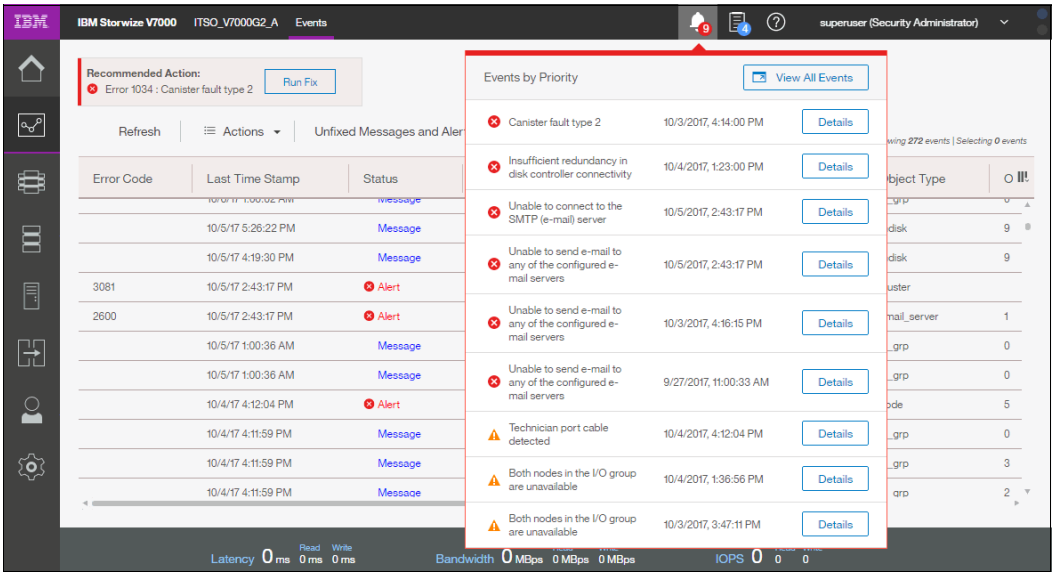

Figure 6-2 Messages in the event log

2. Select **Show All** →**Recommended Actions** to display the most important events to be resolved (Figure 6-3). The Recommended Actions tab shows the highest priority maintenance procedure that must be run. Use the Troubleshooting wizard so that the system can determine the proper order of maintenance procedures.

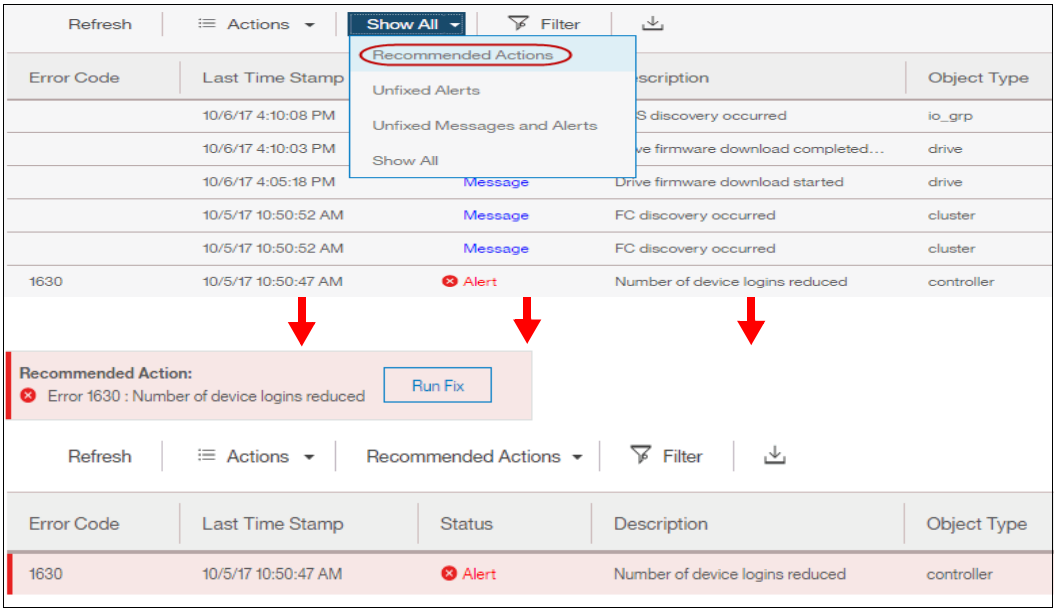

Figure 6-3 Recommended Actions

3. In this example, the *number of device logins reduced* is listed (service error code 1630). Review the physical FC cabling to determine the issue and then click **Run Fix**. At any time and from any GUI window, you can go directly to this menu by clicking the **Status Alerts** icon at the top of the GUI (Figure 6-4).

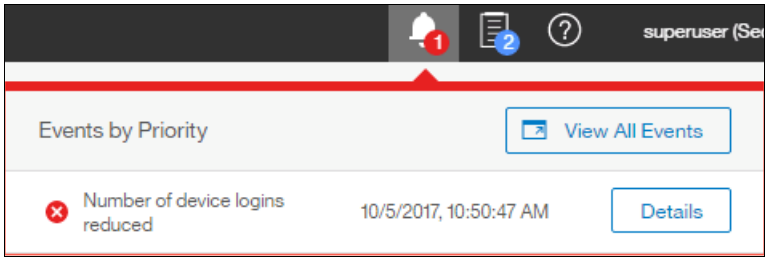

Figure 6-4 Status alerts

6.1.2 Running a fix procedure

If there is an error code for the alert, you should run the fix procedure to assist you in resolving the problem. These fix procedures analyze the system and provide more information about the problem. They suggest actions to take and walk you through the actions that automatically manage the system where necessary, while ensuring availability. Finally, they verify that the problem is resolved.

If an error is reported, always use the fix procedures from the management GUI to resolve the problem. Always use the fix procedures for both software configuration problems and hardware failures. The fix procedures analyze the system to ensure that the required changes will not cause volumes to become inaccessible to the hosts. The fix procedures automatically perform configuration changes that are required to return the system to its optimum state.

The fix procedure displays information that is relevant to the problem, and provides various options to correct the problem. Where possible, the fix procedure runs the commands that are required to reconfigure the system.

Note: After V7.4, you are no longer required to run the fix procedure for a failed drive. Hot plugging of a replacement drive will automatically trigger the validation processes.

The fix procedure also checks that any other existing problem will not result in volume access being lost. For example, if a power supply unit in a node enclosure must be replaced, the fix procedure checks and warns you if the integrated battery in the other power supply unit is not sufficiently charged to protect the system.

Hint: Always use the **Run Fix** button, which resolves the most serious issues first. Often, other alerts are corrected automatically because they were the result of a more serious issue.

The following example demonstrates how to clear the error that is related to a SAN Fabric zoning error to a back-end disk controller:

1. From the GUI menu on the left, click **Monitoring** →**Events**, and list only the Recommended Actions by using the **Actions** menu (Figure 6-5). Click **Run Fix**.

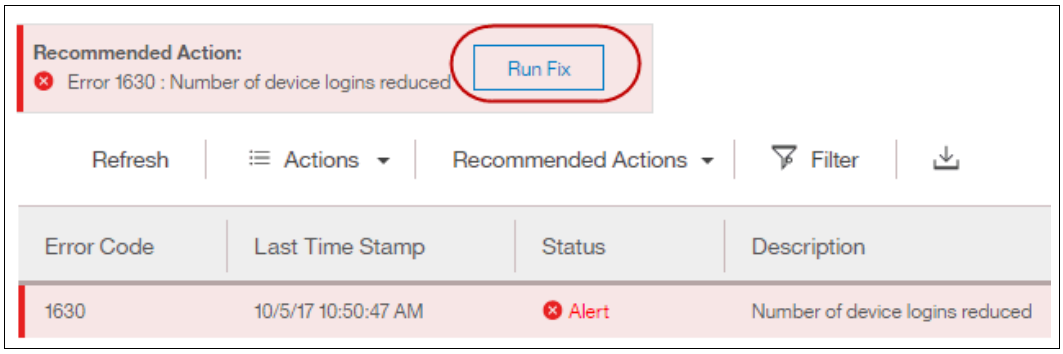

Figure 6-5 Initiate Run Fix procedure from the management GUI

2. The pop-up window prompts you to indicate whether the issue was caused by a planned change or maintenance task, or whether it appeared in an uncontrolled manner (Figure 6-6).

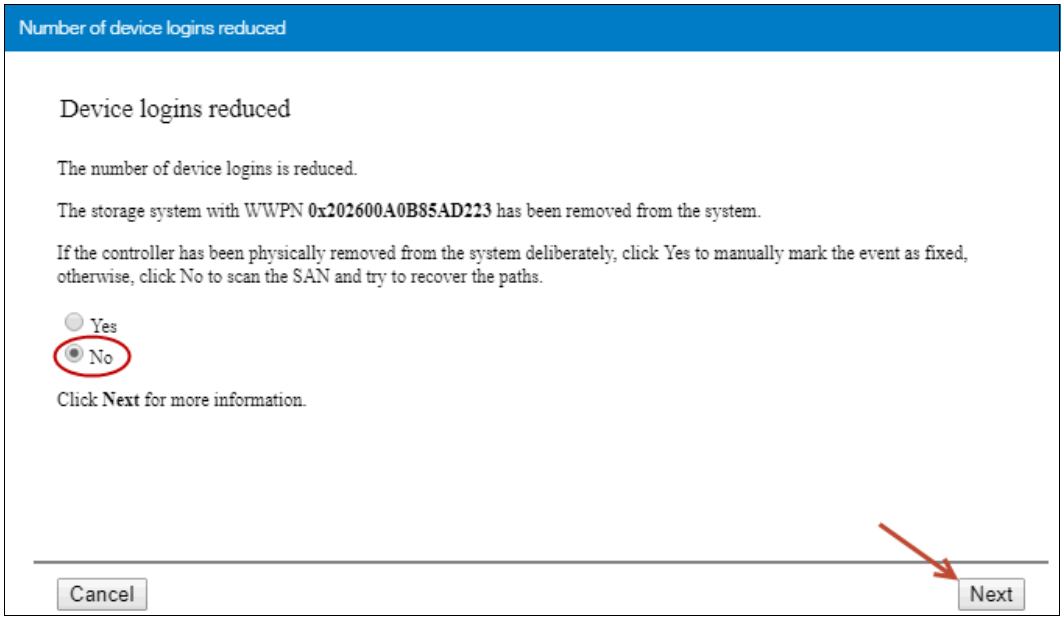

Figure 6-6 Determination of planned action

3. If you answer **Yes**, the fix procedure finishes, assuming that all changes in the system are expected and no other action is necessary, perhaps because you are decommissioning an old disk controller. However, our example simulates a zoning error and we continue to follow the fix procedure. Select **No** and click **Next**.

4. In the next window (Figure 6-7), IBM Spectrum Virtualize lists suggested actions and which components must be checked to fix and close the error. When you are sure that all possible technical requirements are met (in our case, we resolved our SAN zoning config), click **Next**.

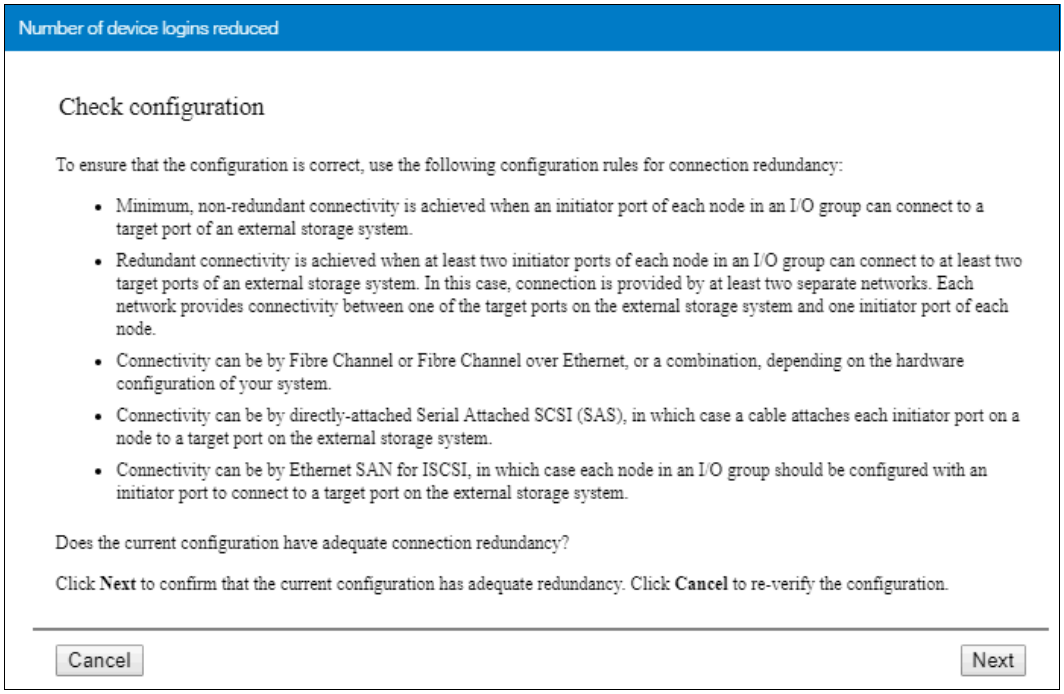

Figure 6-7 Verification steps to eliminate single point of failure

The discovery of managed disks starts (Figure 6-8).

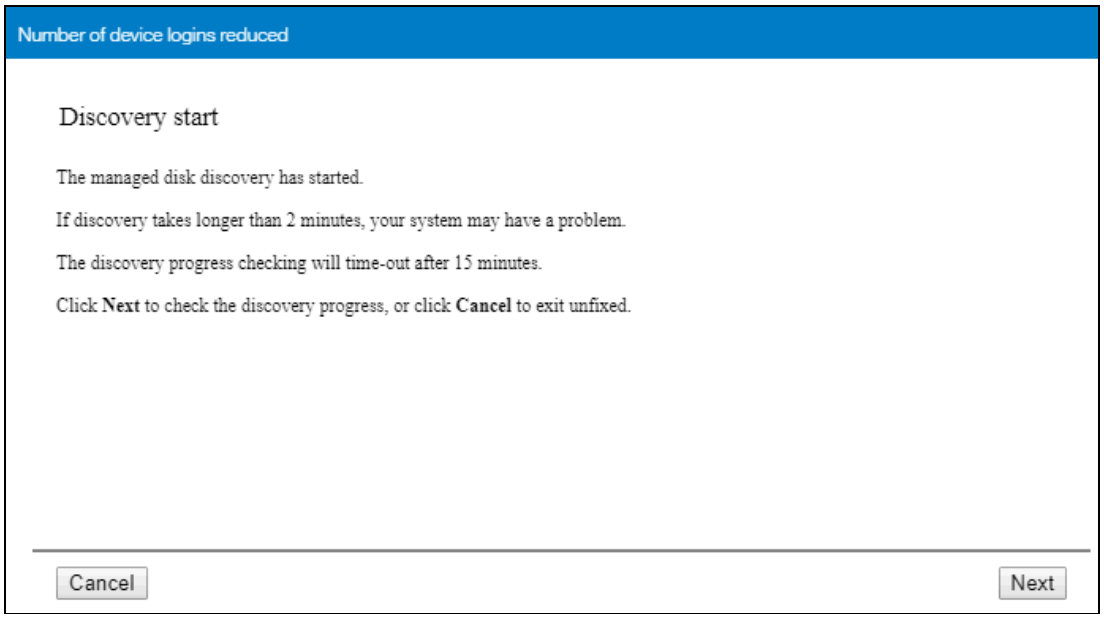

Figure 6-8 Starting the discovery of managed disks

If no other important issue exists, the discovery should finish within 2 minutes, depending on the number of enclosures and installed disk drives (Figure 6-9).

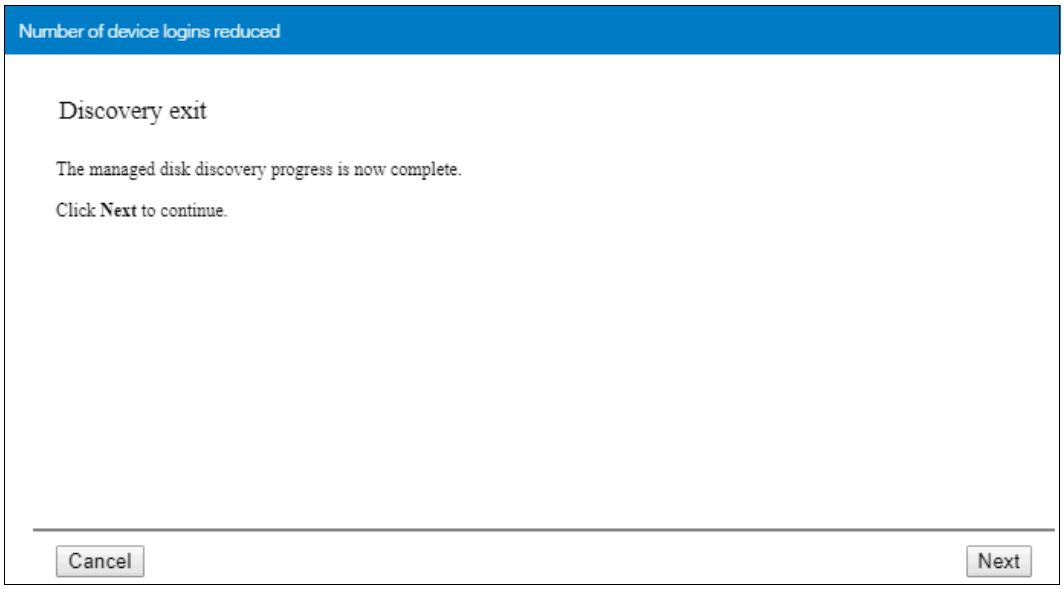

Figure 6-9 Discovery complete

5. The event has been marked as fixed, and you can safely finish the fix procedure. Click **Close** and the event is removed from the list of events (Figure 6-10).

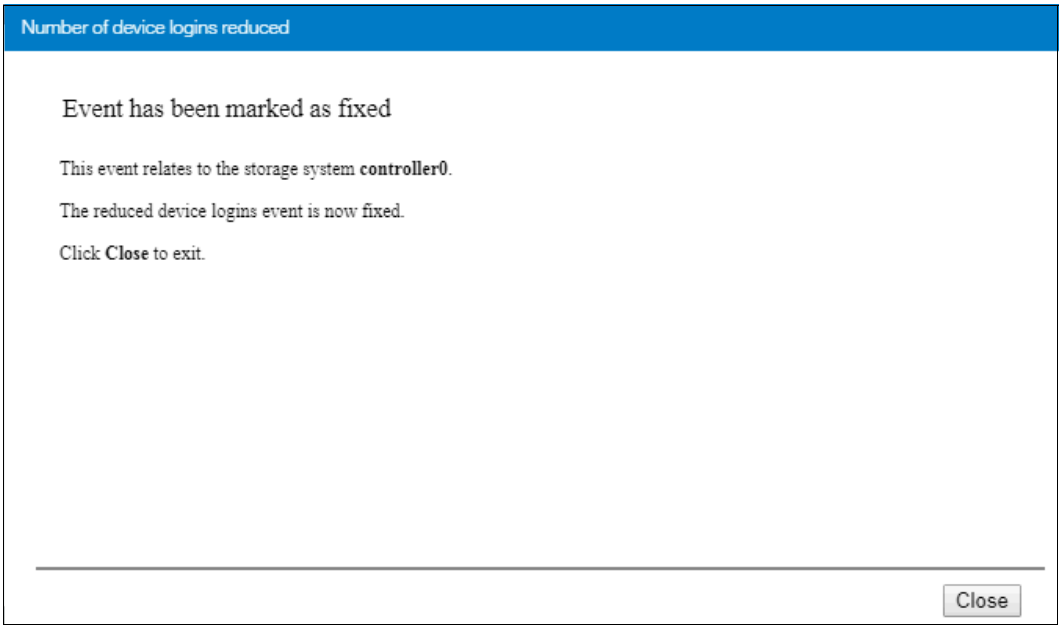

Figure 6-10 Correctly finished fix procedure

Resolve alerts in a timely manner

To minimize any impact to your host systems, always perform the recommended actions as quickly as possible after a problem is reported. Your system is designed to be resilient to most single hardware failures. However, if it operates for any period of time with a hardware failure, the possibility increases that a second hardware failure can result in some volume data that is unavailable. If several unfixed alerts exist, fixing any one alert might become more difficult because of the effects of the others.

6.1.3 Event log details

Multiple views of the events and recommended actions are available. The GUI works like a typical Microsoft Windows pop-up menu, so the event log grid is manipulated through the row that contains the column headings (Figure 6-11):

1. When you click the column icon at the right end of the table heading, a menu for the column choices opens.

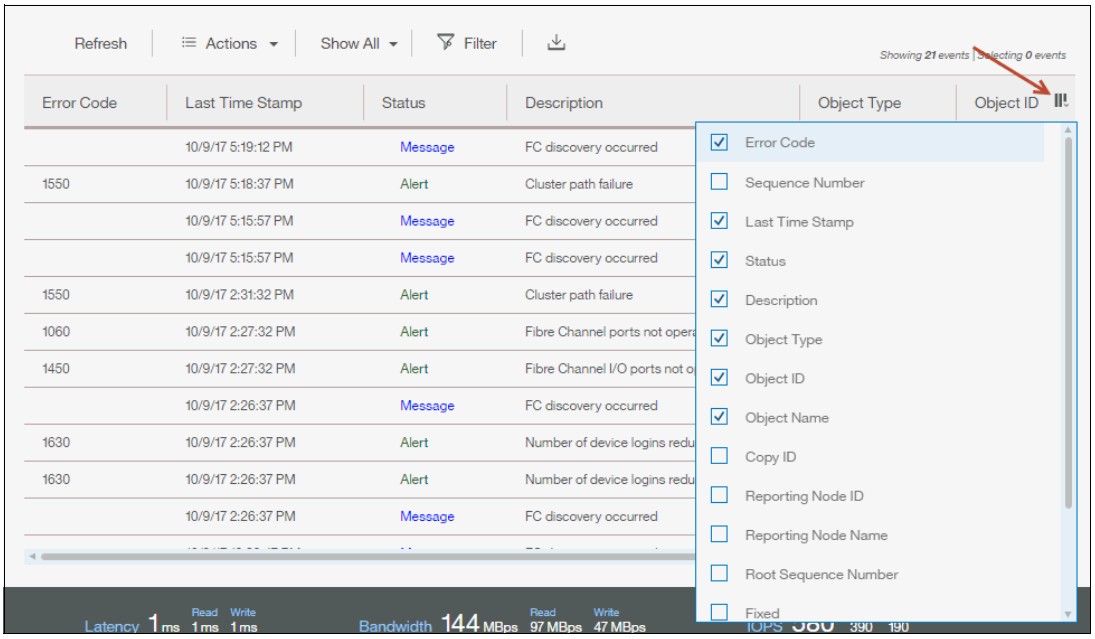

Figure 6-11 Grid options of the event log

2. Select or remove columns as needed. You can then also extend or shrink the width of the column to fit your screen resolution and size. This is the way to manipulate it for most grids in the management GUI of IBM Spectrum Virtualize, not just the events pane.

   Every field of the event log is available as a column in the event log grid. Several fields are useful when you work with IBM Support. The preferred method in this case is to use the **Show All** filter, with events sorted by time stamp. All fields have the sequence number, event count, and the fixed state. Using **Restore Default View** sets the grid back to the defaults.

3. You might want to see more details about each critical event. Some details are not shown in the main grid. To access properties and sense data of a specific event, double-click the specific event anywhere in its row.

The properties window opens (Figure 6-12) with all the relevant sense data. This data includes the first and last time of an event occurrence, worldwide port name (WWPN), and worldwide node name (WWNN), enabled or disabled automatic fix, and so on.

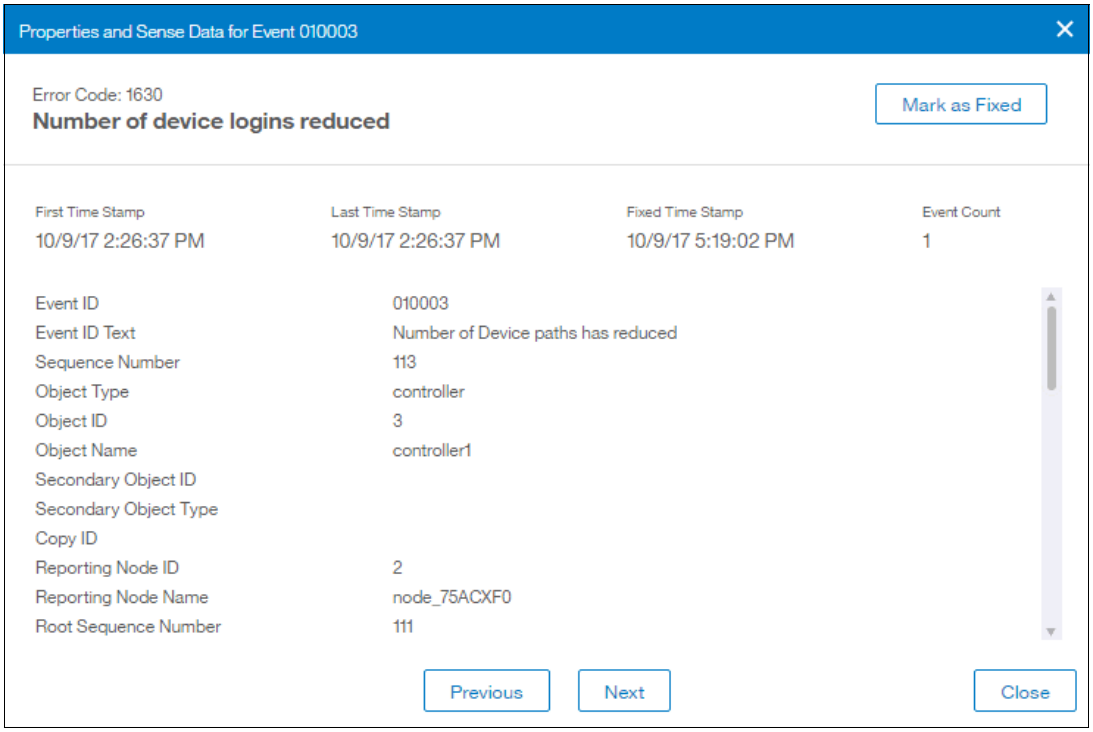

Figure 6-12 Event sense data and properties

For more information about troubleshooting options, see the Troubleshooting section in IBM Knowledge Center:

https://ibm.biz/BdjmgY

6.2 SVC Event Log Definitions for DRP

This section lists the event IDs related to Data Reduction Pools. For more information about troubleshooting options and events, see the Troubleshooting section in IBM Knowledge Center:

https://ibm.biz/Bdjmgz

6.2.1 Event ID 060001

Example 6-1 Event ID 060001

```
Space-efficient VDisk Copy offline due to insufficient space
```

A space-efficient VDisk has been taken offline because there is insufficient allocated real capacity available on the VDisk for the used space to increase further. If the space-efficient VDisk is autoexpand enabled, the managed disk group that it is in also has no free space.

6.2.2 Event ID 060002

Example 6-2 Event ID 060002

```
Space-efficient VDisk Copy offline due to corrupt metadata
```

A space-efficient VDisk Copy has been taken offline because there is an inconsistency in the SAN Volume Controller metadata that describes the disk contents. This might occur because of data corruption on the physical disk (for example, medium error or data miscompare), the loss of cached metadata (because of a T3 recovery) or because of a software error. The event data gives information on the reason.

The SAN Volume Controller maintains backup copies of the metadata, and it might be possible to repair the space-efficient VDisk using this data. Run the repair VDisk option to start the repair process. This repair process can, however, take some time.

In some situations it might be more appropriate to delete the space-efficient VDisk and reconstruct a new one from a backup or mirror copy. If the repair procedure completes successfully, this error is automatically marked as fixed. If not, another error event (error code 1860) is logged to indicate that the repair failed.

6.2.3 Event ID 060003

Example 6-3 Event ID 060003

```
Space-efficient VDisk Copy offline due to failed repair
```

The attempt to repair a space-efficient volume's metadata, which describes the disk contents, has failed because of problems with the automatically maintained backup copy of this data. The error event data describes the problem. Delete the space-efficient VDisk and reconstruct a new one from a backup or mirror copy. Then mark the error fixed. Also mark and the original 1862 error fixed.

6.2.4 Event ID 060004

Example 6-4 Event ID 060004

```
Compressed Virtual Disk Copy offline due to insufficient space
```

The VDisk copy is offline because it needs more real capacity to perform host I/O, and there is no available real capacity: either the copy does not have autoexpand set and it is full, or it does have autoexpand set but there is no free space left in the MDisk group.

6.2.5 Event ID 060005

Example 6-5 Event ID 060005

```
Compressed Virtual Disk Copy offline due to corrupt metadata
```

A compressed VDisk cannot be accessed because of a CRC error in data or metadata on that disk, and it has been taken offline.

6.2.6 Event ID 060006

Example 6-6 Event ID 060006

```
Compressed Virtual Disk Copy offline due to failed repair
```

A disk was corrupt and a repair was issued, but the repair didn't work.

6.2.7 Event ID 060007

Example 6-7 Event ID 060007

```
Compressed Virtual Disk Copy has bad blocks
```

At least one of the blocks on the VDisk copy is unreadable. This is caused by either:

- A replication task, such as IBM FlashCopy® or VDisk Mirroring encountering a medium error on their read and using a virtual medium error to replicate this state; or
- RtC receiving a medium error failure on a data block read and so marking the block as being exempt from optimization.

6.2.8 Event ID 060008

Example 6-8 Event ID 060008

```
Data reduction pool meta data corrupt
```

Some aspect of a volume's metadata that is part of a Data Reduction Pool, is corrupt. All volume copies' data that is part of the Data Reduction Pool is offline as a result.

The SAN Volume Controller maintains backup copies of the metadata, and it might be possible to repair the space-efficient VDisk using this data. Run the repair VDisk option to start the repair process. However, this repair process can take some time.

In some situations, it might be more appropriate to delete the space-efficient VDisk and reconstruct a new one from a backup or mirror copy. If the repair procedure completes successfully, this error is automatically marked as fixed; otherwise, another error event (error code 1860) is logged to indicate that the repair failed.

6.2.9 Event ID 060009

Example 6-9 Event ID 060009

```
Pool's virtual disk copies offline due to failed data reduction pool repair
```

The Data Reduction Pool repair process has encountered an error that prevents the recovery process from completing.

6.2.10 Event ID 060010

Example 6-10 Event ID 060010

```
Virtual Disk Copies offline due to insufficient space in Data Reduction Pool
```

A Data Reduction Pool has insufficient space that prevents thin or compressed virtual disk copies from expanding. In this case, the pool is taken offline. The pool can be brought online through any of the following methods:

- Add extra capacity to the storage pool (add extra MDisks to the pool).
- Delete some of the volumes in the storage pool and wait for garbage collection to reclaim some of the deleted capacity.
- If the pool is offline but has a large amount of reclaimable capacity, wait for garbage collection to reclaim and free some storage.

After a method has been followed, the system administrator needs to manually mark the event as fixed to bring the pool online.

6.2.11 Event ID 060011

Example 6-11 Event ID 060011

```
Virtual Disk Copies offline due to insufficient Journal space in Data Reduction
Pool
```

A Data Reduction Pool has insufficient journal space that prevents metadata modification of thin or compressed virtual disk copies.

6.2.12 Event ID 986001

Example 6-12 Event ID 986001

```
Space-efficient VDisk Copy data in a node is pinned
```

A node in an I/O group has been unable to write a space-efficient VDisk Copy's metadata to MDisk. This might be because a MDisk has gone offline. The metadata remains in the node's cache until it can be written to MDisk. The space-efficient VDisk Copy is offline.

6.2.13 Event ID 986002

Example 6-13 Event ID 986002

```
All space-efficient VDisk Copy data in a node is now unpinned
```

After a period when the space-efficient VDisk Copy's metadata has been pinned, the node has been able to write all the metadata it holds for a space-efficient VDisk to an MDisk.

6.2.14 Event ID 986010

Example 6-14 Event ID 986010

```
Space-efficient Virtual Disk Copy import failure
```

The creation of a new, space-efficient, image mode VDisk copy has been attempted. SAN Volume Controller has determined that it cannot interpret the space-efficient VDisk Copy's control metadata. This is because the image mode disk that is being imported was created by a cluster running a newer version of SAN Volume Controller.

The VDisk Copy is created, but it is offline. To bring the VDisk Copy online, the cluster software version must be upgraded to at least the version that created the image mode VDisk Copy. The import restarts automatically after the software version is upgraded. Alternatively, the VDisk Copy can be deleted.

6.2.15 Event ID 986011

Example 6-15 Event ID 986011

```
Space-efficient VDisk Copy import successful
```

The creation of an image mode space-efficient VDisk Copy, through the import of an existing space-efficient image mode disk, completed successfully.

6.2.16 Event ID 986020

Example 6-16 Event ID 986020

```
Space-efficient VDisk Copy space warning
```

The used capacity of the space-efficient VDisk has exceeded the warning threshold for the VDisk (the threshold is configurable).

Action should be taken to ensure that the space-efficient VDisk's requirements do not continue to grow so they exceed to real disk capacity available.

If the space-efficient VDisk copy has autoexpand enabled, the warning threshold might have been set to indicate that the used capacity of the VDisk Copy is approaching the virtual size of the VDisk Copy. The actions that can be taken depend on whether the space-efficient virtual disk has autoexpand enabled.

If autoexpand is enabled:

- Ignore the event and increase the value of the warning threshold.
- Expand the virtual capacity of the VDisk to provide more storage to the host application.
- If the VDisk is expanding rapidly, you should check for an application error. The autoexpand option can be turned off, or in an extreme case the VDisk can be forced offline by moving it to the recovery I/O group.

If autoexpand is not enabled, this event indicates that the allocated space is filling up. You can take one or more of the following actions:

- Ignore the event and increase the value of the warning threshold
- Expand the real capacity of the VDisk to provide more storage for the space-efficient virtual disk.

- Enable the autoexpand option to allow the cluster to automatically expand the volume when required.

If the used space of the space-efficient VDisk Copy is close to the virtual size of the disk, and the virtual size is not going to be increased, then you should consider migrating the space-efficient VDisk Copy to a fully allocated VDisk Copy, which will provide more efficient access to the data.

6.2.17 Event ID 986030

Example 6-17 Event ID 986030

```
Space-efficient VDisk Copy repair started
```

An attempt to repair a space-efficient VDisk copy has started.

6.2.18 Event ID 986031

Example 6-18 Event ID 986031

```
Space-efficient VDisk Copy repair successful
```

SAN Volume Controller has successfully repaired a space-efficient VDisk copy. The VDisk Copy is online and the 1862 error is marked fixed.

6.2.19 Event ID 986032

Example 6-19 Event ID 986032

```
Space-efficient VDisk Copy validation started
```

A space-efficient VDisk copy validation has started.

6.2.20 Event ID 986033

Example 6-20 Event ID 986033

```
Space-efficient VDisk Copy validate successful
```

SAN Volume Controller has completed a validation of a space-efficient VDisk copy and found no problems. The VDisk Copy is online.

6.2.21 Event ID 986034

Example 6-21 Event ID 986034

```
Compressed Virtual Disk Copy import successful
```

Compressed image-mode VDisk copy has been imported successfully.

6.2.22 Event ID 986035

Example 6-22 Event ID 986035

```
Compressed Virtual Disk Copy space warning
```

Used capacity has exceeded warning capacity.

6.2.23 Event ID 986036

Example 6-23 Event ID 986036

```
Compressed Virtual Disk Copy repair started
```

A compressed VDisk copy repair has started.

6.2.24 Event ID 986037

Example 6-24 Event ID 986037

```
Compressed Virtual Disk Copy repair successful
```

A repair that was started by for example, the **recovervdisk** command, has completed successfully and the VDisk is online again.

6.2.25 Event ID 986038

Example 6-25 Event ID 986038

```
Compressed Virtual Disk Copy has too many bad blocks.
```

A bad block has been requested. The VDisk Copy has so many bad blocks on it that it cannot create another one. The VDisk Copy is not directly taken offline when this happens. However, the operation that is writing the data also responds to the failed write: FlashCopy stops the mapping and VDisk Mirroring holds the VDisk copy offline.

6.2.26 Event ID 986039

Example 6-26 Event ID 986039

```
A data reduction pool repair process has begun.
```

This event is raised to inform the system administrator that the repair process has begun.

6.2.27 Event ID 986040

Example 6-27 Event ID 986040

```
A data reduction pool repair process has completed successfully.
```

This event is raised to inform the system administrator that the repair process has completed successfully.

6.2.28 Event ID 986105

Example 6-28 Event ID 986105

```
Compressed volume copy may require repair due to possible data loss
```

The system has detected that this volume copy could be affected by a data loss issue, APAR HU00898. See the following webpage, which explains how to verify whether the volume copy is affected, and how to repair it:

http://www.ibm.com/support/docview.wss?uid=ssg1S1005361

6.2.29 Event ID 986106

Example 6-29 Event ID 986106

```
Compressed volume virtual size limitation breached
```

Compressed volume size was reduced, and previously created volumes are in breach of the new size.

6.2.30 Event ID 986107

Example 6-30 Event ID 986107

```
Compressed volume real capacity size limitation breached
```

The cluster has at least one compressed volume that uses more real capacity than allowed.

6.3 Monitoring

An important step is to correct any issues that are reported by your IBM Spectrum Virtualize system as soon as possible. Configure your system to send automatic notifications when a new event is reported. To avoid having to monitor the management GUI for new events, select the type of event for which you want to be notified. For example, restrict notifications to just events that require action. Several event notification mechanisms exist:

Email An event notification can be sent to one or more email addresses. This mechanism notifies individuals of problems. Individuals can receive notifications wherever they have email access, including mobile devices.

SNMP An SNMP traps report can be sent to a data center management system, such as IBM Systems Director, that consolidates SNMP reports from multiple systems. With this mechanism, you can monitor your data center from a single workstation.

Syslog A syslog report can be sent to a data center management system that consolidates syslog reports from multiple systems. With this option, you can monitor your data center from a single location.

If your system is within warranty or if you have a hardware maintenance agreement, configure your IBM Spectrum Virtualize system to send email events directly to IBM if an issue that requires hardware replacement is detected. This mechanism is known as *Call Home*. When this event is received, IBM automatically opens a problem report and, if appropriate, contacts you to help resolve the reported problem.

Important: If you set up Call Home to IBM, ensure that the contact details that you configure are correct and kept up to date. Personnel changes can cause delays in IBM making contact.

6.3.1 Email notifications and the Call Home function

The Call Home function of IBM Spectrum Virtualize uses the email notification being sent to the specific IBM support center. Therefore, the configuration is similar to sending emails to the specific person or system owner. The following procedure summarizes how to configure email notifications and emphasizes what is specific to Call Home:

1. Prepare the contact information that you want to use for the email notification and verify the accuracy of the data. From the GUI menu, click **Settings** →**Notifications** (Figure 6-13).

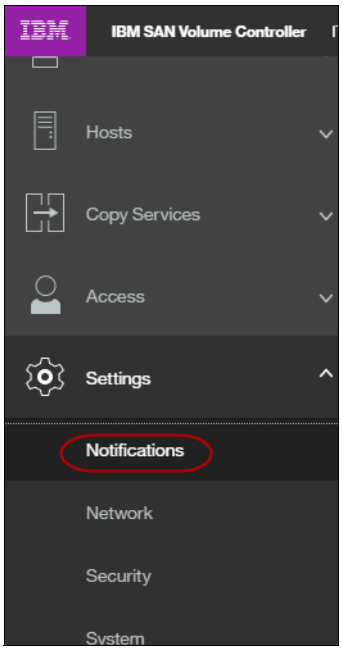

Figure 6-13 Notifications menu

2. Select **Email** and then click **Enable Notifications** (Figure 6-14). You can also access the IBM eLearning movie for more technical details:

   http://www.ibm.com/support/knowledgecenter/ST3FR7/welcome?cp=ST3FR7%2F0&lang=en

   For the correct functionality of email notifications, ask your network administrator if Simple Mail Transfer Protocol (SMTP) is enabled on the management network and is not, for example, blocked by firewalls. Be sure to test accessibility to the SMTP server using the **`telnet`** command (port 25 for a non-secured connection, or port 465 for Secure Sockets Layer (SSL)-encrypted communication) using any server in the same network segment.

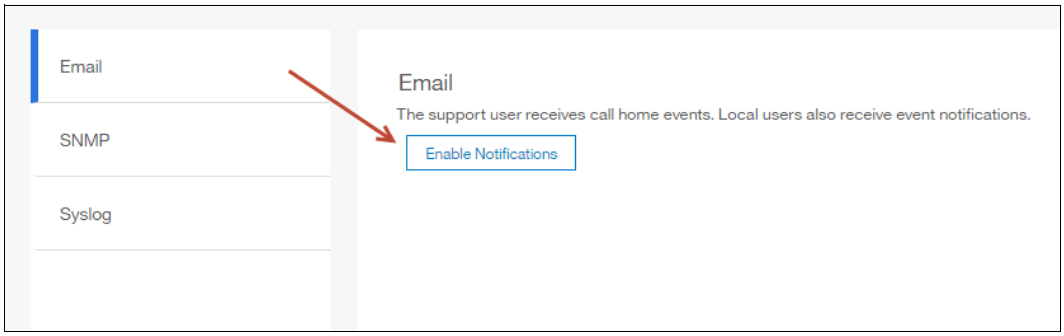

Figure 6-14 Configuration of email notifications

3. After clicking **Next** in the welcome window, provide the information about the location of the system (Figure 6-15) and contact information of the IBM Spectrum Virtualize administrator (Figure 6-16 on page 74) to be contactable by IBM Support. *Always* keep this information current.

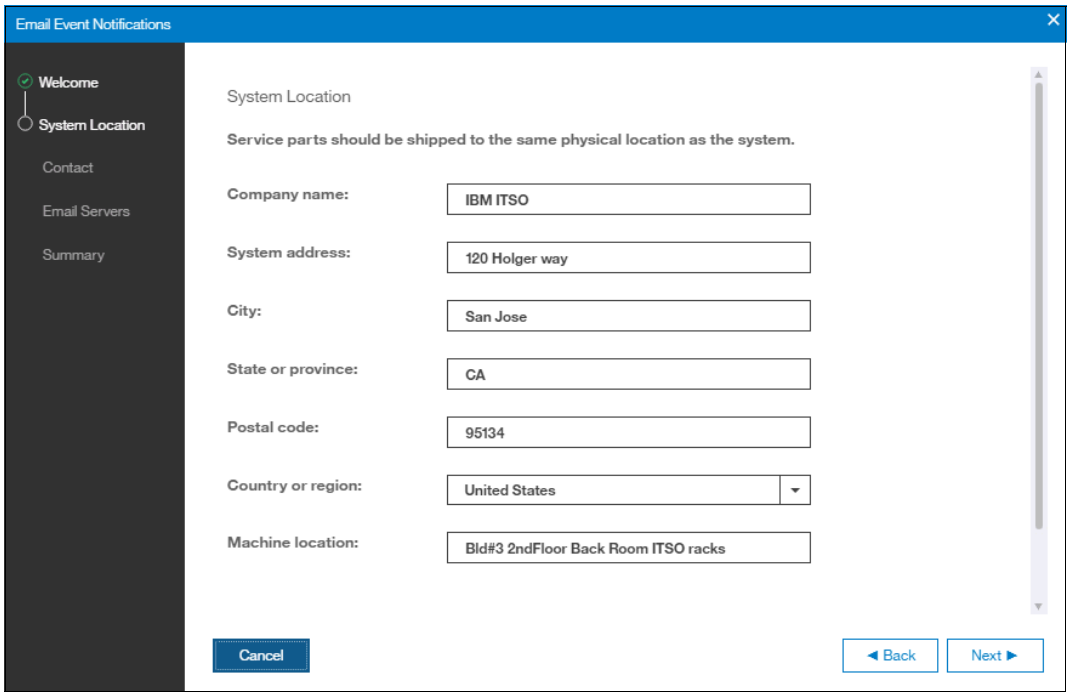

Figure 6-15 Location of the device

Figure 6-16 shows the contact information of the owner.

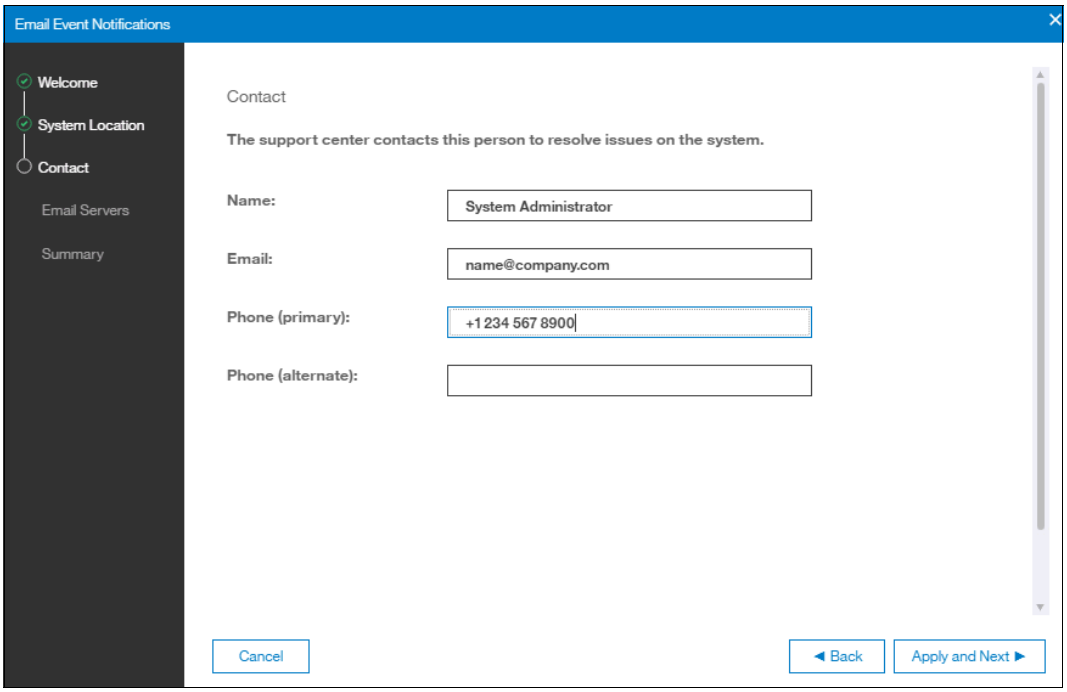

Figure 6-16 Contact information

4. Configure the SMTP server according to the instruction in Figure 6-17. When the correct SMTP server is provided, you can test the connectivity using **Ping** to verify that it can be contacted. Then, click **Apply and Next**.

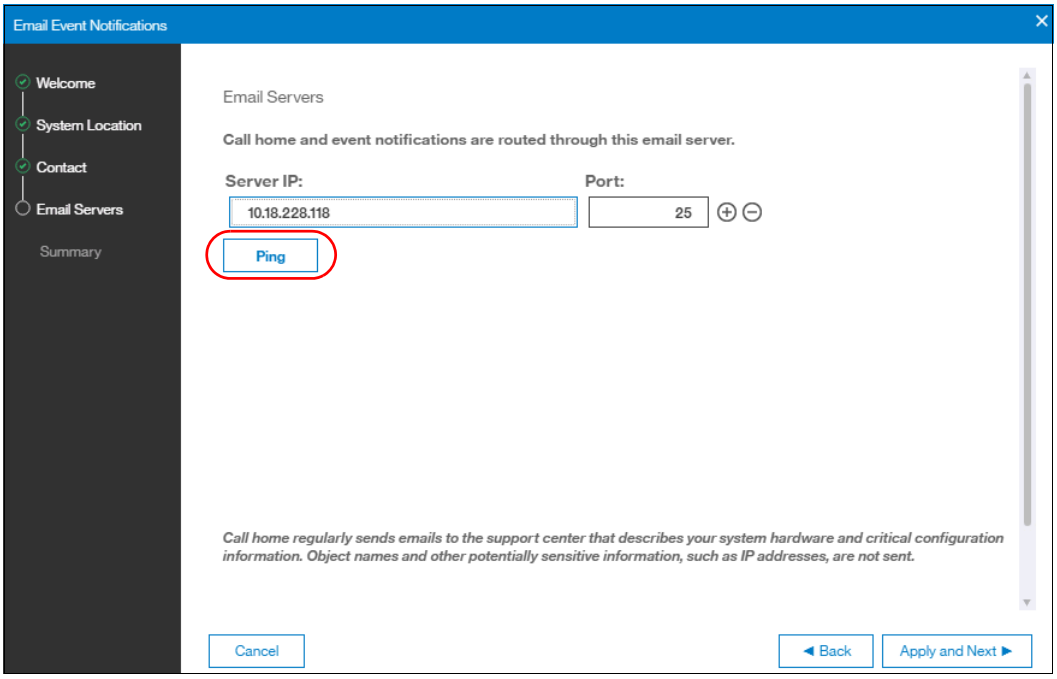

Figure 6-17 Configure email servers and inventory reporting

5. A summary window is displayed. Verify all of the information, then click **Finish**. You are then returned to the Email Settings window where you can verify email addresses of IBM Support (`callhome0@de.ibm.com`) and optionally add local users who also need to receive notifications. See Figure 6-18 for details.

   The default support email address `callhome0@de.ibm.com` is predefined by the system to receive Error Events and Inventory. Do not change these settings or disable the 7 day reporting interval at the bottom of the Settings window.

   You can modify or add local users using **Edit** mode after the initial configuration is saved.

   The **Inventory Reporting** function is enabled by default for Call Home. Rather than reporting a problem, an email is sent to IBM that describes your system hardware and critical configuration information. Object names and other information, such as IP addresses, are not included. By default the inventory email is sent weekly, enabling an IBM Cloud service to analyze and inform you if the hardware or software that you are using requires an update because of any known issue.

   Figure 6-18 shows the configured email notification and Call Home settings.

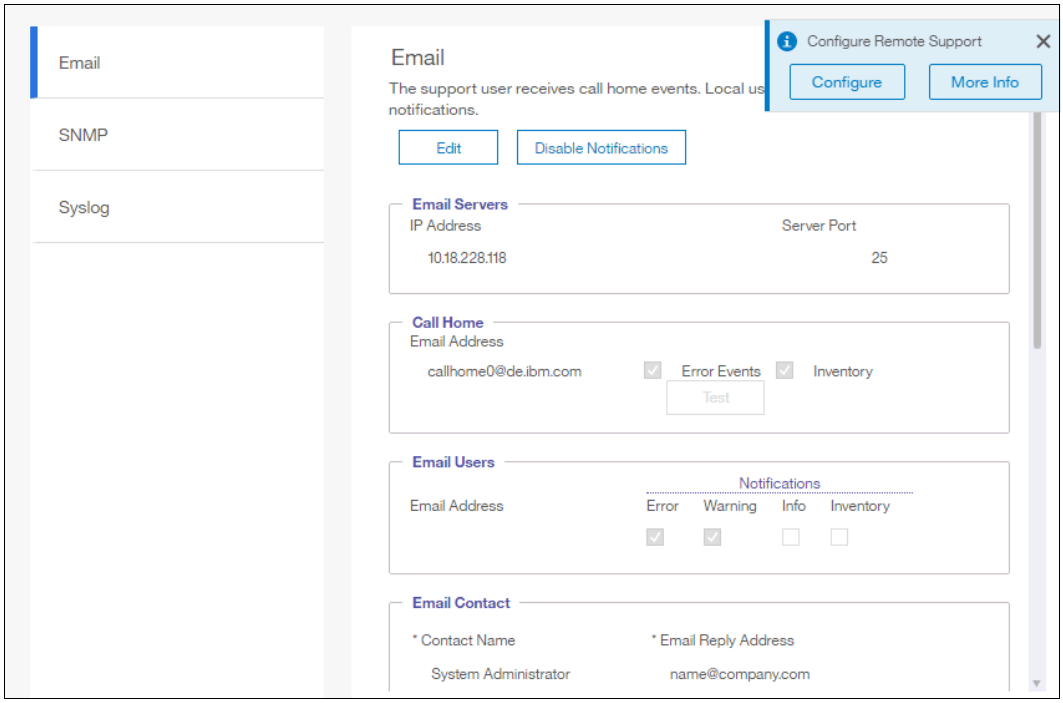

Figure 6-18 Setting email recipients and alert types

6. After completing the configuration wizard, test the email function. To do so, enter **Edit** mode, as illustrated in Figure 6-19. In the same window, you can define additional email recipients or alter any contact and location details as needed.

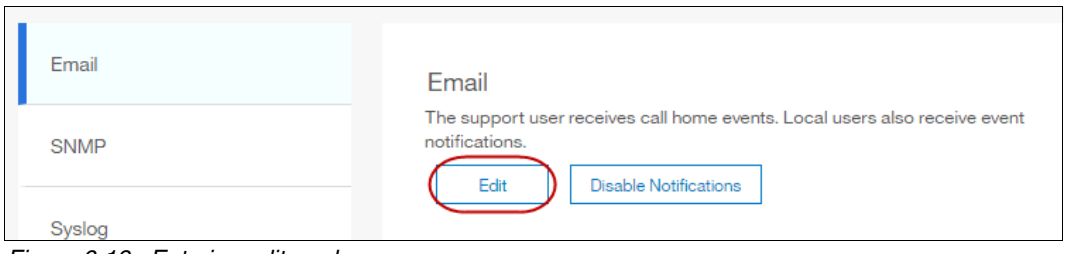

Figure 6-19 Entering edit mode

We strongly suggest that you keep the sending inventory option enabled to IBM support. However, it might not be of interest to local users, although inventory content can serve as a basis for inventory and asset management.

7. In Edit mode, you can change any of the previously configured settings. After you are finished editing these parameters, adding more recipients, or just testing the connection, save the configuration to make the changes take effect (Figure 6-20).

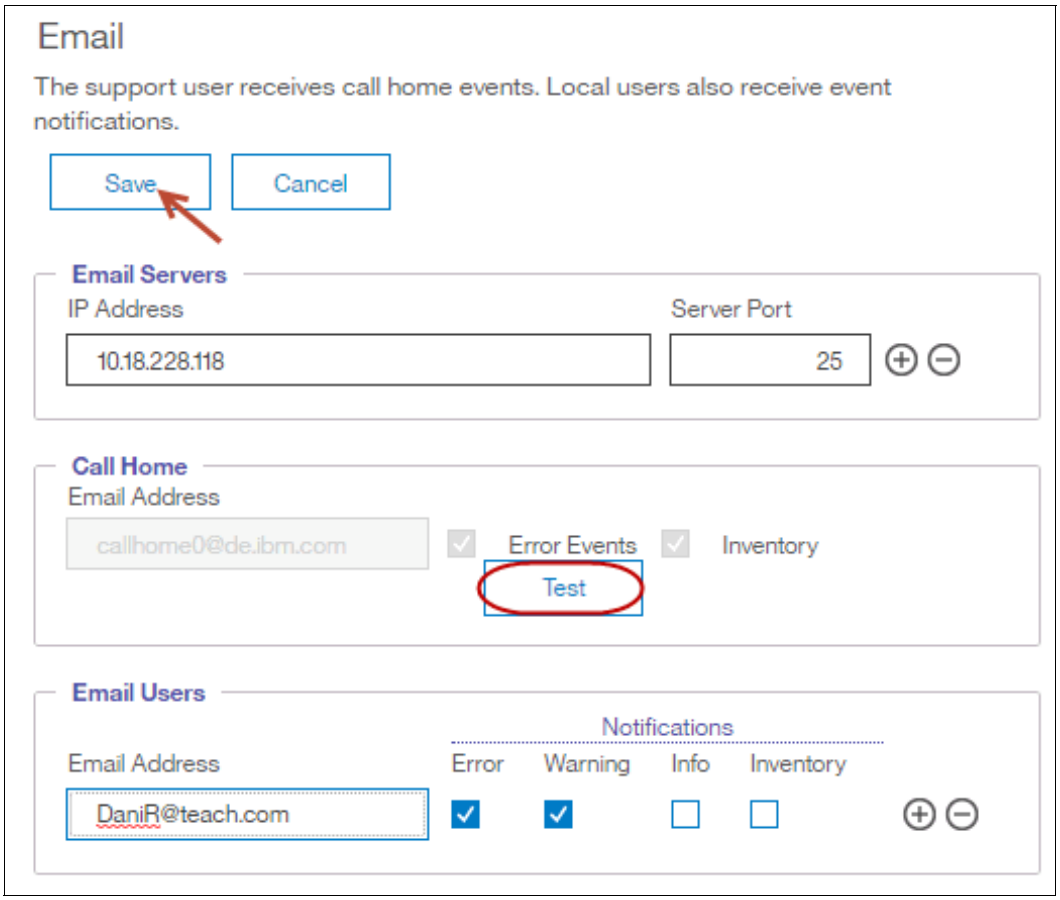

Figure 6-20 Saving modified configuration

Note: The **Test** button displays for new email users after first saving and then editing again.

Disabling and enabling notifications

At any time, you can temporarily or permanently disable email notifications, as shown in Figure 6-21. This is good practice when performing activities in your environment that might generate errors on your IBM Spectrum Virtualize system, such as SAN reconfigurationor replacement activities. After the planned activities, remember to re-enable the email notification function. The same results can be achieved with the CLI **`svctask stopmail`** and **`svctask startmail`** commands.

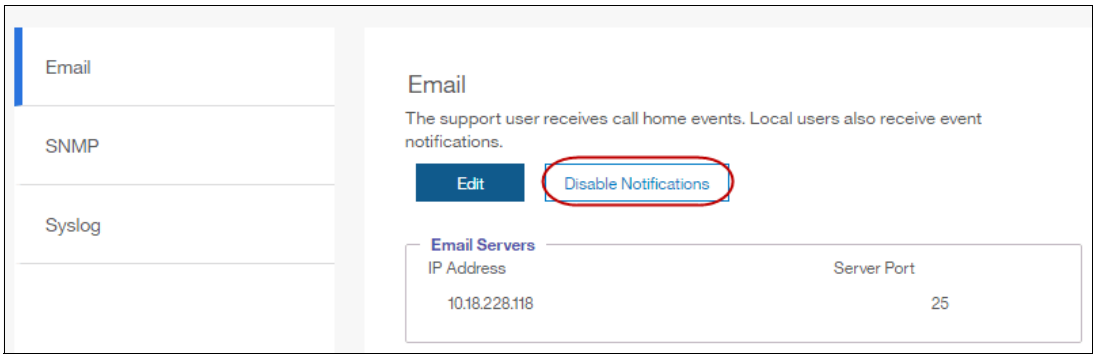

Figure 6-21 Disabling or enabling email notifications

6.3.2 Remote Support Assistance

Remote Support Assistance, introduced with V8.1, allows IBM support to remotely connect to the IBM Spectrum Virtualize system through a secure tunnel to perform analysis, log collection, and software updates. The tunnel can be enabled ad hoc by the client, or you can enable a permanent connection if wanted.

Note: Clients who have purchased Enterprise Class Support (ECS) are entitled to IBM support using Remote Support Assistance to quickly connect and diagnose problems. However, IBM Support might choose to use this feature on non-ECS systems at their discretion. Therefore, configure and test the connection on all systems.

If you are enabling Remote Support Assistance, then ensure that the following prerequisites are met:

1. Ensure that call home is configured with a valid email server.
2. Ensure that a valid service IP address is configured on each node of the IBM Spectrum Virtualize storage system.
3. If your IBM Spectrum Virtualize storage system is behind a firewall, or if you want to route traffic from multiple storage systems to the same place, you must configure a Remote Support Proxy server. Before you configure remote support assistance, the proxy server must be installed and configured separately. During the setup for support assistance, specify the IP address and the port number for the proxy server on the Remote Support Centers window.
4. If you do not have firewall restrictions and the nodes of the IBM Spectrum Virtualize storage system are directly connected to the internet, request your network administrator to allow connections to `129.33.206.139` and `204.146.30.139` on Port 22.
5. Both uploading support packages and downloading software require direct connections to the internet. A DNS server must be defined on your IBM Spectrum Virtualize storage system for both of these functions to work.

6. To ensure that support packages are uploaded correctly, configure the firewall to allow connections to the following IP addresses on port 443: `129.42.56.189`, `129.42.54.189`, and `129.42.60.189`.
7. To ensure that software is downloaded correctly, configure the firewall to allow connections to the following IP addresses on port 22: `170.225.15.105`, `170.225.15.104`, `170.225.15.107`, `129.35.224.105`, `129.35.224.104`, and `129.35.224.107`.

Figure 6-22 shows a pop-up window that appears in the GUI after updating to V8.1. It prompts you to configure IBM Spectrum Virtualize for Remote Support. You can select to not enable it, open a tunnel when needed, or to open a permanent tunnel to IBM.

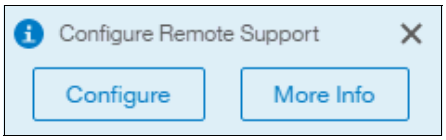

Figure 6-22 Prompt to configure Remote Support Assistance

You can choose to configure IBM Spectrum Virtualize, learn some more about the feature, or just close the window by clicking the **X**. Figure 6-23 shows how you can find the Setup Remote Support Assistance if you have closed the window.

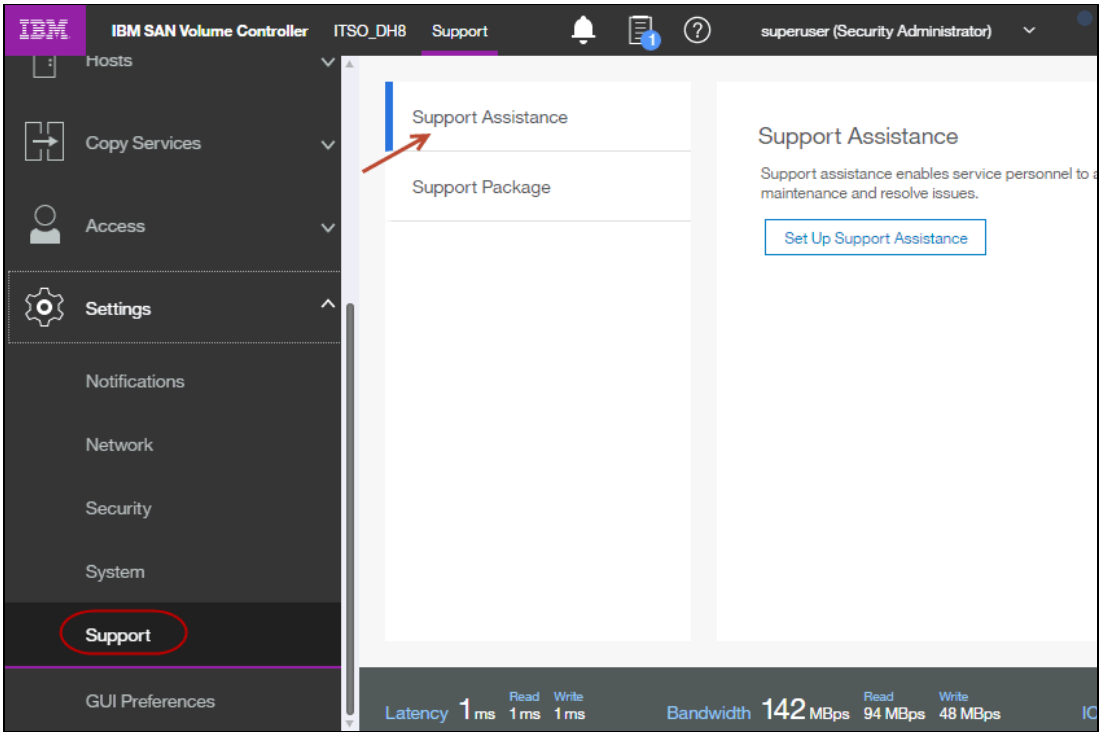

Figure 6-23 Remote Support Assistance menu

Choosing to set up Support Assistance opens a wizard to guide you through the configuration:

1. Figure 6-24 shows the first wizard window. Choose not to enable remote assistance by selecting **I want support personnel to work on-site only** or enable remote assistance by choosing **I want support personnel to access my system both on-site and remotely**. Click **Next**.

> **Note:** Selecting **I want support personnel to work on-site only** does not entitle you to expect IBM support to attend on-site for all issues. Most maintenance contracts are for customer-replaceable units (CRU) support, where IBM diagnoses your problem and sends a replacement component for you to replace if required. If you prefer to have IBM perform replacement tasks for you, then contact your local sales person to investigate an upgrade to your current maintenance contract.

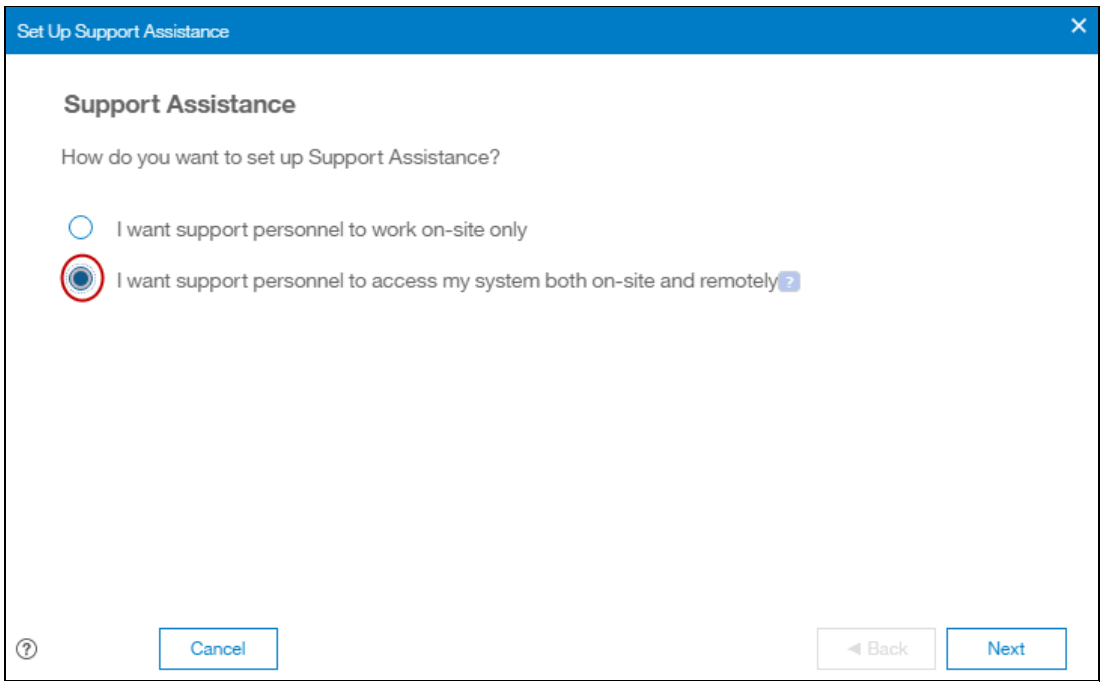

Figure 6-24 Remote Support wizard enable or disable

2. The next window, shown in Figure 6-25, lists the IBM Support center's IP addresses and SSH port that need to be open in your firewall. You can also define a Remote Support Assistance Proxy if you have multiple IBM Spectrum Virtualize storage systems in the data center, allowing for firewall configuration only being required for the Proxy Server rather than every storage system. We do not have a proxy server and leave the field blank, then click **Next**.

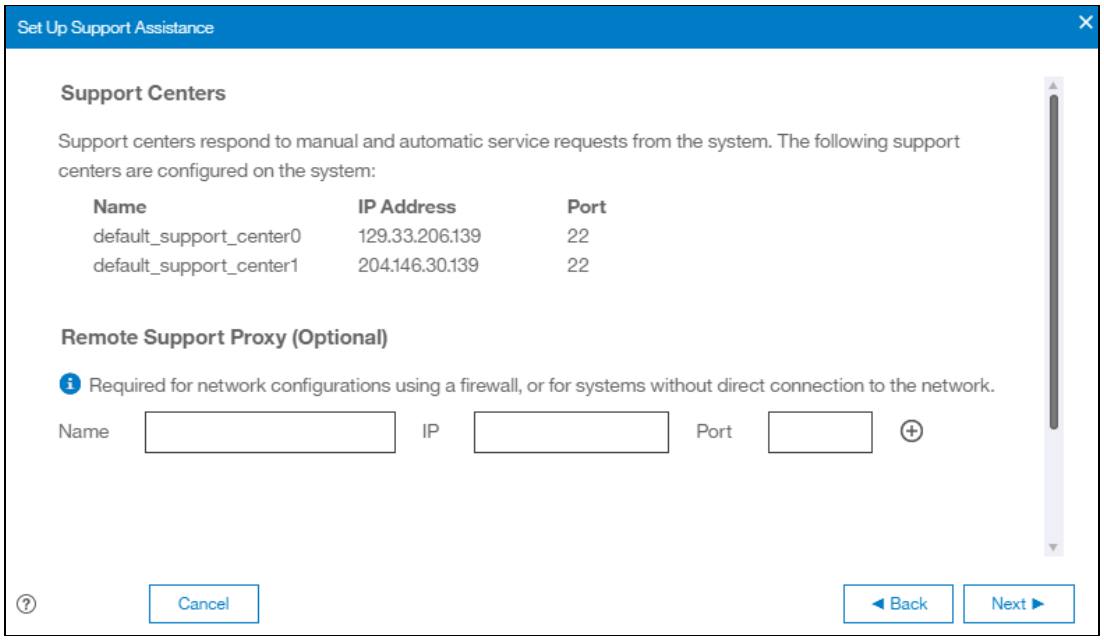

Figure 6-25 Remote Support wizard proxy setup

3. The next window asks if you want to open a tunnel to IBM permanently, allowing IBM to connect to your IBM Spectrum Virtualize storage system **At Any Time**, or **On Permission Only**, as shown in Figure 6-26. **On Permission Only** requires a storage administrator to log on to the GUI and enable the tunnel when required. Click **Finish**.

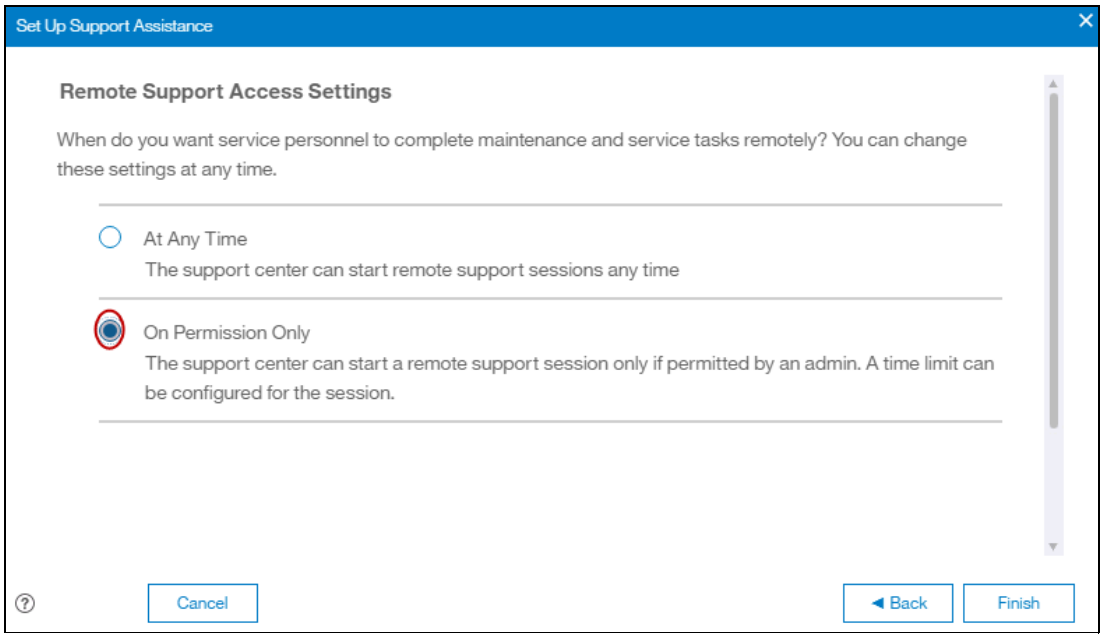

Figure 6-26 Remote Support wizard access choice

4. After completing the remote support setup, you can view the status of any remote connection, start a new session, test the connection to IBM, and reconfigure the setup. In Figure 6-27, we have successfully tested the connection. Click **Start New Session** to open a tunnel for IBM Support to connect through.

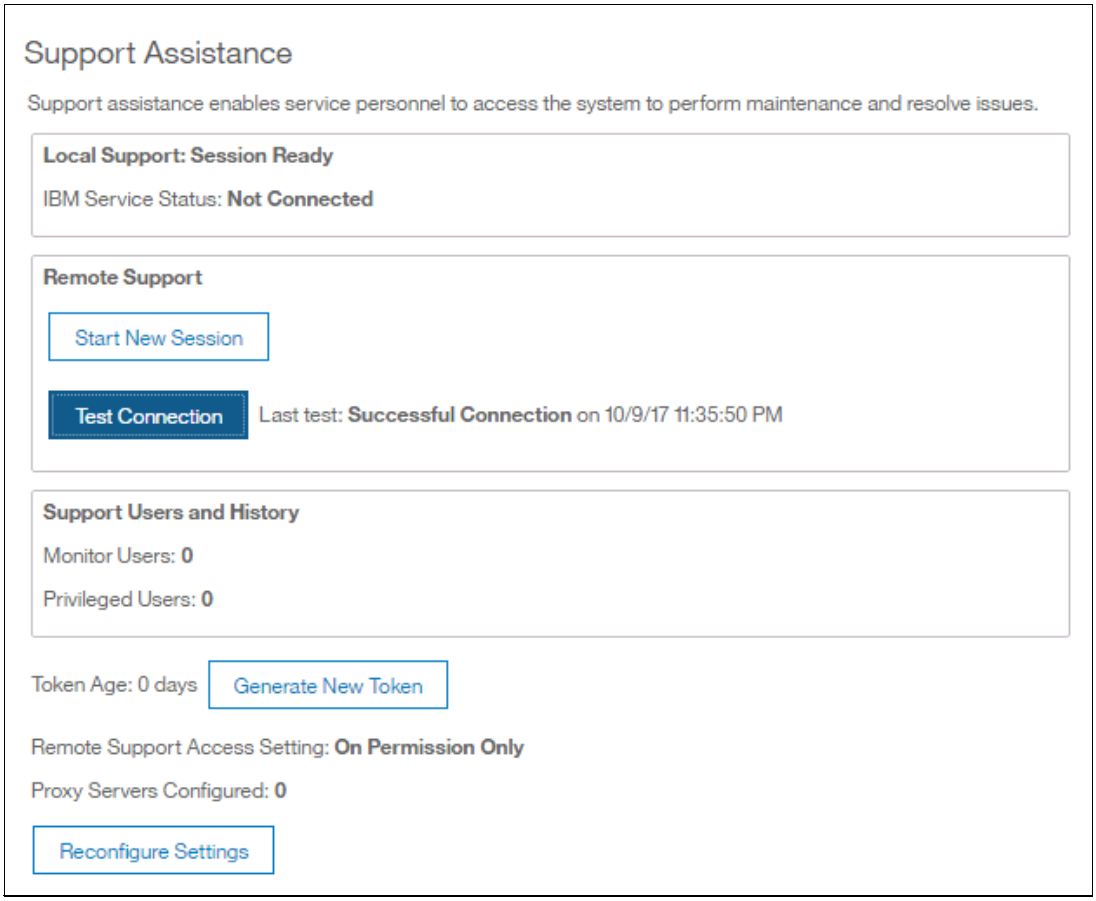

Figure 6-27 Remote Support Status and session management

5. A pop-up window asks how long you would like the tunnel to remain open if there is no activity by setting a timeout value. As shown in Figure 6-28, the connection is established and waits for IBM Support to connect.

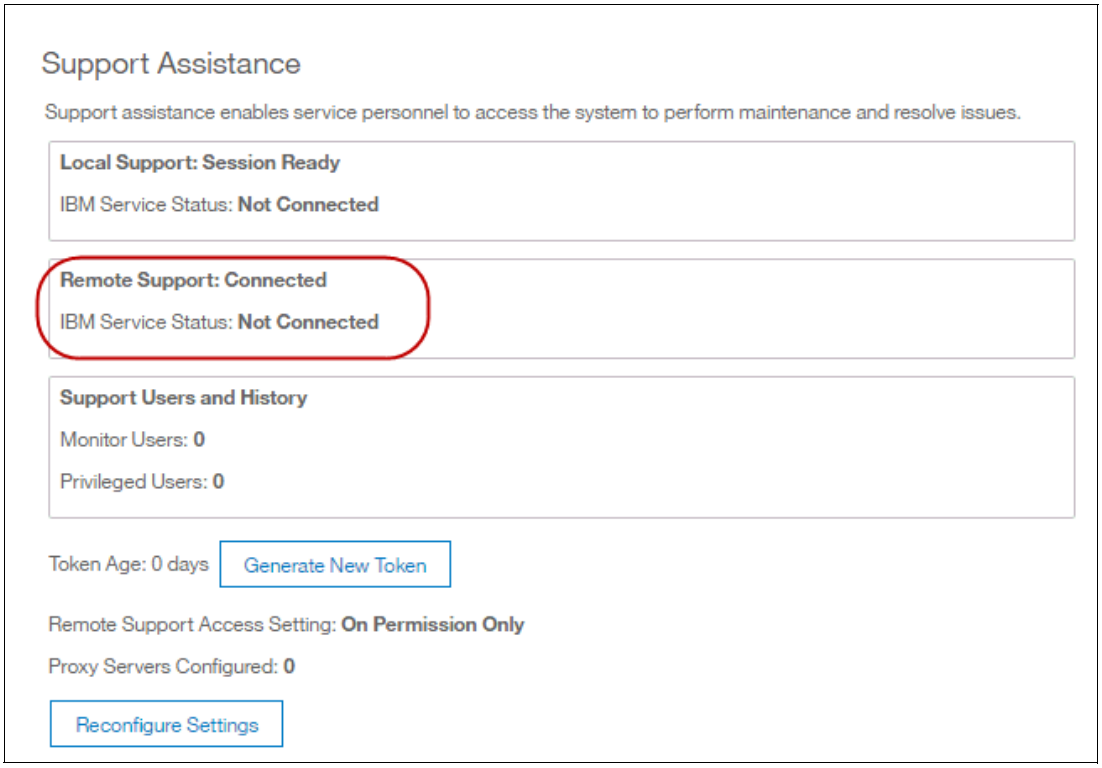

Figure 6-28 Remote Assistance tunnel connected

6.3.3 SNMP configuration

SNMP is a standard protocol for managing networks and exchanging messages. The system can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that are sent by the IBM Spectrum Virtualize storage system.

You can configure an SNMP server to receive various informational, error, or warning notifications by entering the following information (Figure 6-29 on page 83):

- IP Address

  The address for the SNMP server.

- Server Port

  The remote port number for the SNMP server. The remote port number must be a value of 1 - 65535 where the default is port 162 for SNMP.

- Community

  The SNMP community is the name of the group to which devices and management stations that run SNMP belong. Typically, the default of `public` is used.

- Event Notifications:

  Consider the following points about event notifications:

  - Select **Error** if you want the user to receive messages about problems, such as hardware failures, that require prompt action.

    Important: Browse to **Recommended Actions** to run the fix procedures on these notifications.

  - Select **Warning** if you want the user to receive messages about problems and unexpected conditions. Investigate the cause immediately to determine any corrective action such as a space efficient volume running out of space.

    Important: Browse to **Recommended Actions** to run the fix procedures on these notifications.

  - Select **Info** if you want the user to receive messages about expected events. No action is required for these events.

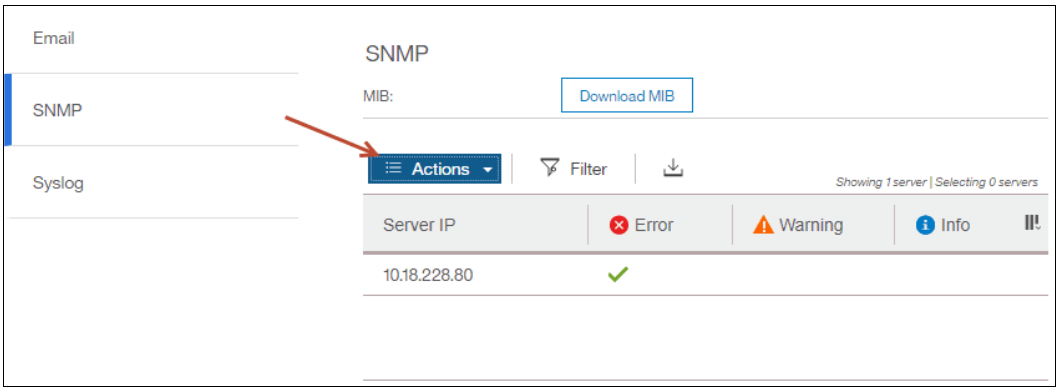

Figure 6-29 SNMP configuration

To add an SNMP server, click **Actions** →**Add** and fill out the Add SNMP Server window, as shown in Figure 6-30. To remove an SNMP server, click the line with the server that you want to remove, and select **Actions** →**Remove**.

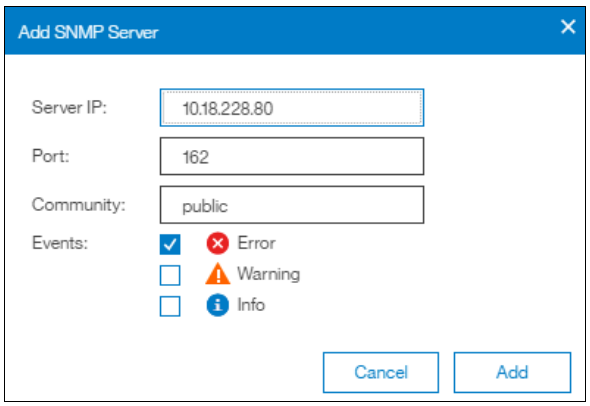

Figure 6-30 Add SNMP Server

6.3.4 Syslog notifications

The syslog protocol is a standard protocol for forwarding log messages from a sender to a receiver on an IP network. The IP network can be IPv4 or IPv6. The system can send syslog messages that notify personnel about an event.

You can configure a syslog server to receive log messages from various systems and store them in a central repository by entering the following information (Figure 6-31):

- IP Address

  The IP address for the syslog server.

- Facility

  The facility determines the format for the syslog messages. The facility can be used to determine the source of the message.

- Message Format

  The message format depends on the facility. The system can transmit syslog messages in the following formats:

  – The concise message format provides standard detail about the event.
  – The expanded format provides more details about the event.

- Event Notifications

  Consider the following points about event notifications:

  – Select **Error** if you want the user to receive messages about problems, such as hardware failures, that must be resolved immediately.

  > **Important:** Browse to **Recommended Actions** to run the fix procedures on these notifications.

  – Select **Warning** if you want the user to receive messages about problems and unexpected conditions. Investigate the cause immediately to determine whether any corrective action is necessary.

  > **Important:** Browse to **Recommended Actions** to run the fix procedures on these notifications.

  – Select **Info** if you want the user to receive messages about expected events. No action is required for these events.

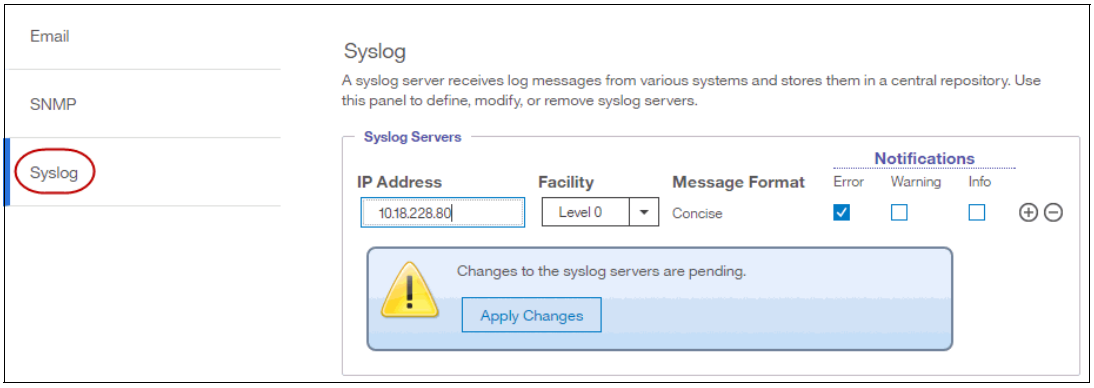

Figure 6-31 Syslog configuration

To remove a syslog server, click the Minus sign (**-**).

To add another syslog server, click the Plus sign (**+**).

The syslog messages can be sent in concise message format or expanded message format.

Example 6-31 shows a compact format syslog message.

Example 6-31 Compact syslog message example

```
IBM2076 #NotificationType=Error #ErrorID=077102 #ErrorCode=1091 #Description=Node
Double fan failed #ClusterName=V7000G2_1 #Timestamp=Wed Jul 02 08:00:00 2017 BST
#ObjectType=Node #ObjectName=Node1 #CopyID=0 #ErrorSequenceNumber=120
```

Example 6-32 shows an expanded format syslog message.

Example 6-32 Full format syslog message example

```
IBM2076 #NotificationType=Error #ErrorID=077102 #ErrorCode=1091 #Description=Node
Double fan failed #ClusterName=V7000G2_1 #Timestamp=Wed Jul 02 08:00:00 2017 BST
#ObjectType=Node #ObjectName=Node1 #CopyID=0 #ErrorSequenceNumber=120 #ObjectID=2
#NodeID=2 #MachineType=2076624#SerialNumber=1234567 #SoftwareVersion=8.1.0.0
(build 13.4.1709291021000)#FRU=fan 31P1847
#AdditionalData(0->63)=0000000046000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000#Additional
Data(64-127)=0000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000
```

6.4 Audit log

The audit log is useful when analyzing past configuration events, especially when trying to determine, for example, how a volume ended up being shared by two hosts, or why the volume was overwritten. The audit log is also included in the `svc_snap` support data to aid in problem determination.

The audit log tracks action commands that are issued through a Secure Shell (SSH) session, through the management GUI or Remote Support Assistance. It provides the following entries:

- Identity of the user who issued the action command.
- Name of the actionable command.
- Time stamp of when the actionable command was issued on the configuration node.
- Parameters that were issued with the actionable command.

The following items are not documented in the audit log:

- Commands that fail are not logged.
- A result code of `0` (success) or `1` (success in progress) is not logged.
- Result object ID of node type (for the **`addnode`** command) is not logged.
- Views are not logged.

Several specific service commands are not included in the audit log:

- **dumpconfig**
- **cpdumps**
- **cleardumps**
- **finderr**
- **dumperrlog**
- **dumpintervallog**
- **svcservicetak dumperrlog**
- **svcservicetask finderr**

Figure 6-32 shows the access to the audit log. Click **Audit Log** in the left menu to see which configuration CLI commands have been run on IBM Spectrum Virtualize storage system.

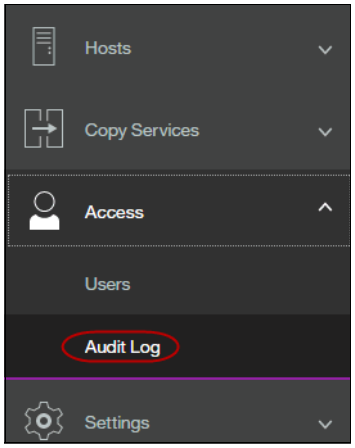

Figure 6-32 Audit Log from Access menu

Figure 6-33 shows an example of the audit log after creating a FlashCopy volume, with a command highlighted. The **Running Tasks** button is available at the top of the window in the status pane. If you click that button, the progress of the currently running tasks can be displayed by clicking the associated **View** button.

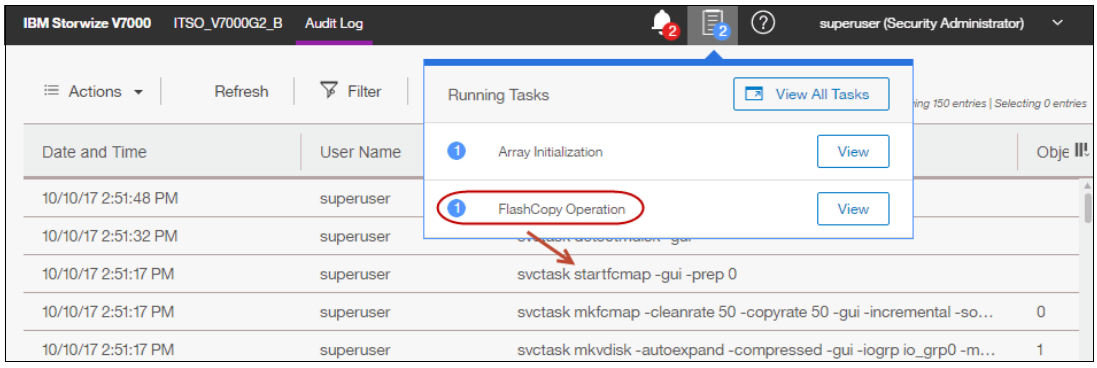

Figure 6-33 Audit log

Changing the view of the Audit Log grid is also possible by right-clicking column headings (Figure 6-34). The grid layout and sorting is completely under the user's control, so you can view everything in the audit log, sort different columns, and reset the default grid preferences.

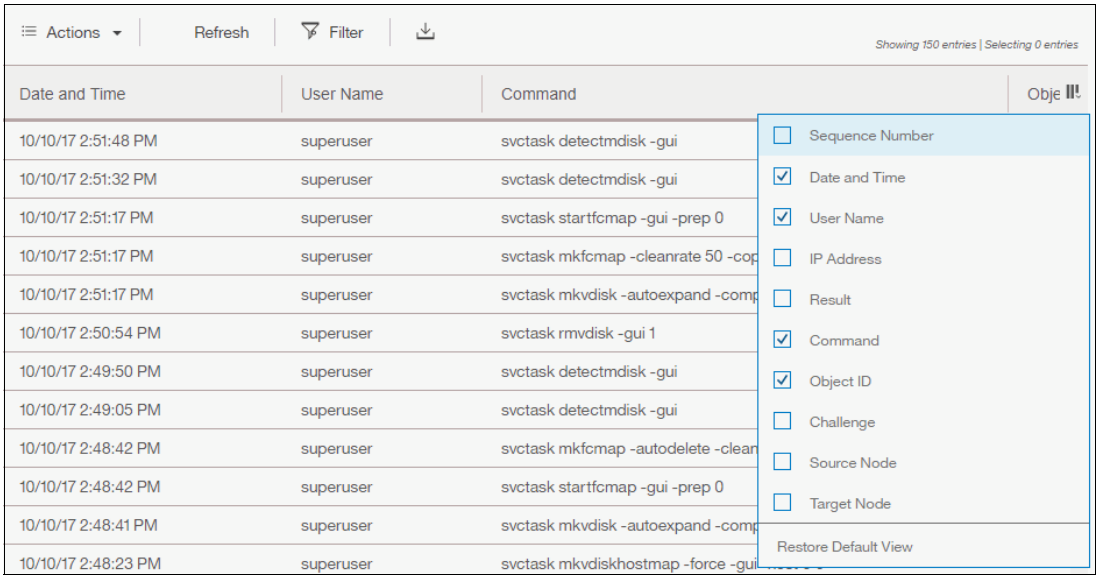

Figure 6-34 Right-click audit log column headings

6.5 Collecting support information using the GUI and the CLI

Occasionally, if you have a problem and call the IBM Support Center, they might ask you to provide a support package. You can collect and upload this package from the **Settings** → **Support** menu.

6.5.1 Collecting information using the GUI

To collect information using the GUI, complete the following steps:

1. Click **Settings** →**Support** and the Support Package tab (Figure 6-35).
2. Click the **Upload Support Package** button.

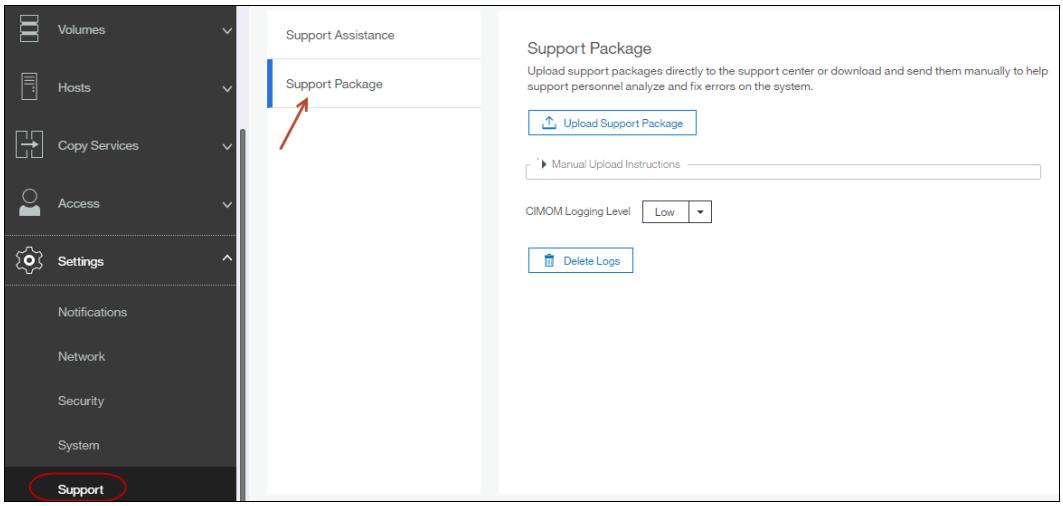

Figure 6-35 Support Package option

Assuming that the problem encountered was an unexpected node restart that has logged a 2030 error, we collect the default logs plus the most recent `statesave` from each node to capture the most relevant data for support.

Note: When a node unexpectedly reboots, it first dumps its current `statesave` information before it restarts to recover from an error condition. This `statesave` is critical for support to analyze what happened. Collecting a snap type 4 creates new `statesaves` at the time of the collection, which is not useful for understanding the restart event.

The Upload Support Package window provides four options for data collection. If you have been contacted by IBM Support due to your system calling home or you have manually opened a call with IBM Support, you will have been given a *PMR number*.

3. Enter that PMR number into the **PMR** field and select the snap type, often referred to as an *option 1, 2, 3, 4 snap*, as requested by IBM Support (Figure 6-36). In our case, we enter our PMR number, select snap type 3 (option 3) because this will automatically collect the `statesave` created at the time the node restarted, and click **Upload**.

Tip: You can use `https://www.ibm.com/support/servicerequest` to open a service request online.

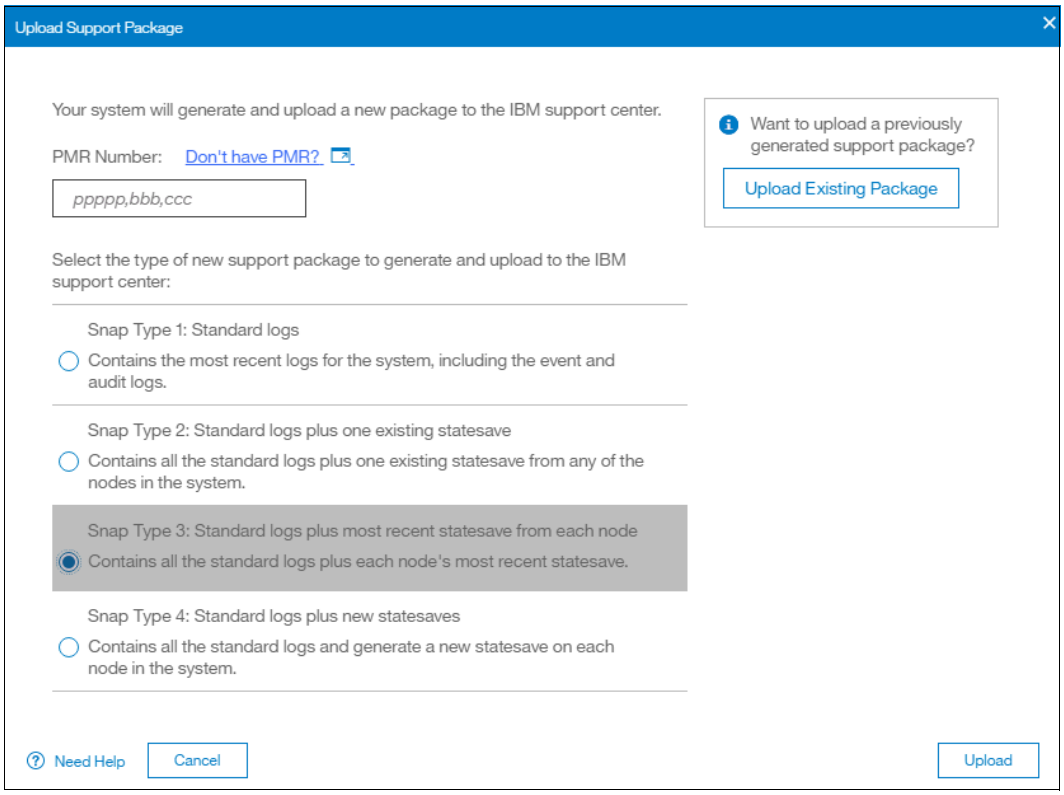

Figure 6-36 Upload Support Package window

4. The procedure to generate the snap on a IBM Spectrum Virtualize storage system, including the most recent statesave from each node canister, starts. This process might take a few minutes (Figure 6-37).

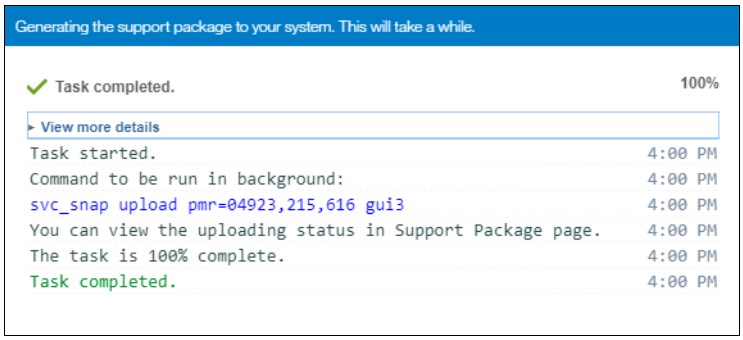

Figure 6-37 Task detail window

6.5.2 Collecting logs using the CLI

The CLI can be used to collect and upload a support package as requested by IBM Support by performing the following steps:

1. Log in to the CLI and issue the **svc_snap** command that matches the type of snap requested by IBM Support:
   - Standard logs (type 1):

     ```
     svc_snap upload pmr=ppppp,bbb,ccc gui1
     ```

   - Standard logs plus one existing statesave (type 2):

     ```
     svc_snap upload pmr=ppppp,bbb,ccc gui2
     ```

   - Standard logs plus most recent statesave from each node (type 3):

     ```
     svc_snap upload pmr=ppppp,bbb,ccc gui3
     ```

   - Standard logs plus new statesaves:

     ```
     svc_livedump -nodes all -yes
     svc_snap upload pmr=ppppp,bbb,ccc gui3
     ```

2. We collect the type 3 (option 3) and have it automatically upload to the PMR number provided by IBM Support, as shown in Example 6-33.

Example 6-33 The svc_snap command

```
ssh superuser@10.18.228.71
Password:
IBM_Storwize:V7000G2:superuser>svc_snap upload pmr=04923,215,616 gui3
```

3. If you do not want to automatically upload the snap to IBM, do not specify the **upload pmr=ppppp,bbb,ccc** part of the commands. When the snap creation completes, it creates a file named using this format:

   /dumps/snap.*<panel_id>*.*YYMMDD*.*hhmmss*.tgz

   It takes a few minutes for the snap file to complete, and longer if including statesaves.

4. The generated file can then be retrieved from the GUI under **Settings** →**Support** → **Manual Upload Instructions** twisty →**Download Support Package** and then click **Download Existing Package**, as shown in Figure 6-38.

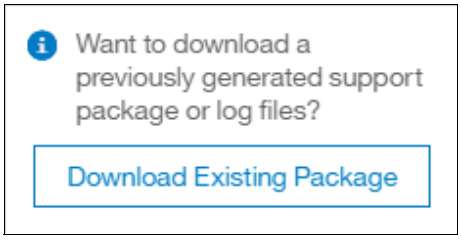

Figure 6-38 Downloaded Existing Package

5. Click in the **Filter** box and enter `snap` to see a list of snap files, as shown in Figure 6-39. Locate the exact name of the snap generated by the **`svc_snap`** command issued earlier, select that file, and click **Download**.

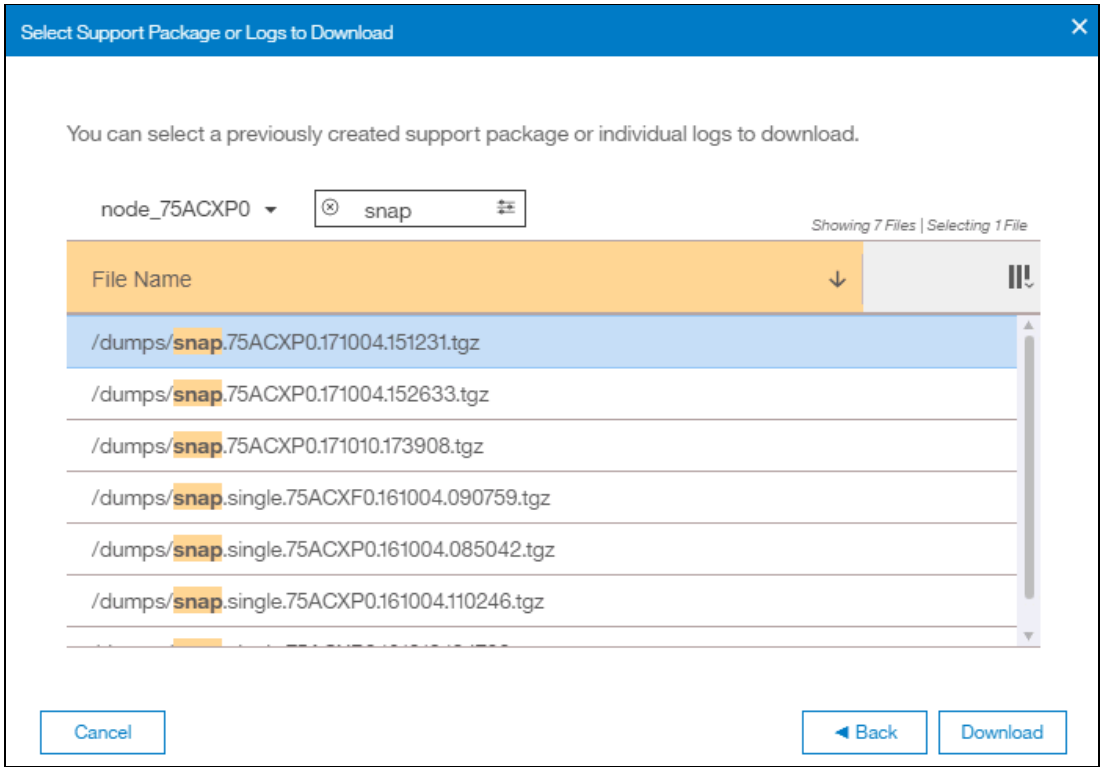

Figure 6-39 Filtering on snap to download

6. Save the file to a folder of your choice on your workstation.

6.5.3 Uploading files to the Support Center

If you have chosen to not have the IBM Spectrum Virtualize storage system upload the support package automatically, it can still be uploaded for analysis from the Enhanced Customer Data Repository (ECuRep). Any uploads should be associated with a specific problem management report (PMR). The PMR is also known as a *service request* and is a mandatory requirement when uploading.

To upload information, use the following procedure:

1. Using a web browser, navigate to ECuRep:

   https://www.secure.ecurep.ibm.com/app/upload

   This link takes you to the Secure Upload page (Figure 6-40).

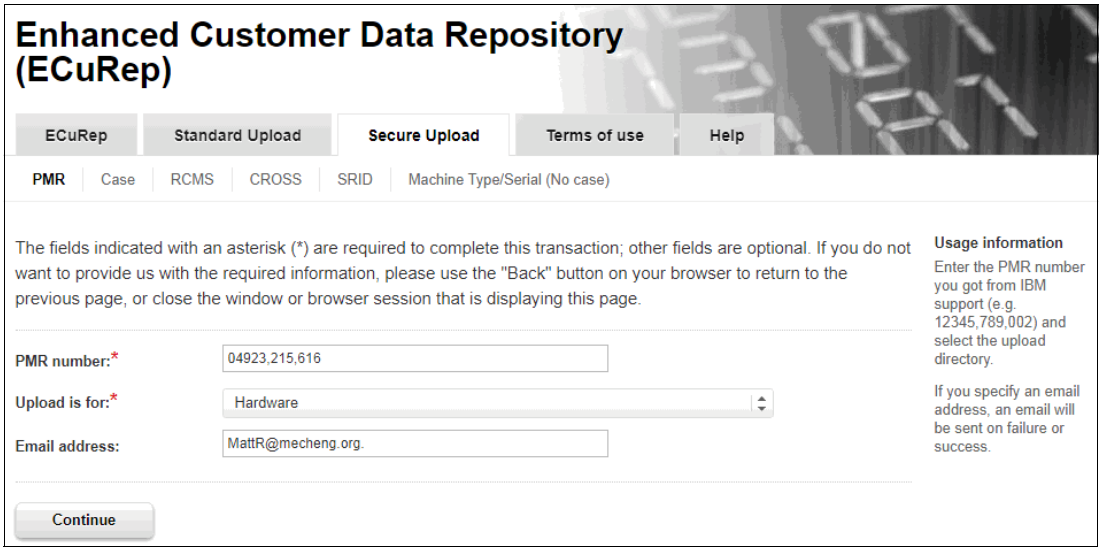

Figure 6-40 ECuRep details

2. Complete the required fields:

   – **PMR number** (mandatory) as provided by IBM support for your specific case. This number should be in the format of `ppppp,bbb,ccc`, for example, 04923,215,616, using a comma (,) as a separator.

   – **Upload is for** (mandatory). Select **Hardware** from the drop-down menu.

   – **Email address** (not mandatory). We advise you to input your email address in this field to be automatically notified of a successful or unsuccessful upload.

3. When the form is completed, click **Continue** to open the input window (Figure 6-41).

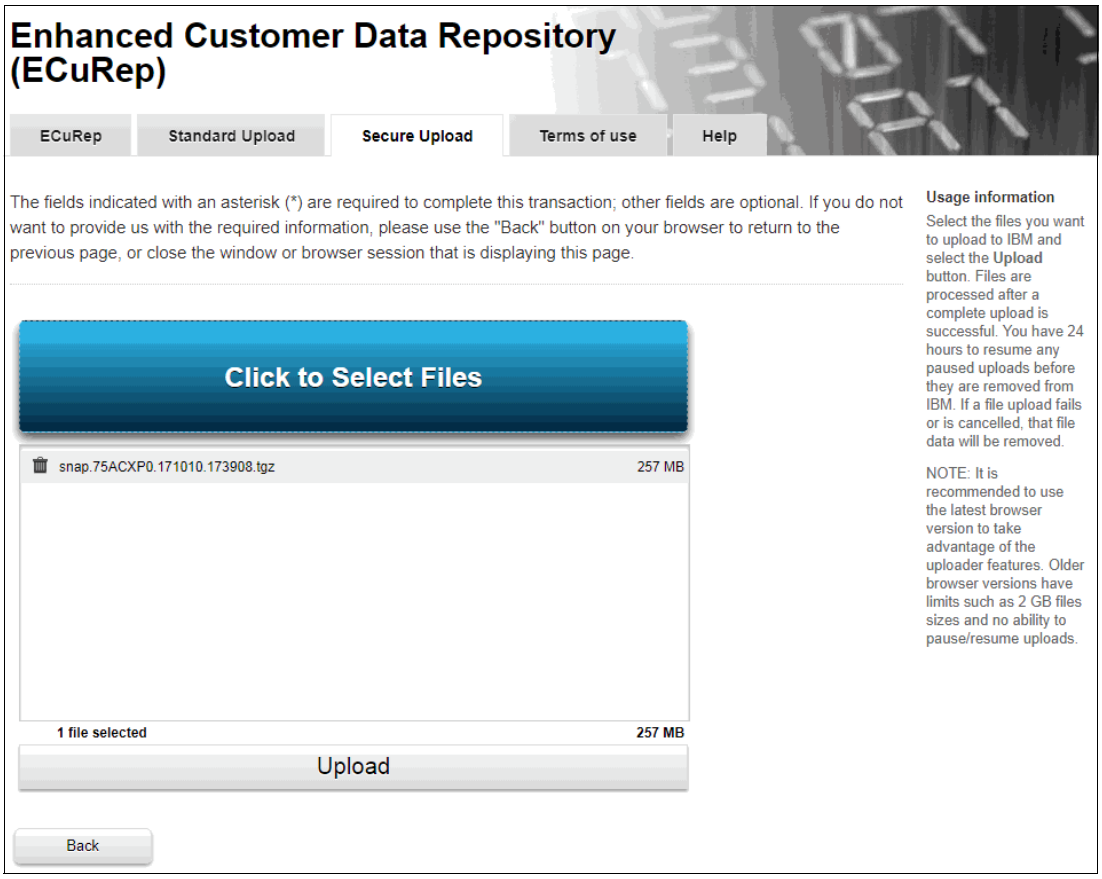

Figure 6-41 ECuRep File upload

4. Select one or more files, click **Upload** to continue, and follow the directions.

6.6 Error codes

This section lists the error codes related to Data Reduction Pools. For more information about troubleshooting options and error codes, see the Troubleshooting section in IBM Knowledge Center:

https://ibm.biz/Bdjmgz

6.6.1 Error code CMMVC7210E

Example 6-34 CMMVC7210E

```
CMMVC7210E The command failed because it would create too many compressed volume
copies that are contained within regular pools in the I/O group.
```

There is a limitation depending on the hardware platform type on the number of compressed copies within regular pools in the I/O group. The limitation is either 200 or 512 compressed copies that are within regular pools within the I/O group, depending on the platform type. There is no such limitation for compressed copies in Data Reduction Pools.

The suggested approaches are to delete compressed copies that are within regular pools, create compressed copies in Data Reduction Pools, or create the copies in other I/O groups.

6.6.2 Error code CMMVC8696E

Example 6-35 CMMVC8696E

```
CMMVC8696E The command failed because the volume or a volume copy is in the
deleting state.
```

An action was requested on a volume or a volume copy that was in the process of being deleted. Make sure that you specified the correct volume. If so, the requested action is not supported on volumes or volume copies which are deleting.

6.6.3 Error code CMMVC9173E

Example 6-36 CMMVC9173E

```
CMMVC9173E The pool specified is a data reduction pool. Thin provisioned or
compressed volume copies created from a data reduction pool must use -autoexpand.
```

A volume copy created from a Data Reduction Pool must have **autoexpand** on. The command fails when a thin provisioned or compressed volume copy is created from a Data Reduction Pool without specifying the **-autoexpand** parameter.

6.6.4 Error code CMMVC9175E

Example 6-37 CMMVC9175E

```
CMMVC9175E The pool specified is a data reduction pool. Volumes or volume copies
which are thin provisioned or compressed and created from a data reduction pool
can not be striped specifying an mdisk, sequential or image mode.
```

A volume or volume copy created from a Data Reduction Pool must not be striped specifying an MDisk, sequential, or image mode. The command fails when a thin provisioned or compressed volume copy is created from a Data Reduction Pool and these parameters are specified.

6.6.5 Error code CMMVC9176E

Example 6-38 CMMVC9176E

```
CMMVC9176E The pool specified is a data reduction pool. For volumes or volume
copies which are thin provisioned or compressed and created from a data reduction
pool, the volume cannot have a cache mode of none or readonly.
```

A volume or volume copy created from a Data Reduction Pool must have a cache mode of readwrite. The command fails when a thin provisioned or compressed volume copy is created from a Data Reduction Pool and the volume cache mode is not `readwrite`.

6.6.6 Error code CMMVC9235E

Example 6-39 CMMVC9235E

```
CMMVC9235E The pool specified is a data reduction pool. Volumes or volume copies
which are thin provisioned and created from a data reduction pool can not use the
-grainsize parameter.
```

A volume or volume copy created from a Data Reduction Pool must not use the **-grainsize** parameter or the command fails. This type of volume or volume copy is created with a grainsize of 8 KB.

6.6.7 Error code CMMVC9236E

Example 6-40 CMMVC9236E

```
CMMVC9236E The pool specified is a data reduction pool. Volumes or volume copies
which are thin provisioned and created from a data reduction pool can not use the
-warning parameter.
```

A volume or volume copy created from a Data Reduction Pool must not use the **-warning** parameter. This type of volume or volume copy is created without a warning threshold.

6.6.8 Error code CMMVC9247E

Example 6-41 CMMVC9247E

```
CMMVC9247E The pool specified is a data reduction pool. Thin or compressed volumes
or volume copies in data reduction pools cannot have the easytier status set or
changed.
```

A volume or volume copy created from a Data Reduction Pool cannot specify the easytier mode of the volume. The `easytier` mode is an attribute of the pool and must be set or changed at the pool level.

6.6.9 Error code CMMVC9245E

Example 6-42 CMMVC9245E

```
CMMVC9245E The pool specified is a data reduction pool and the IO group selected
already contains compressed volumes within a regular pool. An IO group using
software compression can not contain compressed volumes from regular and data
reduction pools at the same time.
```

The I/O group selected supports only software compression and cannot contain compressed volumes or volume copies from regular and Data Reduction Pools at the same time. Select a different I/O group or create the compressed volume or volume copy in a regular pool.

6.6.10 Error code CMMVC9246E

Example 6-43 CMMVC9246E

```
CMMVC9246E The pool specified is a regular pool and the IO group selected already
contains compressed volumes within a data reduction pool. An IO group using
```

```
software compression can not contain compressed volumes from regular and data
reduction pools at the same time.
```

The I/O group selected supports only software compression and cannot contain compressed volumes or volume copies from regular and Data Reduction Pools at the same time. Select a different I/O group or create the compressed volume or volume copy in a Data Reduction Pool.

6.6.11 Error code CMMVC9177E

Example 6-44 CMMVC9177E

```
CMMVC9177E The pool specified is a data reduction pool. Thin provisioned or
compressed volume copies created from a data reduction pool can not use
-noautoexpand.
```

A volume copy created from a Data Reduction Pool must have **autoexpand** on. The command fails when a thin provisioned or compressed volume copy from a Data Reduction Pool is created and the **-noautoexpand** parameter is specified.

6.6.12 Error code CMMVC9250E

Example 6-45 CMMVC9250E

```
CMMVC9250E The pool specified is a data reduction pool. Thin provisioned or
compressed volume copies created from a data reduction pool can not use the
-buffersize parameter.
```

A volume copy created from a Data Reduction Pool cannot use the **-buffersize** parameter. The command fails when a thin provisioned or compressed volume copy is created from a Data Reduction Pool and the **-buffersize** parameter is specified.

6.6.13 Error code CMMVC9248E

Example 6-46 CMMVC9248E

```
CMMVC9248E The volume or volume copy specified is a thin or compressed volume in a
data reduction pool. The easytier status cannot be changed.
```

A volume or volume copy created from a Data Reduction Pool cannot specify the `easytier` mode of the volume. The `easytier` mode is an attribute of the pool and must be set or changed at the pool level.

6.6.14 Error code CMMVC9184E

Example 6-47 CMMVC9184E

```
CMMVC9184E The volume specified is a thin or compressed volume in a data reduction
pool. Autoexpand must be on and can not be changed.
```

Thin or compressed volumes in Data Reduction Pools must have autoexpand on, and this setting cannot be changed.

6.6.15 Error code CMMVC9237E

Example 6-48 CMMVC9237E

```
CMMVC9237E The volume specified is a thin or compressed volume in a data reduction
pool. Cache can not be set to none or readonly.
```

Thin or compressed volumes in Data Reduction Pools must have the cache enabled. This setting cannot be changed.

6.6.16 Error code CMMVC9238E

Example 6-49 CMMVC9238E

```
CMMVC9238E The volume specified is a thin or compressed volume in a data reduction
pool. The -warning parameter can not be used.
```

Thin or compressed volumes in Data Reduction Pools must not have a warning threshold.

6.6.17 Error code CMMVC9185E

Example 6-50 CMMVC9185E

```
CMMVC9185E The volume or volume copy specified is a thin or compressed volume in a
data reduction pool and the action requested is not supported on volumes of this
type.
```

It is not possible to run this command on a thin or compressed volume copy in a Data Reduction Pool. This action can only be performed on fully allocated volumes or thin or compressed volume copies that are in a regular pool.

6.6.18 Error code CMMVC9186E

Example 6-51 CMMVC9186E

```
CMMVC9186E The pool specified is a data reduction pool. It is not possible to
migrate thin or compressed volumes to a data reduction pool.
```

It is not possible to migrate thin or compressed volume copies to a Data Reduction Pool. Use volume mirroring (**addvdiskcopy**) to migrate thin or compressed volumes to a Data Reduction Pool. For more information, see Chapter 2, “Planning” on page 11.

6.6.19 Error code CMMVC9179E

Example 6-52 CMMVC9179E

```
CMMVC9179E The command failed as the resource required for a data reduction pool
could not be allocated.
```

There are not enough resources available. Each Data Reduction Pool requires resources used for managing pools and volumes. Delete some volume copies, volumes, or pools on the system to free enough resources, and then create a Data Reduction Pool.

6.6.20 Error code CMMVC9180E

Example 6-53 CMMVC9180E

```
CMMVC9180E The data reduction pool can not be created as the maximum number of
data reduction pools already exists.
```

There can be four Data Reduction Pools in a cluster. When this limit is reached, no further Data Reduction Pools can be created.

6.6.21 Error code CMMVC9181E

Example 6-54 CMMVC9181E

```
CMMVC9181E Data reduction can only be set to on for a parent pool.
```

It is not allowed to change a child pool to a data reduction child pool.

6.6.22 Error code CMMVC9249E

Example 6-55 CMMVC9249E

```
CMMVC9249E Data reduction pools can not have child pools.
```

A Data Reduction Pool is not allowed to have child pools.

6.6.23 Error code CMMVC9190E

Example 6-56 CMMVC9190E

```
CMMVC9190E The pool specified is a data reduction pool, creating thin or
compressed volumes from this type of pool must have autoexpand on.
```

All thin or compressed volumes in Data Reduction Pools must have autoexpand on, otherwise the command fails.

6.6.24 Error code CMMVC9253E

Example 6-57 CMMVC9253E

```
CMMVC9253E An IO group using software compression can not contain compressed
volumes from regular and data reduction pools at the same time.
```

The I/O group selected supports only software compression and a compressed volume can not be created with copies in both a regular and a Data Reduction Pool. Select a different I/O group or create the compressed volume with both copies in either regular pools or Data Reduction Pools.

6.6.25 Error code CMMVC9242E

Example 6-58 CMMVC9242E

```
CMMVC9242E The volume or volume copy specified is a thin or compressed volume in a
data reduction pool. Thin provisioned or compressed volume copies created from a
data reduction pool can not use the -rsize parameter.
```

A volume copy created from a Data Reduction Pool cannot use the **-rsize** option because Data Reduction Pools automatically manage the physical capacity used across the volumes in the pool. The command fails when the **-rsize** of a thin provisioned or compressed volume copy from a Data Reduction Pool is expanded or shrunk.

6.6.26 Error code CMMVC9241E

Example 6-59 CMMVC9241E

```
CMMVC9241E Action failed because Quorum device is unavailable.
```

The action failed due to issues communicating with the Quorum device. Check the event log and resolve any outstanding issues with the Quorum device.

6.6.27 Error code CMMVC9185E

Example 6-60 CMMVC9185E

```
CMMVC9185E The volume specified is a thin or compressed volume in a data reduction
pool and the action requested is not supported on volumes of this type.
```

It is not possible to run this command on a thin or compressed volume copy in a Data Reduction Pool. Only perform this action on fully allocated volumes, thin or compressed volumes, or volume copies that are in a regular pool.

6.6.28 Error code CMMVC5772E

Example 6-61 CMMVC5772E

```
CMMVC5772E The operation requested could not be performed because software upgrade
is in progress.
```

The operation failed because a software upgrade is in progress. Wait for the software upgrade to complete and then reissue the command.

6.6.29 Error code CMMVC9296E

Example 6-62 CMMVC9296E

```
CMMVC9296E The command has failed because the volume LBA address supplied is not
recoverable and contains a virtual medium error.
```

The operation failed because the data at the supplied volume LBA cannot be recovered, and the volume LBA cannot look up the physical address. The data to the virtual LBA need to be restored from a backup or from a previously generated backup of the volume or the volume needs to be repaired from the host application.

6.6.30 Error code CMMVC9185E

Example 6-63 CMMVC9185E

```
CMMVC9185E The volume specified is a thin or compressed volume in a data reduction
pool and the action requested is not supported on volumes of this type.
```

This command cannot be executed on a thin or compressed volume copy in a Data Reduction Pool. Only perform this action on fully allocated volumes, or thin or compressed volumes or volume copies that are in a regular pool.

6.6.31 Error code CMMVC9277E

Example 6-64 CMMVC9277E

```
CMMVC9277E The command failed because there are out of sync volume mirrors and
some copies are in data reduction pools that may also be deleted.
```

There are volume mirrors that are out of synchronisation and at least one thin or compressed copy is in a Data Reduction Pool. For the volume mirrors where there is at least one copy in a Data Reduction Pool, bring the volume mirrors back into synchronisation before re-attempting the command.

6.6.32 Error code CMMVC9265E

Example 6-65 CMMVC9265E

```
CMMVC9265E The volume specified has a thin or compressed volume copy in a data
reduction pool. Thin provisioned or compressed volumes created from a data
reduction pool can not use -size parameter to reduce the size of the volume.
```

It is not possible to shrink thin provisioned or compressed volumes created from a Data Reduction Pool, when **-size** parameter was used. Select a regular disk or a disk from a non Data Reduction Pool to perform this action.

6.6.33 Error code CMMVC9262E

Example 6-66 CMMVC9262E

```
CMMVC9262E The pool specified is a data reduction pool. The action requested is
not supported with pools of this type.
```

It is not possible to run this command if a Data Reduction Pool is selected. Select a regular pool to perform this action.

6.6.34 Error code CMMVC9252E

Example 6-67 CMMVC9252E

```
CMMVC9252E The command failed because the I/O group contains compressed volumes.
The new hardware configuration does not support compressed volumes.
```

The I/O group this node belongs to contains compressed volumes and the hardware configuration being submitted stops the node from supporting compressed volumes. Revert the change to the hardware configuration or remove all the compressed volumes from the I/O group.

6.6.35 Error code CMMVC9286E

Example 6-68 CMMVC9286E

```
CMMVC9286E The amount of memory configured for FlashCopy cannot be greater than
1.5 gigabytes (GB) because data reduction volumes exist in this I/O group.
```

For I/O groups containing 8 GB node types, the FlashCopy bitmap cannot be increased beyond 1.5 GB if there are thin or compressed volumes created within that I/O group. Upgrade the node types within the I/O group to contain more than 8 GB of memory. Do not use data reduction volumes within that I/O group, set the FlashCopy bitmap size under 1.5 GB, or select another I/O group to use for flash copies if you want to have more than 1.5 GB of flash copy bitmap.

6.6.36 Error code CMMVC9285E

Example 6-69 CMMVC9285E

```
CMMVC9285E A thin or compressed volume in a data reduction pool cannot be created
within this I/O group because the amount of memory configured for FlashCopy is
greater than 1.5 gigabytes (GB) and there is insufficient free memory in the I/O
group.
```

For I/O groups containing 8 GB node types, thin or compressed volumes cannot be created if the FlashCopy bitmap is set beyond 1.5 GB for the I/O group. Upgrade the node types within the I/O group to contain more than 8 GB of memory, do not use data reduction volumes within that I/O group, select another I/O group to create the thin or compressed volumes in a Data Reduction Pool, or reduce the amount of FlashCopy bitmap memory that is in use for the I/O group to below 1.5 GB.

6.6.37 Error code CMMVC9283E

Example 6-70 CMMVC9283E

```
CMMVC9283E The command failed because the pool specified is a data reduction pool
that has insufficient capacity.
```

There is insufficient capacity within the pool specified to be able to create a thin or compressed volume within a Data Reduction Pool. Add extra capacity to the specified storage pool and then re-attempt the command.

6.6.38 Error code CMMVC9284E

Example 6-71 CMMVC9284E

```
CMMVC9284E The command failed because the pool specified is a data reduction pool
that contains offline volumes.
```

Volumes cannot be created in Data Reduction Pools while there are other thin or compressed volumes in Data Reduction Pools. Repair the offline volumes in the Data Reduction Pool, or bring online the MDisks or bring online the offline Data Reduction Pool or add additional capacity to the pool.

6.6.39 Error code CMMVC8710E

Example 6-72 CMMVC8710E

```
CMMVC8710E The command failed because there were not enough extents in storage
pool.
```

There is insufficient capacity within the pool specified to be able to create a fully allocated volume within a data reduction or regular storage pool. The user needs to add extra capacity to the specified storage pool and then re-attempt the command.

6.6.40 Error code CMMVC9185E

Example 6-73 CMMVC9185E

```
CMMVC9185E The volume specified is a thin or compressed volume in a data reduction
pool and the action requested is not supported on volumes of this type.
```

This action is not supported on thin or compressed volumes in Data Reduction Pools.

6.6.41 Error code CMMVC9186E

Example 6-74 CMMVC9186E

```
CMMVC9186E The pool specified is a data reduction pool. It is not possible to
migrate thin or compressed volumes to a data reduction pool.
```

It is not possible to migrate thin or compressed volume copies to a Data Reduction Pool. Use volume mirroring (**addvdiskcopy**) to migrate thin or compressed volumes to a Data Reduction Pool.

6.6.42 Error code CMMVC9287E

Example 6-75 CMMVC9287E

```
CMMVC9287E The volume or volume copy cannot be deleted because the data reduction
pool is corrupt.
```

Volumes cannot be deleted from Data Reduction Pools.

6.6.43 Error code CMMVC9288E

Example 6-76 CMMVC9288E

```
CMMVC9288E The node could not be added as a data reduction pool containing a thin
or compressed volume exists within this I/O group and the new node does not meet
the minimum CPU requirements.
```

When the first thin or compressed volume in a Data Reduction Pool is created for an I/O group, the I/O group sets CPU parameters based on the lowest number of CPU resources available based on the nodes in the I/O group. A new node with less CPU resources cannot be added to the I/O group. Delete all thin and compressed volumes in all Data Reduction Pools for the specified I/O group then retry the `addnode`.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

The following IBM Redbooks publications provide additional information about the topic in this document (note that some publications referenced in this list might be available in softcopy only):

- *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933
- *Implementing the IBM Storwize V7000 with IBM Spectrum Virtualize V8.1*, SG24-7938
- *Implementing the IBM Storwize V5000 Gen2 (including the Storwize V5010, V5020, and V5030) with IBM Spectrum Virtualize V8.1*, SG24-8162

You can search for, view, download, or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

The following Redbooks domains related to this book are also useful resources:

- IBM Storage Networking Redbooks

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/san
- IBM Flash Storage Redbooks

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/flash
- IBM Software Defined Storage Redbooks

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/sds
- IBM Disk Storage Redbooks

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/disk
- IBM Storage Solutions Redbooks

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/storagesolutions
- IBM Tape storage Redbooks

  http://www.redbooks.ibm.com/Redbooks.nsf/domains/tape

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Redbooks

Introduction and Implementation of Data Reduction Pools and

(0.2”spine)
0.17”<->0.473”
90<->249 pages

SG24-8430-00

ISBN 0738457310

Printed in U.S.A.